

บทที่ 4

ผลการทดลอง

ผลการศึกษาสร้างแบบจำลองการทำงานของสมองขณะพิมพ์ ด้วยเครื่องมือ และ ประยุกต์ใช้เทคนิคต่าง ๆ ซึ่งงานวิจัยนี้ได้แบ่งการทดลองออกเป็น 3 รูปแบบที่จะทดลองค้นหา คำตอบที่แตกต่างกันของแต่ละส่วนการทดลองที่ข้อมูล และรูปแบบการทดลองแตกต่างกัน ซึ่งการ ทดลองวัดประสิทธิภาพของเครื่องมือสร้างแบบจำลอง และการทดลองที่จะศึกษาข้อมูลเวลาการ พิมพ์ของมนุษย์ที่เปลี่ยนแปลงตามลำดับเวลา ซึ่งจะทดลองวัดประสิทธิภาพการทำงานนายของ แบบจำลองที่สร้างด้วยวิธีการที่กล่าวมาในข้างต้น

ข้อมูลที่แสดงในบทนี้คือ ผลการทดลองจากวิธีการต่าง ๆ ซึ่งได้แสดงในรูปแบบ ได้แก่ ตารางแสดงค่าประสิทธิภาพในการทำงานของแบบจำลองซึ่งจะเปรียบเทียบระหว่างค่าเวลาที่ผู้ใช้ แต่ละคนทำได้ กับค่าเวลาที่แบบจำลองที่สร้างขึ้นด้วยวิธีการต่าง ๆ สามารถทำนายได้ ซึ่งการวัด ประสิทธิภาพการทำงานของแบบจำลองดังกล่าวนี้จะคำนวณหาค่าความผิดพลาดกำลังสอง (Root Means Square Error: RMSE) โดยที่ค่าดังกล่าวแสดงให้เห็นว่าแบบจำลองที่สร้างด้วย วิธีการใดเหมาะสมในการทำนายค่าเวลาการพิมพ์ของผู้ใช้มากกว่ากัน นอกจากนี้ยังได้แสดงกราฟ ให้เห็นถึงแนวโน้มของค่าเวลาการพิมพ์ที่มีความเปลี่ยนแปลงไปตามทักษะที่เปลี่ยนแปลงไปตาม จำนวนครั้งที่พิมพ์หรือไม่ โดยข้อมูลเวลาการพิมพ์มาสร้างกราฟการเรียนรู้ ซึ่งงานวิจัยนี้ได้ ยกตัวอย่างกลุ่มข้อมูลที่ทดลอง 1 คู่อักษรคือ คู่อักษร “ฟว” เป็นตัวอย่างจำนวน 73 ตัว จากกลุ่ม ข้อมูลที่ใช้การทดลองทั้งหมด 331 คู่อักษร ซึ่งผลการทดลองที่ได้แสดงให้เห็นดังต่อไปนี้

4.1 ผลการทดลองการสร้างแบบจำลองแบบที่ 1

ผลการทดลองของแบบจำลองที่ข้อมูลอินพุตไม่ได้ผ่านการจัดการข้อมูลด้วยวิธีการใด ๆ เลยคือ ไม่มีการจัดการข้อมูลอินพุตด้วยเทคนิคหน้าต่างเลื่อน และการแบ่งกลุ่มด้วยเทคนิคการ แบ่งกลุ่มข้อมูลทั้ง 2 เทคนิคที่งานวิจัยนี้เลือกใช้ในการสร้างแบบจำลองของผู้ใช้แต่ละคนด้วย เครื่องมือทั้ง 3 ที่งานวิจัยนี้ได้เลือกใช้ไม่ว่าจะเป็น การถดถอยเชิงเส้น โครงข่ายประสาทเทียม และ แบบจำลองต้นไม้เอนโทรปี ผลที่ได้จากการวัดประสิทธิภาพการทำงานของแบบจำลองที่ถูกสร้าง ด้วยวิธีการที่กล่าวมาในข้างต้นแสดงให้เห็นในตารางที่ 4.1

ตารางที่ 4.1

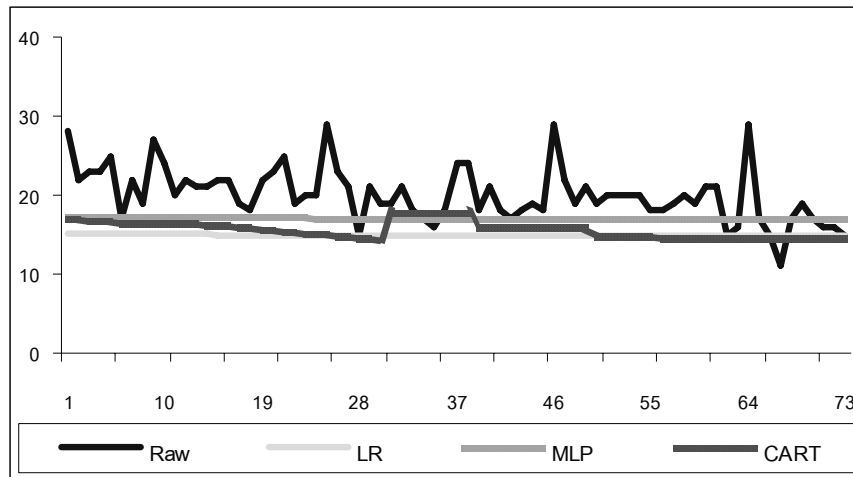
ค่าความผิดพลาดกำลังสองที่ทำนายด้วยแบบจำลองแบบที่ 1

Fold	Non Cluster		
	LR	MLP	M5P
1	3.924	3.720	3.249
2	3.936	4.227	3.246
3	3.905	3.553	3.252
4	4.002	3.913	3.367
5	3.971	4.166	3.334
6	4.012	4.031	3.371
7	4.020	3.763	3.336
8	3.961	3.776	3.295
9	3.913	3.906	3.228
10	4.043	4.462	3.408
ค่าเฉลี่ย	3.969	3.952	3.309

จากตารางที่ 4.1 ค่าประสิทธิภาพในการทำงานของแบบจำลองที่สร้างด้วยเครื่องมือต่าง ๆ ซึ่งค่าดังกล่าวเป็นค่าความผิดพลาดที่แบบจำลองทำนายเวลาการพิมพ์ที่ได้ใกล้เคียงกับค่าเวลาจริงที่ผู้ใช้สามารถทำได้มากที่สุดเพียงใด ซึ่งงานวิจัยนี้ได้ทดลองทั้งหมด 10 ครั้งตามวิธีการไขว้ข้าม 10 กลุ่ม จะเห็นได้ว่าค่าเฉลี่ยความผิดพลาดกำลังสองของค่าเวลาที่ทำนายด้วยแบบจำลองต้นไม้เอ็มไพร์ฟี่คือ 3.309 เซนติวินาที ซึ่งเป็นค่าเวลาที่ใกล้เคียงกับค่าเวลาที่ผู้ใช้สามารถทำได้มากกว่าแบบจำลองที่สร้างขึ้นด้วย การถดถอยเชิงเส้น และโครงข่ายประสาทเทียมที่ได้ค่าเท่ากับ 3.969 เซนติวินาที และ 3.952 เซนติวินาที ตามลำดับ นอกจากนี้ได้ใช้เครื่องมือทางสถิติคือ การทดสอบค่าทีแบบหางเดียว (T-Test One Tail) ที่ใช้จับคู่ทดสอบความแตกต่างของเครื่องมือที่นำมาใช้สร้างแบบจำลองการพิมพ์พบว่าผลการทำนายเวลาการพิมพ์ที่ได้จากแบบจำลองต้นไม้เอ็มไพร์ฟี่เปรียบเทียบกับ โครงข่ายประสาทเทียม และการถดถอยเชิงเส้นสามารถทำนายได้ใกล้เคียงกับค่าเวลาที่ผู้ใช้มากกว่าที่ระดับนัยสำคัญ 0.05

ภาพที่ 4.1

เปรียบเทียบค่าเวลาการพิมพ์ของผู้ใช้กับแบบจำลองแบบที่ 1



จากภาพที่ 4.1 แสดงให้เห็นถึงค่าเวลาที่ใช้ทำได้ กับค่าเวลาที่แบบจำลองทั้ง 3 ทำได้ ซึ่งผลการทำนายของ การถดถอยเชิงเส้น และโครงข่ายประสาทเทียม ค่าเวลาที่ได้จากการทำนาย เป็นไปในลักษณะที่เป็นเส้นตรงคือ ค่าที่มีความคงที่แต่มีแนวโน้มในการเปลี่ยนแปลงลดลงตามลำดับ แต่ในแบบจำลองต้นไม้เอ็มไพร์พีจะเห็นได้ว่าค่าของข้อมูลที่แสดงในกราฟนั้นมีความเปลี่ยนแปลงเป็นช่วง ๆ ที่พิมพ์ถึงจะไม่ได้เปลี่ยนไปในทุกครั้งที่พิมพ์ แต่เปลี่ยนในแต่ละช่วงของชุดค่าที่เปลี่ยนไป ซึ่งส่งผลให้ค่าประสิทธิภาพในการทำนายที่ดีกว่าคือ สามารถทำนายใกล้เคียงกับค่าเวลาที่ใช้ทำได้มากกว่า

4.2 ผลการทดลองการสร้างแบบจำลองแบบที่ 2

เนื่องด้วยงานวิจัยนี้ได้ศึกษาสร้างแบบจำลองที่สามารถทำนายเวลาที่ผู้ใช้แต่ละคนพิมพ์คู่อักขระ เมื่อจำนวนครั้งของการพิมพ์คู่อักขระ และทักษะการพิมพ์เปลี่ยนแปลงไป การทดลองข้อมูลที่มีการเปลี่ยนแปลงตามลำดับเวลา ซึ่งงานวิจัยนี้เลือกใช้การแบ่งกลุ่มของข้อมูล อินพุตมีเทคนิคในการแบ่งกลุ่มข้อมูล 2 เทคนิคได้แก่ การแบ่งกลุ่มข้อมูลแยกตามแถวของคู่อักขระ และค่าเฉลี่ยเคกลุ่ม

4.2.1 ผลการทดลองของแบบจำลองที่แบ่งกลุ่มข้อมูลแยกตามแถวของคู่ อักขระ (Row-Based)

การทดลองที่การศึกษานี้ได้เพิ่มเทคนิคในการจัดการข้อมูลก่อนที่จะนำข้อมูลนั้นเข้าไปวิเคราะห์ด้วยเครื่องมือทั้ง 3 ที่การศึกษานี้ได้เลือกใช้คือ การถดถอยเชิงเส้น โครงข่ายประสาทเทียม และแบบจำลองต้นไม้เอนโทรปี ซึ่งในการทดลองในส่วนนี้จะมีการจัดการข้อมูลเบื้องต้นก่อนที่จะวิเคราะห์สร้างแบบจำลอง โดยที่การทดลองในส่วนนี้ใช้เครื่องมือในการแบ่งกลุ่มได้แก่ การแบ่งกลุ่มข้อมูลแยกตามแถวของคู่อักขระ (Row-Based Clustering) และผลการทดลองที่ได้จากการแบ่งกลุ่มข้อมูลด้วยเทคนิคข้างต้นแสดงให้เห็นดังต่อไปนี้

ตารางที่ 4.2

ค่าความผิดพลาดกำลังสองที่ทำนายด้วยแบบจำลองแบบที่ 2 (Row-Based)

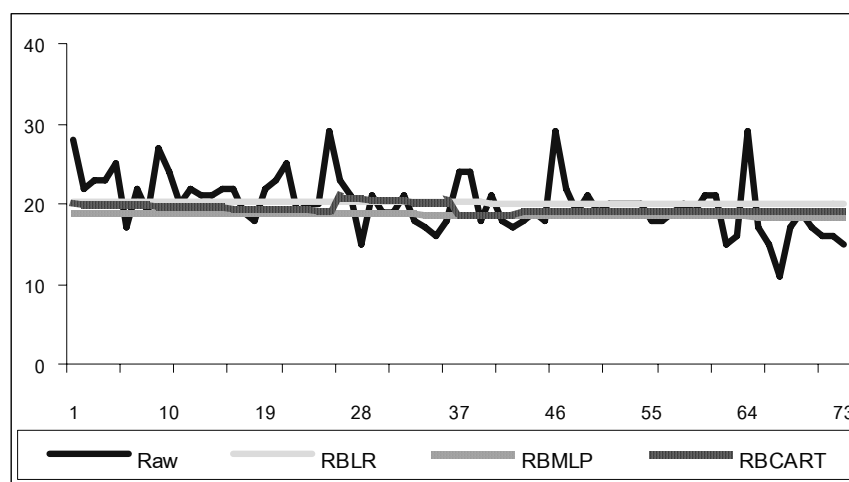
Fold	Row-Based Cluster		
	LR	MLP	M5P
1	3.805	3.969	3.256
2	3.939	3.928	3.355
3	3.819	3.798	3.214
4	3.929	3.958	3.354
5	3.905	3.862	3.290
6	3.798	3.827	3.204
7	3.835	3.655	3.300
8	3.900	3.747	3.316
9	3.842	3.755	3.316
10	3.812	3.770	3.264
ค่าเฉลี่ย	3.858	3.827	3.287

จากค่าประสิทธิภาพในการทำงานของแบบจำลองที่สร้างด้วยวิธีการต่าง ๆ แสดงให้เห็นในตารางที่ 4.2 ซึ่งค่าดังกล่าวเป็นค่าความผิดพลาดที่แบบจำลองทำนายเวลาการพิมพ์ได้ใกล้เคียงกับค่าเวลาจริงที่ผู้ใช้สามารถทำได้มากน้อยเพียงใด ซึ่งงานวิจัยนี้ได้ทดลองทั้งหมด 10 ครั้งตามวิธีการไขว้ข้าม 10 กลุ่ม นอกจากนี้แบบจำลองยังสร้างมาด้วยเครื่องมือที่แตกต่างกัน การ

ทดลองนี้แสดงให้เห็นถึงประสิทธิภาพในการเพิ่มเทคนิคจัดการกลุ่มข้อมูลอินพุตก่อนที่จะนำข้อมูลดังกล่าวนั้นช่วยให้แบบจำลองที่สร้างขึ้นด้วยเครื่องมือต่าง ๆ ให้ผลดีขึ้น จากการที่ได้แบ่งกลุ่มของคู่อักษรที่มีแถวเหมือนกัน และวิเคราะห์ข้อมูลเป็นกลุ่มย่อยช่วยเพิ่มประสิทธิภาพได้ ซึ่งจากค่าเฉลี่ยความผิดพลาดกำลังสองในตารางแสดงให้เห็นว่าการแบ่งกลุ่มค่าของข้อมูลตามแถวนั้นส่งผลให้ประสิทธิภาพในการทำนายดีขึ้นกว่าการที่ข้อมูลอินพุตไม่ได้ผ่านขั้นตอนในการแบ่งกลุ่มข้อมูล จะเห็นได้ว่าค่าเฉลี่ยความผิดพลาดกำลังสองของค่าเวลาที่ทำนายด้วยแบบจำลองต้นไม้เอนโอฟีฟคือ 3.287 เซนติวินาที ซึ่งเป็นค่าเวลาที่ใกล้เคียงกับค่าเวลาที่ผู้ใช้สามารถทำได้มากกว่าแบบจำลองที่สร้างขึ้นด้วย โครงข่ายประสาทเทียม และการถดถอยเชิงเส้นที่ได้ค่าเท่ากับ 3.858 เซนติวินาที และ 3.827 เซนติวินาที ตามลำดับ นอกจากนี้ได้ใช้เครื่องมือทางสถิติคือ การทดสอบค่าที่แบบหางเดี่ยวที่ใช้จับคู่ทดสอบความแตกต่างของเครื่องมือที่นำมาใช้สร้างแบบจำลองการพิมพ์พบว่าผลการทำนายเวลาการพิมพ์ที่ได้จากแบบจำลองต้นไม้เอนโอฟีฟเปรียบเทียบกับโครงข่ายประสาทเทียม และการถดถอยเชิงเส้นสามารถทำนายได้ใกล้เคียงกับค่าเวลาที่ผู้ใช้มากกว่าที่ระดับนัยสำคัญ 0.05

ภาพที่ 4.2

เปรียบเทียบค่าเวลาการพิมพ์ของผู้ใช้กับแบบจำลองแบบที่ 2 (Row-Based)



จากภาพที่ 4.2 แสดงให้เห็นถึงค่าเวลาที่ผู้ใช้ทำได้ กับค่าเวลาที่แบบจำลองทั้ง 3 ทำได้ ซึ่งการทดลองนี้ได้เพิ่มเทคนิคการแบ่งกลุ่มข้อมูลแยกตามแถวของคู่อักษรช่วยเพิ่มประสิทธิภาพการทำนายผลข้อมูลของแบบจำลองให้ดีขึ้น ค่าเวลาการพิมพ์ของแบบจำลองต้นไม้เอนโอฟีฟมี

การเปลี่ยนแปลงไปในทิศทางที่ใกล้เคียงกับข้อมูลเวลาของผู้ใช้มากขึ้น แต่ค่าข้อมูลเวลาของการ ถดถอยเชิงเส้น และโครงข่ายประสาทเทียม ค่าเวลาที่ได้จากการทำนายเป็นไปในลักษณะที่เป็น เส้นตรงคือ เป็นค่าที่มีความคงที่แต่มีแนวโน้มในการเปลี่ยนแปลงที่ลดลงตามลำดับ แต่ใน แบบจำลองต้นไม้เอนโทรปีจะเห็นได้ว่าค่าของข้อมูลที่แสดงในกราฟนั้นมีความเปลี่ยนแปลงเป็น ช่วง ๆ ที่พื้มพีถึงจะไม่ได้เปลี่ยนไปในทุกครั้งที่พื้มพี แต่เปลี่ยนในแต่ละช่วงของชุดค่าที่เปลี่ยนไป ซึ่งส่งผลให้ค่าประสิทธิภาพในการทำนายที่ดีกว่าคือ สามารถทำนายใกล้เคียงกับค่าเวลาที่ผู้ใช้ทำ ได้มากกว่า

4.2.2 ผลการทดลองของแบบจำลองที่แบ่งกลุ่มข้อมูลด้วยค่าเฉลี่ยเคกลุ่ม (K-Means)

การทดลองในส่วนที่ข้อมูลอินพุตผ่านการจัดการก่อนด้วยเทคนิคการแบ่งกลุ่มข้อมูลที่ ช่วยเพิ่มประสิทธิภาพในการทำนายของแบบจำลองที่สร้างขึ้นด้วยเครื่องมือต่าง ๆ ที่การศึกษาครั้งนี้เลือกใช้เทคนิคดังกล่าวจะแบ่งข้อมูลที่มีความเหมือนกันผ่านการคำนวณด้วยสมการวัดค่าความ ห่างกันของข้อมูลแต่ละตัวซึ่งการศึกษานี้ได้กำหนดกลุ่มของข้อมูลที่น่ามาทดลองแบ่งออกเป็น 10 กลุ่ม เมื่อแบ่งกลุ่มด้วยเทคนิคค่าเฉลี่ยเคกลุ่มแล้วจึงนำค่าข้อมูลที่ผ่านมาผ่านการแบ่งกลุ่มดังกล่าวมา สร้างแบบจำลองด้วยเครื่องมือที่เลือกใช้ ได้แก่ การถดถอยเชิงเส้น โครงข่ายประสาทเทียม และ แบบจำลองต้นไม้เอนโทรปี หลังจากนั้นจึงนำค่าที่ได้จากการทำนายค่าข้อมูลกลุ่มตัวอย่างที่ได้มา คำนวณหาค่าความผิดพลาดกำลังสองของแต่ละเครื่องมือที่ได้เลือกมานั้นค่าความผิดพลาด ดังกล่าวได้แสดงให้เห็นในตารางที่ 4.3

ตารางที่ 4.3

ค่าความผิดพลาดกำลังสองที่ทำนายด้วยแบบจำลองแบบที่ 2 (K-Means)

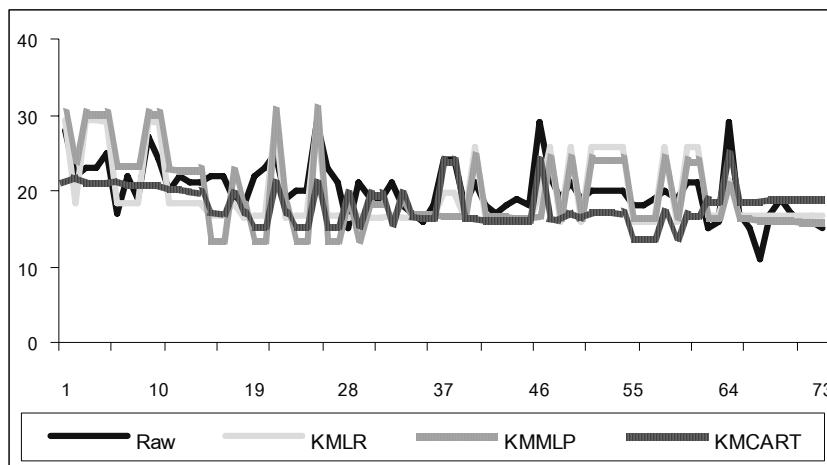
Fold	K-Means Cluster		
	LR	MLP	M5P
1	3.569	3.609	3.233
2	3.538	3.603	3.154
3	3.488	3.636	3.089
4	3.460	3.615	3.068
5	3.457	3.485	3.095
6	3.506	3.727	3.155
7	3.566	3.566	3.163
8	3.391	3.435	3.001
9	3.479	3.810	3.141
10	3.557	3.716	3.252
ค่าเฉลี่ย	3.501	3.620	3.135

จากค่าประสิทธิภาพในการทำงานของแบบจำลองที่สร้างด้วยวิธีการต่าง ๆ แสดงให้เห็นในตารางที่ 4.3 ซึ่งค่าดังกล่าวเป็นค่าความผิดพลาดที่แบบจำลองทำนายเวลาการพิมพ์ได้ใกล้เคียงกับค่าเวลาจริงที่ผู้ใช้สามารถทำได้มากน้อยเพียงใด ซึ่งงานวิจัยนี้ได้ทดลองทั้งหมด 10 ครั้งตามวิธีการไขว้ข้าม 10 กลุ่ม นอกจากนี้แบบจำลองยังสร้างมาด้วยเครื่องมือที่แตกต่างกัน การทดลองนี้แสดงให้เห็นถึงประสิทธิภาพในการเพิ่มเทคนิคจัดการกลุ่มข้อมูลอินพุตก่อนที่จะนำข้อมูลดังกล่าวมาช่วยในการแบ่งกลุ่มที่สร้างขึ้นด้วยเครื่องมือต่าง ๆ ให้ผลดีขึ้น จากการที่ได้แบ่งกลุ่มข้อมูลด้วยค่าเฉลี่ยแต่ละกลุ่ม และวิเคราะห์ข้อมูลเป็นกลุ่มย่อยช่วยเพิ่มประสิทธิภาพในการทำนายได้ นอกจากนี้การแบ่งกลุ่มข้อมูลเข้าไปจะเห็นได้ว่าค่าความผิดพลาดกำลังสองของทั้ง 3 แบบจำลองนั้นได้ค่าที่ลดลงไปในทิศทางเดียวกัน และมีย่านน้อยกว่าวิธีการทดลองที่กล่าวมาในช่วงต้น ซึ่งสามารถทำนายค่าระยะเวลาการพิมพ์ของผู้ใช้ได้ใกล้เคียงกับค่าเวลาพิมพ์ที่แท้จริงที่ผู้ใช้สามารถทำได้ จะเห็นได้ว่าค่าเฉลี่ยความผิดพลาดกำลังสองของค่าเวลาที่ทำนายด้วยแบบจำลองต้นไม้เอนิมไฟว์พีคือ 3.135 เซนติวินาที ซึ่งเป็นค่าเวลาที่ใกล้เคียงกับค่าเวลาที่ผู้ใช้สามารถทำได้มากกว่า

แบบจำลองที่สร้างขึ้นด้วยการถดถอยเชิงเส้น และโครงข่ายประสาทเทียมที่ได้ค่าเท่ากับ 3.620 เซนติวินาที และ 3.501 เซนติวินาที ตามลำดับ นอกจากนี้ได้ใช้เครื่องมือทางสถิติคือ การทดสอบค่าที่แบบหางเดียวที่ใช้จับคู่ทดสอบความแตกต่างของเครื่องมือที่นำมาใช้สร้างแบบจำลองการพิมพ์พบว่าผลการทำนายเวลาการพิมพ์ที่ได้จากแบบจำลองต้นไม้เอ็มไพร์พีเปรียบเทียบกับโครงข่ายประสาทเทียม และการถดถอยเชิงเส้นสามารถทำนายได้ใกล้เคียงกับค่าเวลาที่ผู้ใช้มากกว่าที่ระดับนัยสำคัญ 0.05

ภาพที่ 4.3

เปรียบเทียบค่าเวลาการพิมพ์ของผู้ใช้กับแบบจำลองแบบที่ 2 (K-Means)



จากภาพที่ 4.3 นั้นได้แสดงถึงแนวโน้ม และประสิทธิภาพในการทำนายค่าระยะเวลาการพิมพ์ด้วยทั้ง 3 เครื่องมือว่ามีความใกล้เคียง และเปลี่ยนแปลงไปตามจำนวนครั้งที่พิมพ์ ถึงแม้ว่าค่าประสิทธิภาพการทำนาย และค่าข้อมูลที่เห็นในกราฟของแบบจำลองที่สร้างด้วยโครงข่ายประสาทเทียมจะแย่ง แต่ค่าประสิทธิภาพการทำงานนั้นยังดีมากกว่าการแบ่งกลุ่มแยกตามแถวของคู่อักษร แต่ในการถดถอยเชิงเส้น และแบบจำลองต้นไม้เอ็มไพร์พีให้ผลการทำนายที่ดีขึ้น และมีความเปลี่ยนแปลงในชุดค่าที่เปลี่ยนแปลงไปตามค่าประสบการณ์ สามารถทำนายได้ใกล้เคียงมากขึ้นกว่าการทดลองในรูปแบบอื่น

4.3 ผลการทดลองการสร้างแบบจำลองแบบที่ 3

ข้อมูลที่ถูกนำมาใช้ในการทดลองในส่วนนี้เป็นข้อมูลที่ผ่านการจัดการข้อมูลด้วยเทคนิคหน้าต่างเลื่อน โดยการนำข้อมูลอินพุตเวลามาผ่านเทคนิคการจัดการข้อมูลของเทคนิคหน้าต่างเลื่อนที่จะเป็นเทคนิคที่ปรับให้มีค่าใกล้เคียงกันของข้อมูล เพื่อลดความคลาดเคลื่อนของข้อมูลที่จะเกิดขึ้นจากทุก ๆ ครั้งที่ผู้พิมพ์ตัวคู่อักษร จากกลุ่มข้อมูลที่นำเข้ามาวิเคราะห์สร้างแบบจำลองที่การศึกษาครั้งนี้ได้เลือกใช้นอกจากนี้ยังได้ทดลองออกเป็น 2 ส่วน เช่นเดียวกับการทดลองของข้อมูลที่ไม่ผ่านการจัดการข้อมูลด้วยเทคนิคหน้าต่างเลื่อนคือ การทดลองข้อมูลที่ผ่านมาและไม่ผ่านการแบ่งกลุ่มของค่าคุณลักษณะของข้อมูลด้วยทั้ง 2 เทคนิคการแบ่งกลุ่มข้อมูลได้แก่เทคนิคการแบ่งกลุ่มข้อมูลแบบแยกตามแถว กับเทคนิคการแบ่งกลุ่มด้วยค่าเฉลี่ยเคกลุ่ม

ในส่วนของการทดลองที่เพิ่มเทคนิคการจัดการข้อมูลอินพุตเบื้องต้นอีกรูปแบบหนึ่งเข้าไปนอกจากที่ได้เพิ่มเทคนิคที่ใช้ในการแบ่งกลุ่มข้อมูลซึ่งให้ได้ผลการทดลองที่ดีกว่า และจากที่ข้อมูลเวลาการพิมพ์ที่งานวิจัยนี้ศึกษา เพื่อสร้างแบบจำลองการพิมพ์จากข้อมูลที่อยู่ในรูปแบบดังกล่าว เพื่อสร้างแบบจำลองที่มีความสามารถใกล้เคียง หรือเทียบเท่าได้กับสิ่งที่ผู้ใช้ทำได้มากที่สุด ซึ่งการที่จะพัฒนาความสามารถของแบบจำลองให้สามารถทำนายได้ใกล้เคียงมากขึ้นมานั้นก็ได้มีงานวิจัยที่ได้ศึกษาว่าค่าเวลาการพิมพ์ หรือพฤติกรรมในการพิมพ์นั้นเป็นค่าที่มีความเปลี่ยนแปลงไปตามระยะเวลา และเสนอเทคนิคหน้าต่างเลื่อน เพื่อช่วยจัดการข้อมูลดังกล่าวเพื่อให้ค่าข้อมูลดังกล่าวมีความเสถียรมากยิ่งขึ้นไม่เปลี่ยนแปลงไปในทุก ๆ ครั้งที่พิมพ์ และช่วยกำจัดค่าข้อมูลที่รบกวนต่อข้อมูลจริงที่จะได้มา ซึ่งก็ได้ประยุกต์ใช้เทคนิคดังกล่าวกับเครื่องมือและเทคนิคในการแบ่งกลุ่มข้อมูลทดลองมาในช่วงต้นแล้วผลการทดลองที่ได้แสดงให้เห็นในตารางที่ 4.4

ตารางที่ 4.4

ค่าความผิดพลาดกำลังสองที่ทำนายด้วยแบบจำลองแบบที่ 3

	LR	MLP	M5P
Non Cluster	4.446	4.343	3.725
Row-Based Cluster	4.315	4.392	3.728
K-Means Cluster	3.799	3.990	3.382

จากตารางที่ 4.4 เป็นค่าเฉลี่ยความผิดพลาดของแบบจำลองที่ได้มาจากข้อมูลอินพุตที่ผ่านการจัดการข้อมูลอินพุตค่าประสิทธิภาพในการทำนายที่ได้จากการทดลองสร้างแบบจำลองขึ้นมาไม่ว่าจะเป็นการทดลองด้วยเครื่องมือทั้ง 3 คือ การถดถอยเชิงเส้น โครงข่ายประสาทเทียม และแบบจำลองต้นไม้เอ็มไพร์พี นอกจากการทดลองด้วยเครื่องมือดังกล่าวยังได้ผนวกกับเทคนิคแบ่งกลุ่มข้อมูล 2 แบบคือ แบ่งกลุ่มข้อมูลแยกตามแถวของคู่อักษร และค่าเฉลี่ยเคกลุ่ม ผลการทดลองที่ได้พบว่าค่าประสิทธิภาพในการทำนายที่ได้กลับได้ผลที่แย่กว่าการทำนายด้วยวิธีการทั้งหมดที่ผ่านมาทั้งหมด แสดงให้เห็นว่าวิธีการที่ใช้ในการจัดการข้อมูลแบบหน้าต่างเลือนที่เลือกใช้นี้ไม่เหมาะสมที่จะนำมาใช้ในการทดลองกับข้อมูลที่ใช้ในการศึกษาครั้งนี้

เนื่องจากค่าที่แบบจำลองที่ถูกสร้างขึ้นด้วยกระบวนการต่าง ๆ ที่การศึกษานี้เลือกใช้มีความผิดพลาด และคาดเคลื่อนไปจากค่าเวลาที่ผู้ใช้สามารถทำนายได้มากกว่า ถึงแม้ว่าจะทดลองด้วยค่าเฉลี่ยเคกลุ่ม ในการแบ่งกลุ่มข้อมูลค่าที่แบบจำลองทำให้เกิดความผิดพลาดไปเป็นอย่างมาก ซึ่งความผิดพลาดอาจเกิดขึ้นจากการปรับปรุงเทคนิคจัดการข้อมูลที่ทำให้ข้อมูลอินพุตมีความผิดพลาดทำให้ผลการทำนายของแบบจำลองผิดพลาดไปจากค่าเวลาที่ผู้ใช้แต่ละคนทำได้จริงเป็นอย่างมากกว่าการทดลองในขั้นต้นที่ไม่ผ่านการจัดการข้อมูลอินพุต