



บทที่ 1
บทนำ

แนวเหตุผล

เทคโนโลยีต่าง ๆ ถูกคิดค้นและพัฒนาขึ้นมา เพื่ออำนวยความสะดวกให้กับมนุษย์หรือเพื่อใช้ทำงานแทนมนุษย์ในสภาพแวดล้อมที่อาจเป็นอันตรายต่อมนุษย์ได้ ในการใช้งานมนุษย์มีวิธีติดต่อสื่อสารหรือควบคุมเทคโนโลยีเหล่านี้ได้หลายวิธี การรู้จำเสียงพูด (speech recognition) เป็นอีกรูปแบบหนึ่งของการติดต่อสื่อสารระหว่างมนุษย์กับคอมพิวเตอร์ในรูปแบบของเสียง ซึ่งเป็นรูปแบบหนึ่งของการติดต่อสื่อสารที่มนุษย์คุ้นเคย ทั้งสะดวกในการใช้งาน และยังช่วยให้คนพิการทางประสาทตาสามารถติดต่อสื่อสารกับคอมพิวเตอร์ได้อีกด้วย งานวิจัยที่เกี่ยวกับการรู้จำเสียงพูดนั้นมีอยู่ด้วยกันหลายวิธี เช่น KNN (K nearest neighbour), DTW (dynamic time warping), HMM (hidden markov model) ซึ่งเทคนิคเหล่านี้ถูกนำมาใช้ในการรู้จำเสียงภาษาไทยด้วย (ธีระ ภัทรพรนันท์ 2538), (เสาวลักษณ์ อารีย์พงศา 2538) ปัจจุบันมีการนำนิเวรอลเน็ตเวิร์ก (neural network) มาประยุกต์ใช้กับงานด้านต่าง ๆ มากมาย รวมทั้งการรู้จำเสียงพูดด้วย จากงานวิจัยในต่างประเทศ (Lubensky ,1991) พบว่าการรู้จำเสียงพูดโดยใช้นิเวรอลเน็ตเวิร์ก เป็นวิธีการที่มีประสิทธิภาพเทียบเท่าหรืออาจดีกว่าเทคนิคที่ใช้กันทั่วไป โดยมีข้อดีที่สำคัญคือ ใช้เวลาในการทดสอบการรู้จำน้อยกว่าวิธีอื่น ๆ รวมทั้ง DTW จากข้อมูลใน (Reynolds and Tarassenko, 1991), (Kammerer and Kupper ,1989) และ HMM จากข้อมูลใน (Schurer ,1994), (Bedworth ,1989) ด้วย นอกจากนี้ยังมีการค้นพบเทคนิคที่ทำให้นิเวรอลเน็ตเวิร์ก สามารถเรียนรู้สิ่งใหม่เพิ่มเติมขึ้นด้วย (Haiyan and Chengyi ,1992), (Zeng and Yu ,1992) ซึ่งมีข้อดีคือ ทำให้สามารถขยายความสามารถในการเรียนรู้ของนิเวรอลเน็ตเวิร์ก เช่นทำให้ระบบสามารถรู้จำคำศัพท์เพิ่มขึ้นได้ งานวิจัยนี้เลือกใช้ back propagation neural network เนื่องจากมีข้อดีเหนือกว่านิเวรอลเน็ตเวิร์กแบบอื่นดังนี้ (Elvira and Carrasco ,1992), (Kung ,1993)

- 1 ความสามารถในการจำดีกว่า hopfield net ซึ่งอยู่ในประเภท fixed-weight associative memory network
- 2 วิธีการในการวิเคราะห์การลู่เข้า (convergence) แน่นอน เมื่อเทียบกับของ self-organizing feature map ซึ่งอยู่ในประเภท competitive learning network
- 3 ความสามารถในการจำแนกกลุ่มสูงกว่า linear perceptron ซึ่งอยู่ในประเภท decision-based neural network

อย่างไรก็ตาม ยังมีนิเวรอลเน็ตเวิร์กแบบอื่นที่นิยมใช้ในการรู้จำหน่วยเสียง (phoneme) เช่น TDNN (time delay neural network) (Waibel, et al. ,1988) แต่ TDNN มีข้อเสียคือมีโครงสร้างที่ซับซ้อน

ข้ออื่น อีกเหตุผลหนึ่งที่เลือกใช้ back propagation neural network ก็เป็นเทคนิคที่สามารถนำไปประยุกต์ใช้กับ continuous speech recognition ได้ เช่น งานวิจัยที่ใช้ back propagation neural network เพื่อทำ Tone recognition (Chen and Wang ,1995) และการนำ back propagation neural network ไปใช้ร่วมกับ HMM ในการรู้จำเสียงพูด (Jin and Chung ,1993)

วัตถุประสงค์

1. เพื่อพัฒนาวิธีการรู้จำเสียงตัวเลขภาษาไทย โดยใช้ Back propagation neural network
2. เพื่อเป็นแนวทางสำหรับการนำเทคนิคนี้ไปใช้กับงานรู้จำเสียงต่อเนื่อง

เป้าหมายและขอบเขต

1. สามารถรู้จำเสียงตัวเลขภาษาไทย แบบพยางค์เดียวโดด ๆ โดยมีอัตราการรู้จำ 80 % ขึ้นไป
2. สามารถรู้จำเสียงตัวเลขภาษาไทย เฉพาะกลุ่มคนเพศชายอายุ 17-25 ปี
3. ขยายขีดความสามารถให้ใช้กับการรู้จำเสียงตัวเลขภาษาไทยแบบหลายพยางค์คือ 2 และ 3 พยางค์

ประโยชน์ที่คาดว่าจะได้รับ

1. วิธีการรู้จำเสียงตัวเลขภาษาไทยแบบคำโดด ๆ ได้โดยใช้นิวรอลเน็ตเวิร์กและแอลพีซี
2. เป็นแนวทางในการศึกษาวิธีการรู้จำเสียงอักขระภาษาไทยโดยใช้นิวรอลเน็ตเวิร์ก