

## บทที่ 1

### บทนำ

#### 1.1 ความเป็นมาและความสำคัญของปัญหา

การป้อนข้อมูลภาษาไทยเป็นวิธีการหนึ่งที่ใช้ติดต่อระหว่างผู้ใช้กับเครื่องคอมพิวเตอร์ให้ทำงานตอบสนองต่อความต้องการไม่ว่าจะเป็นการพิมพ์เอกสาร หรืองานด้านการป้อนข้อมูลรูปแบบอื่น ๆ ดังนั้นการพัฒนาทักษะการป้อนข้อมูลจึงเป็นสิ่งจำเป็นในการทำงานร่วมกับคอมพิวเตอร์ เพื่อให้สามารถทำงานได้รวดเร็ว และถูกต้องแม่นยำเพิ่มขึ้น ซึ่งจะส่งผลให้ประสิทธิภาพในการทำงานเพิ่มขึ้น ในการที่จะพัฒนาทักษะดังกล่าวปัญหาที่จะเกิดขึ้นประการหนึ่งคือ ภาษาไทยเป็นภาษาที่มีจำนวนตัวอักขระมากกว่าจำนวนของแป้นอักขระบนแป้นพิมพ์ (Keyboard) โดยทั่วไปจึงส่งผลให้ขนาดของแผงแป้นพิมพ์มีขนาดใหญ่ เพื่อให้มีตำแหน่งประจำของตัวอักขระแต่ละตัว ดังนั้นจึงมีการเพิ่มแป้นยกแคว่ (Shift Key) ซึ่งจะช่วยเพิ่มจำนวนตำแหน่งของตัวอักขระบนแป้นพิมพ์ทำให้สามารถกำหนดตำแหน่งตัวอักขระ 1 แป้นมีตัวอักขระมากกว่า 1 ตัวทำให้ขนาดของแป้นพิมพ์มีขนาดไม่ใหญ่มากแต่จะทำให้ระยะเวลาที่จะทำการพิมพ์ได้ลดลงโดยถึงแม้จะได้ฝึกพิมพ์จนถึงที่สุดแล้วก็ยังคงไม่สามารถทำความเร็วได้เทียบเท่ากับการพิมพ์ภาษาอังกฤษ และระยะเวลาที่ใช้จดจำตำแหน่งของตัวอักขระต้องใช้เวลาานกว่า อย่างไรก็ตามได้มีการพัฒนาการป้อนข้อมูลภาษาไทย เพื่อแก้ไขปัญหาดังต้นนั้นสามารถทำได้ 2 แนวทาง ได้แก่ การพัฒนาอุปกรณ์นำเข้าสู่ข้อมูล และการพัฒนาทักษะของผู้ใช้

การพัฒนาอุปกรณ์การนำเข้าสู่ข้อมูลหรือแป้นพิมพ์นั้น เป็นการพัฒนารูปแบบการจัดวางตัวอักขระ (Layout Keyboard) โดยในปัจจุบันแป้นพิมพ์ภาษาไทยที่ถูกพัฒนาขึ้นมี 2 แบบ คือ แป้นพิมพ์เกษมณี และแป้นพิมพ์ปัตตโชติ การวางตัวอักขระแบบเกษมณีนั้นเป็นรูปแบบการจัดวางตัวอักขระมาตรฐานที่ได้รับความนิยมใช้กันอยู่ในปัจจุบัน ส่วนการวางตัวอักขระแบบปัตตโชตินั้นถูกพัฒนาโดยมุ่งเน้นการทำให้ภาระงานของมือทั้งสองข้างเท่าเทียมกัน เพื่อที่จะเพิ่มประสิทธิภาพในการพิมพ์ให้สูงขึ้น ทั้งนี้จากการศึกษาของ Koanantakool Thaweesak (1993) พบว่าแป้นพิมพ์แบบปัตตโชติสามารถทำให้ภาระงานของมือทั้งสองเท่าเทียมกัน และไม่ทำให้ภาระงานนั้นไปอยู่ที่มือข้างใดข้างหนึ่งมากเกินไป โดยภาระงานของแป้นพิมพ์แบบเกษมณีจะอยู่ที่มือข้างขวา และนิ้วก้อยข้างขวาก็จะมีภาระงานมากกว่านิ้วชี้ข้างซ้าย ซึ่งเป็นนิ้วที่มีความถนัด

มากกว่า แต่อย่างไรก็ตามรูปแบบการจัดวางตัวอักษรแบบปัตตโชตินั้นไม่ได้รับความยอมรับอย่างแพร่หลายจากผู้ใช้งาน ถึงแม้ว่าแป้นพิมพ์แบบปัตตโชติจะสามารถเพิ่มประสิทธิภาพในการพิมพ์ของผู้ใช้ได้ 27% ก็ตาม

การพัฒนาทักษะของผู้ใช้ สามารถพัฒนาได้ใน 2 รูปแบบ ได้แก่ การพัฒนาแบบฝึกและเทคนิคหรือวิธีการฝึก โดยการพัฒนาทั้ง 2 ประการนั้นได้มีการพัฒนาควบคู่กันไปอย่างต่อเนื่อง โดยในเทคนิคการพิมพ์ขั้นพื้นฐานที่ผู้ใช้งานส่วนมากใช้ในการพิมพ์ขั้นเริ่มต้นตอนหัดฝึกพิมพ์ใหม่ ๆ คือ เทคนิคการพิมพ์ใช้นิ้วเดียวกดแป้น (Hunt and Peak) ต่อมาจึงมีการพัฒนาเทคนิคที่จะเพิ่มประสิทธิภาพการพิมพ์ ซึ่งจะกำหนดนิ้วในการกดแป้นตัวอักษร และตำแหน่งในการวางนิ้วบนแป้นพิมพ์เทคนิคดังกล่าวคือ แป้นเหย้า (Home Key) ได้มุ่งเน้นพัฒนาตามทักษะการพิมพ์ของมนุษย์ ซึ่งจะช่วยลดระยะเวลาในการเรียนรู้แบบฝึกพิมพ์ โดยผู้ใช้งานสามารถจดจำตัวอักษร และตำแหน่งบนแป้นพิมพ์ได้รวดเร็ว ซึ่งส่งผลให้ผู้พิมพ์ตัวอักษรที่ต้องการได้แม่นยำมากขึ้น

ในการพัฒนาการป้อนข้อมูลภาษาไทยตามแนวทางที่กล่าวมาในข้างต้นนี้ การที่จะวัดประสิทธิภาพการเรียนรู้ที่ได้จากแบบฝึก หรือแป้นพิมพ์ที่พัฒนาขึ้นมา การทดลองดังกล่าวจะใช้มนุษย์เป็นกลุ่มตัวอย่างทดลอง โดยจะแบ่งกลุ่มตัวอย่างทดลองตามวิธีที่ถูกพัฒนาขึ้น เช่น จำนวนวิธีที่พัฒนาขึ้นมี 10 วิธี ในการทดลองจะต้องแบ่งกลุ่มตัวอย่างทดลองออกเป็น 10 กลุ่ม เป็นต้น ซึ่งการทดลองในแต่ละวิธีดังกล่าวนั้นจะมีความแตกต่างของทักษะการพิมพ์ของผู้ทดลองแต่ละคนในกลุ่มตัวอย่างทดลองที่ไม่เท่ากัน และในแต่ละกลุ่มตัวอย่างทดลองที่ใช้ทดลองกับแต่ละวิธีที่พัฒนาที่จะวัดประสิทธิภาพของวิธีที่พัฒนาออกมา ซึ่งความแตกต่างดังกล่าวส่งผลให้การวัดประสิทธิภาพของการเรียนรู้ของผู้ใช้ไม่ได้ค่าที่ตรงกับความเป็นจริงคือ ค่าที่ได้จะเป็นค่าเฉลี่ยของกลุ่มตัวอย่างทดลองนั้น ๆ ในการแก้ปัญหาดังกล่าวได้มีการนำกลุ่มตัวแทนชาญฉลาด (Intelligent Agent) มาใช้แทนกลุ่มบุคคลในการทดลอง โดยกลุ่มตัวแทนชาญฉลาดที่สร้างขึ้นจะทดลองกับทุกวิธีที่ถูกพัฒนาขึ้น โดยจะเริ่มทดลองในวิธีแรกหาอัตราการเรียนรู้ และเมื่อทดลองในวิธีแรกเสร็จกลุ่มตัวแทนชาญฉลาดจะลบข้อมูลที่ได้จากการฝึกเพื่อกลับไปยังค่าทักษะเริ่มต้นก่อนทดลองด้วยวิธีแรก แล้วจึงทดลองด้วยวิธีที่สองต่อ และจะทำเช่นนี้ต่อไปเรื่อย ๆ จนครบตามจำนวนวิธีที่ถูกพัฒนาขึ้น ทั้งนี้การทดลองโดยกลุ่มตัวแทนชาญฉลาดจะช่วยขจัดปัญหาที่เกิดจากความแตกต่างของทักษะการพิมพ์ หรือข้อจำกัดทางด้านเวลา และงบประมาณที่ใช้ในการทดลอง อีกทั้งมนุษย์นั้นสามารถทดลองได้เพียงครั้งเดียว ซึ่งในการเสนอวิธีการดังกล่าวช่วยให้ได้ค่าประสิทธิภาพการเรียนรู้ของผู้ใช้ตามวิธีต่าง ๆ ที่ได้พัฒนาขึ้นมา เพื่อเลือกแบบฝึก หรือแป้นพิมพ์ที่เหมาะสมกับผู้ใช้งานแต่ละคน

งานวิจัยนี้ศึกษาแนวทางการพัฒนาการบ่อนข้อมูลภาษาไทยของผู้ใช้ เพื่อหาแบบฝึก และแป้นพิมพ์ที่พัฒนามาอย่างหลากหลายให้เหมาะสมกับผู้ใช้แต่ละคนที่มีทักษะการพิมพ์ ซึ่งได้มาจากการฝึกพิมพ์ดีดด้วยแบบฝึก หรือแป้นพิมพ์ที่มีประสิทธิภาพในการเรียนรู้ที่ได้แตกต่างกัน ซึ่งแนวทางที่งานวิจัยนี้เสนอตัวแทนชาวจังหวัดทดลองหาแบบฝึก หรือแป้นพิมพ์ที่เหมาะสมนั้น ซึ่งสามารถใช้ตัวแทนชาวจังหวัดทดลองกับแบบฝึกจำนวนมาก และแตกต่างกันได้ ทั้งนี้ ในการสร้างตัวแทนชาวจังหวัดจำเป็นต้องสร้างแบบจำลองเวลาที่ใช้ในการพิมพ์ของผู้ใช้แต่ละคนก่อน ซึ่งแบบจำลองดังกล่าวเป็นตัวแทนของผู้ใช้แต่ละคนที่ฝึกพิมพ์กับแบบฝึก หรือแป้นพิมพ์นั้น ๆ ในการสร้างแบบจำลองจะใช้ค่าเวลาการพิมพ์ ซึ่งเป็นค่าที่แสดงถึงพฤติกรรมกรรมการพิมพ์ของผู้ใช้แต่ละคน เป็นข้อมูลอินพุตมาวิเคราะห์สร้างแบบจำลองด้วยเครื่องมือที่งานวิจัยนี้เลือกใช้ เพื่อทำนายเวลาการพิมพ์ที่ผู้ใช้แต่ละคนทำได้ขณะฝึกพิมพ์กับแบบฝึก หรือแป้นพิมพ์ นอกจากนี้ยังได้เปรียบเทียบผลการทดลองที่วัดประสิทธิภาพการทำนาย เพื่อค้นหาเครื่องมือที่สามารถจำลองพฤติกรรมกรรมการพิมพ์ผู้ใช้ได้ดีที่สุดระหว่าง แบบจำลองต้นไม้เอ็มไพร์ฟี (Model Tree: M5P) โครงข่ายประสาทเทียม (Artificial Neural Network) และการถดถอยเชิงเส้น (Linear Regression)

ในส่วนข้อมูลเวลาการพิมพ์ของมนุษย์ที่นำมาใช้เป็นข้อมูลสร้างแบบจำลองเป็นค่าไม่คงที่ และเปลี่ยนแปลงตามลำดับเวลา (Dix et al., 1992) โดยการเปลี่ยนแปลงนั้นจะเกิดขึ้นในแต่ละครั้งที่พิมพ์ตัวอักษร ดังนั้นการที่จะนำข้อมูลดังกล่าวมาใช้ในการสร้างแบบจำลองจำเป็นต้องใช้เครื่องมือช่วยวิเคราะห์ข้อมูลรูปแบบดังกล่าว ซึ่งโครงข่ายประสาทเทียมเป็นเครื่องมือหนึ่งที่ถูกเสนอ (Hill et al., 1996) และ (Foka, 1999) เพื่อใช้วิเคราะห์ข้อมูล และทำนายข้อมูลว่าจะเป็นไปได้ในทิศทางใด เนื่องจากข้อมูลลักษณะนี้ส่วนใหญ่เป็นงานทางด้านการเงิน (Finance) หรือในการทำนายค่าการซื้อขายในตลาดหุ้น โดยโครงข่ายประสาทเทียมที่ใช้ในการวิเคราะห์ข้อมูลนั้นมีอยู่ 2 รูปแบบคือ โครงข่ายประสาทเทียมแบบรีเคอร์เรนต์ (Recurrent Neural Network) (Dorffner, 1996) และ (Oeda et al., 2006) ที่มีความสามารถในการจดจำสถานะการทำงาน เพื่อนำสถานะดังกล่าวมาปรับค่าน้ำหนักให้ทำนายผลได้แม่นยำมากขึ้น และอีกรูปหนึ่งคือ โครงข่ายประสาทเทียมแบบทิศทางการบ่อนไปข้างหน้า (Feed-Forward Neural Network) (Plummer, 2000) กับเทคนิคจัดการข้อมูลแบบหน้าต่างเลื่อน (Sliding windows) ที่จะจัดการข้อมูลอินพุตก่อนวิเคราะห์ด้วยเครื่องมือที่เลือกใช้ ซึ่งงานวิจัยนี้ได้ประยุกต์ใช้เทคนิคการจัดการข้อมูลด้วยรูปแบบที่กล่าวมาในข้างต้นคือ เทคนิคหน้าต่างเลื่อน และอีกเทคนิคหนึ่งคือ เทคนิคการแบ่งกลุ่มข้อมูล (Clustering Data) ด้วยค่าเฉลี่ยเคกลุ่ม (K-Means) (Frühwirth-Schnatter, 2004) และการแบ่งกลุ่มแยกตามแถวของคู่อักษร (Row-Based) ที่จะแบ่งกลุ่มข้อมูล

อินพุตก่อนที่จะนำมาวิเคราะห์สร้างแบบจำลองด้วยเครื่องมือทั้ง 3 ที่งานวิจัยนี้เลือกใช้ และเปรียบเทียบประสิทธิภาพของแบบจำลองที่พัฒนาว่าวิธีการจัดการข้อมูลที่เพิ่มไปนั้นช่วยลดอัตราความผิดพลาดที่ได้จากการทำนายของแบบจำลอง และเวลาที่แบบจำลองทำนายได้ใกล้เคียงกับสิ่งที่ผู้ใช้แต่ละคนสามารถทำได้มากที่สุด

ผลการศึกษาของงานวิจัยสามารถนำไปประยุกต์ใช้ เพื่อช่วยในการออกแบบแบบฝึกหัดหรือแป้นพิมพ์ภาษาไทยที่เหมาะสมสำหรับผู้ใช้แต่ละคน เนื่องจากแบบฝึกหัดพิมพ์ภาษาไทยโดยทั่วไปนั้นจัดทำขึ้นมาสำหรับผู้ใช้ทั่วไปไม่ได้ออกแบบมาสำหรับผู้ใช้คนใดคนหนึ่ง ซึ่งผู้ใช้งานบางคนอาจจะพอพิมพ์บางอักขระได้ หรือจำบางอักขระได้แล้ว เมื่อเริ่มฝึกในแบบฝึกก็อาจมีความสามารถอยู่ในแบบฝึกหัดบทแรก ๆ เมื่อต้องฝึกซ้ำในสิ่งที่ทำได้แล้ว อาจทำให้ผู้ใช้เกิดความเบื่อหน่าย ซึ่งผลจากงานวิจัยนี้สามารถนำไปใช้แก้ปัญหาดังกล่าวได้ และกระบวนการในการสร้างแบบจำลองที่งานวิจัยนี้เสนอสามารถประยุกต์ใช้ทดลองแทนมนุษย์ในการเรียนรู้ทักษะอื่น ๆ ที่มีความใกล้เคียงกัน

## 1.2 วัตถุประสงค์ของงานวิจัย

1. เพื่อสร้างแบบจำลองที่ใช้ในการพิมพ์ของผู้ใช้แต่ละคน
2. เพื่อศึกษาแนวทางสำหรับใช้พัฒนาทักษะการพิมพ์ของผู้ใช้
3. เพื่อศึกษาพัฒนาการของทักษะการพิมพ์ของผู้ใช้
4. เพื่อเปรียบเทียบแบบจำลองที่ถูกสร้างด้วย การถอดถอดเชิงเส้น โครงข่ายประสาทเทียม และแบบจำลองต้นไม้เอ็มไพร์พี
5. เพื่อศึกษาการแบ่งกลุ่มข้อมูลส่งผลต่อประสิทธิภาพการทำนายของแบบจำลองที่สร้างขึ้น

## 1.3 ขอบเขตของงานวิจัย

1. กลุ่มผู้ทดลองเป็นนักศึกษาปริญญาตรีจำนวน 5 คนที่ไม่มีได้รับการฝึกพิมพ์ดีดมาก่อนเป็นกรณีศึกษา
2. แบบฝึกที่ใช้ได้มาจากบทเรียนฝึกพิมพ์ดีดภาษาไทย
3. ข้อมูลเวลาการพิมพ์ของคู่อักขระที่มีช่องว่าง (Space Key) ไม่ได้นำมาใช้วิเคราะห์สร้างแบบจำลอง

4. งานวิจัยนี้ทดลองบนแป้นพิมพ์แบบเกษมณี
5. คัดกรองค่าระยะเวลาการพิมพ์ที่เกินจาก 30 เซนติวินาที (Centiseconds) ออกจากกลุ่มข้อมูล

#### 1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้แบบจำลองการพิมพ์ของผู้ใช้ เพื่อพัฒนาตัวแทนชาวนาถนัดทดลองแทนผู้ใช้
2. ได้แบบจำลองที่ช่วยในการพัฒนาแบบฝึก และรูปแบบการจัดวางตัวอักษรบนแป้นพิมพ์ที่เหมาะสมกับผู้ใช้
3. ได้แบบจำลองที่ช่วยลดระยะเวลาในกระบวนการทดลองกับผู้ใช้

#### 1.5 รายละเอียดของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ได้แบ่งการนำเสนอรายละเอียดของเนื้อหาออกเป็น 5 บท ดังนี้  
 บทที่ 1 บทนำ ซึ่งกล่าวถึงความสำคัญของปัญหา วัตถุประสงค์ของงานวิจัย ขอบเขตของงานวิจัย ประโยชน์ที่คาดว่าจะได้รับ และรายละเอียดของวิทยานิพนธ์

บทที่ 2 กล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง ประกอบด้วยทฤษฎีต่าง ๆ ที่เกี่ยวข้องกับงานวิจัยนี้ แบบจำลองการทำงานของสมองรูปแบบต่าง ๆ การถอดออยเชิงเส้น โครงข่ายประสาทเทียม แบบจำลองต้นไม้เอ็มไพร์ฟี ข้อมูลอนุกรมเวลา และค่าเฉลี่ยเคกลุ่ม (K-Means) วิธีการทดสอบแบบไขว้ข้ามสับกลุ่ม

บทที่ 3 วิธีการดำเนินงานวิจัย ประกอบด้วยรายละเอียดข้อมูลที่ใช้ในการวิจัย เครื่องมือที่ใช้ ขั้นตอนของการวิจัย และวิธีวัดผลการวิจัย

บทที่ 4 ผลการทดลอง แสดงผลการทดลองที่ได้แต่ละกระบวนการ

บทที่ 5 เสนอสรุปผลการศึกษาและข้อเสนอแนะ ประกอบด้วย ผลสรุปและอภิปรายผลการศึกษาวิจัย และข้อเสนอแนะต่าง ๆ