

ชื่อเรื่องการค้นคว้าแบบอิสระเชิงวิทยานิพนธ์

การรู้จำเสียงคำโดดด้วยโครงข่ายประสาทเทียม

ชื่อผู้เขียน

นางสาวกาญจนา ทองบุญนาท

วิทยาศาสตร์มหาบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

คณะกรรมการสอบการค้นคว้าแบบอิสระเชิงวิทยานิพนธ์

อาจารย์ ดร.เสมอแซ	สมหอม	ประธานกรรมการ
รองศาสตราจารย์วิไลพร	สิริมังครารัตน์	กรรมการ
อาจารย์ปรียานุช	วงศ์ประภาส	กรรมการ

## บทคัดย่อ

การค้นคว้าแบบอิสระเชิงวิทยานิพนธ์นี้มีวัตถุประสงค์เพื่อศึกษาและพัฒนาในส่วนของ การจัดเตรียมข้อมูลนำเข้าให้กับโครงข่ายประสาทเทียมเพื่อการรู้จำเสียงคำโดด ซึ่งกระบวนการใน การรู้จำเสียงประกอบด้วย 4 ขั้นตอนคือ การประมวลผลสัญญาณเบื้องต้น การสกัดลักษณะเด่น การ ทดสอบความคล้ายคลึงกันของรูปแบบ และการตัดสินใจ

ในการประมวลผลสัญญาณเบื้องต้นใช้หลักการพิจารณาค่าพลังงานและอัตราการตัดศูนย์ การสกัดลักษณะเด่นใช้การหาค่าสัมประสิทธิ์การประมาณพหุเชิงเส้นลำดับ 10 ทั้งสองขั้นตอนนี้ ใช้โปรแกรมแมทแลป

ในขั้นตอนการทดสอบความคล้ายคลึงกันของรูปแบบ ใช้อัลกอริทึมแบบการแพร่กระจาย กลับและขั้นตอนการตัดสินใจใช้ความคลาดเคลื่อนต่ำสุดในการเปรียบเทียบ ในสองส่วนนี้ใช้ โปรแกรมภาษาซี

การทดสอบประสิทธิภาพของการจัดเตรียมข้อมูลนำเข้าให้กับโครงข่ายประสาทเทียม ใช้ ข้อมูลเสียงตัวอย่างสามชุด จากกลุ่มตัวอย่างจำนวน 30 คน โดย 20 คนแรกใช้เป็นชุดฝึกฝนและ ทดสอบ และอีก 10 คนเป็นชุดทดสอบ แต่ละชุดเป็นข้อมูลเสียงที่บันทึกจากผู้พูดทั้งเพศหญิงและ ชายที่มีอายุอยู่ในช่วง 18-25 ปี โดยผู้พูดจะพูดคำว่า ฉิน ดู ข้าว กิน ไป คนละสองครั้ง หลังจากนั้น ข้อมูลจะถูกแบ่งดังนี้ ข้อมูลชุดฝึกฝนและชุดทดสอบที่หนึ่งเป็นข้อมูลที่บันทึกจากผู้พูดกลุ่มเดียวกัน

แต่บันทึกข้อมูลไว้คนละชุด ส่วนข้อมูลชุดทดสอบที่สอง เป็นข้อมูลที่บันทึกจากผู้พูดต่างกลุ่มออกไป ค่าพารามิเตอร์ที่ใช้คือ อัตราการเรียนรู้เท่ากับ 0.3 ค่าความผิดพลาดเฉลี่ยเท่ากับ 0.02 ใช้จำนวนโหนดในระดับซ่อนตัวเท่ากับ 100 โหนด ผลการทดลองได้อัตราการรู้จำเฉลี่ยของแต่ละกลุ่มเรียงตามลำดับคือร้อยละ 97.50 79.50 และ 89 จากค่าที่ได้พบว่า การประยุกต์ใช้โครงข่ายประสาทเทียมทำให้ระบบสามารถรู้จำเสียงได้เกินกว่า 80 เปอร์เซ็นต์ และใช้เวลาน้อยในส่วนของการรวบรวมการจัดเตรียมข้อมูลนำเข้าให้กับโครงข่ายประสาทเทียม โดยใช้โปรแกรมแมทแลป

มหาวิทยาลัยเชียงใหม่  
Chiang Mai University

Research Title	Single Word Recognition with Artificial Neural Network	
Author	Miss. Kanjana Thongboonnak	
M.S.	Computer Science	
Examining Committee	Dr.Samerkae Somhom	Chair Person
	Assoc.Prof. Wilaiporn Sirimungkrarat	Member
	Miss Preeyanuj Wongprapas	Member

### Abstract

The research for Independent study was performed in the purpose of improve input data preparation process for Artificial Neural Network (ANN). The process of recognition was separated into 4 steps; Voice Signal Preprocessing, Voice Feature Extraction, Pattern Similarity and Decision Rule

Voice Signal Preprocessing step considered the zero-crossing rate of recognized voice. After passed this step, the zero-crossing rate helped improve output voice signal by using end-point detection. The second step, Voice Feature Extraction, considered Linear Prediction Coefficient (LPC) to help finding the test pattern. This step was performed, by using Matlab program, to support complicated match calculation. The third step, Pattern Similarity, used back-propagation algorithm to get the distance score. The final step, Decision Rule, used the minimum error distance method to help matching the distance score with nearest value in the pattern. The third and fourth steps were performed by using C programming.

In testing process, 3 different sampling tests were distributed to 30 samplers to get the most accuracy value. Samplers come from both female and male, age between 18-25 years. All samplers were separated into 2 groups; 20 samplers, 10 samplers. The first group performed 2 different sampling tests; training set and testing set I. The second group performed testing set II.

From the testing process, the result has shown 3 major proper parameters used in this process. The learning rate should be 0.3, the proper summary-square error rate should be 0.02 and the proper hidden nodes should be 100 nodes. The average recognition rates from the 3 tests are 97.50%, 79.50%, and 89% respectively.

Finally, the testing results had shown that applying ANN technology to create isolated word recognition can help improve voice recognition more than 80 percents and take a little time in input data preparation process.