บทที่ 5

สรุปการวิจัยและข้อเสนอแนะ

สรุปการวิจัย

ปัญหาของระบบไอแอลพี คือ เมื่อกฎที่สร้างขึ้นไปใช้จำแนกตัวอย่างที่ไม่เคยพบมาก่อน หรือตัวอย่างที่มีสัญญาณรบกวนแล้วบางครั้งไม่มีกฎข้อใดในเซตของกฎที่ตรงกับตัวอย่างพอดี ระบบไอแอลพีเพียงอย่างเดียวจึงไม่สามารถจำแนกตัวอย่างในกรณีนี้ได้ ในการวิจัยครั้งนี้ได้ทำการทดลองเพื่อหาวิธีการที่สามารถประมาณกฎลำดับที่หนึ่งซึ่งได้จากกระบวนการสร้างของระบบไอแอลพี โดยใช้กระบวนการดึงลักษณะสำคัญประกอบกับวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก กระบวนการดึงลักษณะสำคัญจะทำการดึงลักษณะสำคัญภายในกฎลำดับที่หนึ่งซึ่งสร้างจากระบบไอแอลพี (PROGOL และ GOLEM) จากนั้นนำมาเปรียบเทียบกับตัวอย่างที่ใช้เรียนรู้และตัวอย่างที่ใช้ทดสอบ ร่วมกับความรู้ภูมิหลังซึ่งอยู่ในรูปตรรกะลำดับที่หนึ่งเช่นกัน เพื่อสร้างอินพุตและเอาต์พุตเวกเตอร์ให้กับนิวรอลเน็ตเวิร์กเพื่อทำการเรียนรู้ด้วยวิธีการแบ็กพรอพาเกชัน จากนั้นจึงนำไปทดสอบกับตัวอย่างที่ใช้ทดสอบ โดยอาศัยชุดข้อมูล 4 ชุด คือ การรู้จำตัวพิมพ์อักษรไทย การวิเคราะห์โครงสร้างไฟไนต์เอลิเมนต์ การวิเคราะห์ความสามารถก่อกลายพันธุ์ และการวิเคราะห์ตำแหน่งตัวหมากรุกสากล ผลการทดลองที่ได้แสดงให้เห็นถึงเปอร์เซ็นต์ความถูกต้องที่เพิ่มขึ้นของระบบ BANNAR เมื่อเทียบกับกฎเดิมซึ่งได้จากระบบ PROGOL หรือ GOLEM และเมื่อเทียบกับระบบอื่นอีก 4 ระบบใน 4 ชุดข้อมูล รวมเปรียบเทียบ 16 ครั้ง ผลปรากฏว่า ระบบ PROGOL หรือ GOLEM ให้เปอร์เซ็นต์ความถูกต้องต่ำกว่าระบบ TILDE และ LINUS ในชุดข้อมูลที่เป็นแบบหลายกลุ่ม และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าระบบ TILDE และ LINUS ในชุดข้อมูลที่มีลักษณะเป็นสองกลุ่ม ผลการทดลองนี้แสดงให้เห็นว่ากฎที่ได้จากระบบ PROGOL และ GOLEM สามารถใช้ได้เป็นอย่างดีในปัญหาที่มีลักษณะเป็นสองกลุ่ม แต่ให้เปอร์เซ็นต์ความถูกต้องไม่ดีเมื่อใช้กับปัญหาที่มีลักษณะเป็นหลายกลุ่ม เมื่อใช้กระบวนการดึงลักษณะสำคัญและนิวรอลเน็ตเวิร์กในระบบ BANNAR ทำให้เปอร์เซ็นต์ความถูกต้องสูงขึ้นและสูงกว่าระบบอื่นในทุกชุดข้อมูล โดยให้ความถูกต้องสูงกว่าอย่างมีนัยสำคัญทางสถิติสูงกว่า 90.0% จำนวน 12 ครั้ง จากการเปรียบเทียบ 16 ครั้ง และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าอย่างมีนัยสำคัญทางสถิติต่ำกว่า 90.0% จำนวน 4 ครั้ง จากการเปรียบเทียบ 16 ครั้ง

เมื่อพิจารณาเปอร์เซ็นต์ความถูกต้องที่เพิ่มขึ้นของระบบ BANNAR โดยแยกกลุ่มตัวอย่างทดสอบออกเป็นสองกลุ่ม คือ กลุ่มตัวอย่างที่ถูกครอบคลุมด้วยกฎและกลุ่มตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎ เปรียบเทียบกันระหว่างระบบ BANNAR กับระบบ PROGOL หรือ GOLEM ผลปรากฏว่า ระบบ

BANNAR สามารถจำแนกตัวอย่างได้ถูกต้องเพิ่มขึ้นจาก 3 ตัวอย่าง เป็น 453 ตัวอย่างในชุดข้อมูล TCR และเพิ่มขึ้นจาก 14 ตัวอย่างเป็น 34 ตัวอย่างในชุดข้อมูล FEM แสดงให้เห็นว่าเปอร์เซ็นต์ความถูกต้องที่สูงขึ้นของระบบ BANNAR เมื่อเทียบกับระบบ PROGOL หรือ GOLEM เป็นผลมาจากกลุ่มตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม ซึ่ง PROGOL หรือ GOLEM จำแนกตัวอย่างได้ถูกต้องน้อยกว่าระบบ BANNAR

เมื่อทดสอบระบบ BANNAR กับชุดข้อมูลที่ได้รับการเพิ่มสัญญาณรบกวน โดยทดลองเปรียบเทียบระหว่างระบบ GOLEM, BANNAR และ FOSSIL ในชุดข้อมูล KRK พบว่า เปอร์เซ็นต์ความถูกต้องที่ได้จากระบบ GOLEM ลดลงเร็วกว่าระบบ BANNAR และ FOSSIL มาก เมื่อพิจารณาค่าน้ำหนักที่ได้จากนิวรอลเน็ตเวิร์ก พบว่าค่าน้ำหนักของลักษณะสำคัญซึ่งมีความสำคัญต่างกันจะมีค่าไม่เท่ากัน สัญพจน์บางสัญพจน์ในกฎแต่ละข้ออาจไม่จำเป็นต้องใช้ในการจำแนกตัวอย่าง แต่ถูกเพิ่มเข้ามาในกฎในกระบวนการลร้างกฎของระบบ GOLEM ทำให้กฎข้อนั้นมีความเฉพาะเจาะจงกับตัวอย่างมากเกินไป เมื่อใช้นิวรอลเน็ตเวิร์กกำหนดความสำคัญให้แก่ลักษณะสำคัญแต่ละข้อภายในกฎ ทำให้กฎแต่ละข้อมีความยืดหยุ่นมากขึ้น ในกรณีที่ตัวอย่างทดสอบไม่ตรงกับบางสัญพจน์ที่ไม่สำคัญภายในกฎข้อใดข้อหนึ่ง ตัวอย่างนั้นอาจถูกจำแนกด้วยกฎข้อนั้นก็ได้

นอกจากการทดลองต่างๆ ข้างต้นแล้ว ในงานวิจัยนี้ยังได้ทำการทดลองเปรียบเทียบประสิทธิภาพของลักษณะสำคัญที่ได้จากระบบ BANNAR กับลักษณะสำคัญจากระบบ LINUS เพื่อแสดงให้เห็นประสิทธิภาพของลักษณะสำคัญจากระบบ BANNAR เมื่อเปรียบเทียบกับระบบอื่น โดยใช้นิวรอลเน็ตเวิร์กและระบบ C4.5 ทำการเรียนรู้และทดสอบ ผลการทดลองปรากฏว่า ลักษณะสำคัญที่ได้จากระบบ BANNAR ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าลักษณะสำคัญจากระบบ LINUS ในทุกชุดข้อมูลเมื่อใช้นิวรอลเน็ตเวิร์ก และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าลักษณะสำคัญจากระบบ LINUS 3 ชุดข้อมูล จาก 4 ชุดข้อมูล เมื่อใช้ระบบ C4.5 ในการทดสอบ

ประสิทธิภาพที่เพิ่มขึ้นของระบบ BANNAR เป็นผลมาจากองค์ประกอบหลักสองส่วนของระบบ นั่นคือ กระบวนการดึงลักษณะสำคัญ และนิวรอลเน็ตเวิร์ก งานวิจัยนี้ได้ทดลองเพื่อแสดงให้เห็นถึงความถูกต้องที่เพิ่มขึ้นจากทั้งสองส่วน โดยนำลักษณะสำคัญที่ได้จากระบบ BANNAR เป็นอินพุตป้อนให้กับระบบ 1BC, TILDE และ C4.5 ซึ่งได้ทำการทดลองเปรียบเทียบเปอร์เซ็นต์ความถูกต้องกับระบบ BANNAR ไว้ พบว่าเมื่อใช้ลักษณะสำคัญจากระบบ BANNAR แทนค่าอินพุตปกติของระบบดังกล่าว ทำให้เปอร์เซ็นต์ความถูกต้องของระบบทั้งสามสูงขึ้นในทุกชุดข้อมูล ยกเว้นเมื่อใช้ระบบ C4.5 ในชุดข้อมูล KRK

เพื่อแสดงถึงประสิทธิภาพของลักษณะสำคัญที่ได้จากกระบวนการดึงลักษณะสำคัญของระบบ BANNAR และประสิทธิภาพของนิวรอลเน็ตเวิร์ก ผู้วิจัยได้ทำการทดลองโดยใช้ลักษณะสำคัญเพียงอย่างเดียวมาทำการจำแนกตัวอย่างโดยใช้ฟังก์ชัน *matchRatio* เปรียบเทียบกับความถูกต้องที่ได้จากระบบ PROGOL หรือ GOLEM และความถูกต้องเมื่อใช้นิวรอลเน็ตเวิร์กจำแนกตัวอย่าง ผลการทดลองปรากฏว่า

เมื่อใช้ลักษณะสำคัญเพียงอย่างเดียวสามารถจำแนกตัวอย่างได้ถูกต้องมากขึ้น และสูงกว่ากฎจากระบบ PROGOL หรือ GOLEM โดยเฉพาะอย่างยิ่งกลุ่มตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม และความถูกต้องยังเพิ่มขึ้นอีกเมื่อใช้นิวรอลเน็ตเวิร์กทำการเรียนรู้และทดสอบ ซึ่งแสดงให้เห็นถึงประสิทธิภาพในการจำแนกตัวอย่างของระบบ BANNAR ที่เพิ่มขึ้นจากระบบ PROGOL หรือ GOLEM โดยใช้กระบวนการดึงลักษณะสำคัญร่วมกับแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก

จากการทดลองทั้งหมดสามารถสรุปได้ว่า กฎจากระบบ PROGOL หรือ GOLEM ซึ่งสามารถใช้ได้เป็นอย่างดีในชุดข้อมูลหรือกลุ่มของปัญหาที่มีลักษณะเป็นสองกลุ่ม แต่ให้ผลไม่ดีเท่าที่ควรในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม เมื่อนำกระบวนการดึงลักษณะสำคัญและแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์กของระบบ BANNAR ไปใช้ประกอบกับกฎเหล่านั้น จะทำให้ความสามารถในการรู้จำสูงขึ้นโดยเฉพาะอย่างยิ่งในกลุ่มตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎของชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม ซึ่งจากการทดลองเพื่อทดสอบประสิทธิภาพของลักษณะสำคัญ แสดงให้เห็นว่าเปอร์เซ็นต์ความถูกต้องที่เพิ่มขึ้นนี้เป็นผลมาจากองค์ประกอบทั้งสองส่วนของระบบ BANNAR นอกจากนี้ในการทดลองกับกลุ่มตัวอย่างที่ได้รับการเพิ่มสัญญาณรบกวนยังแสดงให้เห็นว่า นิวรอลเน็ตเวิร์กสามารถกำหนดค่าน้ำหนักที่ต่างกันของเส้นเชื่อมภายในนิวรอลเน็ตเวิร์กเพื่อให้ความสำคัญแก่ลักษณะสำคัญที่มีสำคัญไม่เท่ากันได้อีกด้วย

ข้อเสนอแนะ

1. กระบวนการดึงลักษณะสำคัญ

ในการวิจัยครั้งนี้ จะต้องนำกฎที่ได้จากระบบไอแอลพีมาผ่านกระบวนการดึงลักษณะสำคัญ เพื่อให้ได้ลักษณะสำคัญของกฎแต่ละข้อ ซึ่งต่างจากระบบ LINUS และ 1BC ซึ่งในระบบ LINUS สามารถสร้างลักษณะสำคัญได้จากกลุ่มความรู้ภูมิหลัง ในขณะที่ระบบ 1BC ให้ผู้ใช้กำหนดเพรดิเคตโครงสร้างเพื่อสร้างตัวแปรใหม่และเพรดิเคตลักษณะเพื่อบอกลักษณะของตัวอย่างเอง ทั้งสองระบบสามารถสร้างลักษณะสำคัญกลุ่มความรู้ภูมิหลังได้ ในขณะที่ข้อจำกัดของระบบ BANNAR คือ จะต้องสร้างกฎจากระบบไอแอลพีก่อนจึงจะทำการดึงลักษณะสำคัญได้ หากสามารถหาวิธีการที่ทำการสร้างลักษณะสำคัญจากกลุ่มความรู้ภูมิหลังโดยไม่ต้องผ่านกระบวนการสร้างกฎได้ จะเป็นการลดขั้นตอนและเวลาที่ใช้ในการเรียนรู้รวมทั้งเวลาที่ใช้ในการทดสอบก็จะลดลงตามไปด้วย

2. การหาค่าความจริงของลักษณะสำคัญ

ในการวิจัยครั้งนี้ เมื่อได้ลักษณะสำคัญจากกฎ ลักษณะสำคัญเหล่านี้จะถูกนำไปเทียบกับตัวอย่างที่ใช้เรียนรู้และตัวอย่างที่ใช้ทดสอบ ถ้าตัวอย่างนั้นตรงกับลักษณะสำคัญจะกำหนดค่าให้อินพุตนิวรอนสำหรับลักษณะสำคัญข้อนั้นเป็นจริงซึ่งมีค่าเท่ากับ 1 ถ้าตัวอย่างไม่ตรงกับลักษณะสำคัญ จะกำหนดให้อินพุตนิวรอนสำหรับลักษณะสำคัญข้อนั้นมีค่าเป็นเท็จหรือเท่ากับ -1 สำหรับลักษณะสำคัญบางข้อซึ่งมีจำนวนสัญพจน์มาก หากมีบางสัญพจน์ที่ไม่ตรงกับตัวอย่างนั้น ค่าสำหรับอินพุตนิวรอนจะถูกกำหนดให้เป็นเท็จโดยไม่สนใจว่าตัว

อย่างนั้นใกล้เคียงกับลักษณะสำคัญมากหรือน้อย วิธีการที่น่าสนใจที่จะนำมาปรับใช้เพื่อกำหนดให้เป็นค่าความจริงเป็นอย่างอื่น คือ วิธีฟัซซีลอจิก ซึ่งสามารถกำหนดค่าอื่นซึ่งไม่ใช่ค่าจริงหรือค่าเท็จให้กับลักษณะสำคัญได้

3. การใช้เครื่องมือรู้จำแบบอื่นกับลักษณะสำคัญที่ได้จากระบบ BANNAR

เนื่องจากค่าที่ได้เมื่อนำลักษณะสำคัญไปเทียบกับตัวอย่างจะอยู่ในรูปแบบค่าจริงหรือเท็จเท่านั้น กลุ่มปัญหาที่อยู่ในรูปตรรกะลำดับที่หนึ่ง จะเปลี่ยนไปอยู่ในรูปแบบตรรกศาสตร์ประพจน์ ซึ่งมีแต่ค่าคุณลักษณะเท่านั้น ในการวิจัยนี้มีการทดลองที่นำค่าความจริงของลักษณะสำคัญไปป้อนเป็นตัวอย่างให้กับระบบอื่นๆ เพื่อทำการเรียนรู้และทดสอบ นอกเหนือจากวิธีการนิวรอลเน็ตเวิร์ก, ระบบ 1BC, TILDE และ C4.5 ที่ทดลองในครั้งนี้แล้ว ยังสามารถใช้ระบบเรียนรู้อื่นซึ่งสามารถทำงานกับปัญหาที่มีลักษณะเป็นตรรกศาสตร์ประพจน์ หรือตัวอย่างที่เป็นค่าคุณลักษณะได้ ตัวอย่างของระบบเหล่านี้ได้แก่ ระบบ CN2 [Clark และ Niblett, 1989] หรือ ID3 [Quinlan, 1986] เป็นต้น

4. การเปลี่ยนกฎให้อยู่ในรูปแบบที่สามารถเข้าใจได้ง่าย

ข้อดีสำคัญประการหนึ่งของตรรกะลำดับที่หนึ่ง คือ การที่มนุษย์สามารถเข้าใจรูปแบบของกฎได้ง่าย ต่างกับระบบเรียนรู้อื่นๆ ซึ่งเมื่อนำกฎจากกระบวนการไอแอลพีมาทำการประมาณด้วยนิวรอลเน็ตเวิร์กแล้ว จะทำให้ความมนุษย์ไม่สามารถเข้าใจเอาต์พุตของระบบได้ง่าย ดังนั้นการปรับเปลี่ยนนิวรอลเน็ตเวิร์กให้อยู่ในรูปแบบที่สามารถเข้าใจได้ง่าย จึงเป็นส่วนที่น่าสนใจในการพัฒนาต่อไป