บทที่ 4

ผลการทดลอง

เนื้อหาในบทที่ 4 ประกอบด้วย เครื่องมือที่ใช้ในการทดลอง ชุดข้อมูลที่ใช้ในการทดลอง เวลาที่ใช้ในการทดลอง ผลการทดลองเปรียบเทียบระหว่างวิธีการประมาณกฎด้วยวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก (Backpropagation Neural Network for Approximation Rules: BANNAR) กับวิธีการอื่นๆ ผลการทดลองในกลุ่มตัวอย่างที่มีสัญญาณรบกวน การทดลองเปรียบเทียบลักษณะสำคัญที่ได้กับลักษณะสำคัญจากระบบไอแอลพีอื่น

เครื่องมือที่ใช้ในการทดลอง

ระบบไอแอลพีอื่นที่นำมาใช้ทดลองเพื่อเปรียบเทียบผลกับระบบ BANNAR ประกอบด้วย PROGOL, GOLEM, TILDE, 1BC, LINUS และ FOSSIL ซึ่งรายละเอียดของแต่ละระบบมีดังนี้

1. PROGOL

ระบบ PROGOL [Muggleton, 1995] เป็นระบบไอแอลพีที่มีประสิทธิภาพระบบหนึ่งที่ใช้กันอย่างแพร่หลาย รับอินพุตเป็นเซตของตัวอย่างบวก ตัวอย่างลบ และกลุ่มความรู้ภูมิหลัง ผู้ใช้จะต้องประกาศลักษณะการทำงานของสัญพจน์ (mode declaration) เพื่อให้ระบบรู้ว่าแต่ละอาร์กิวเมนต์ของสัญพจน์จะมีลักษณะเป็นอย่างไร โดยอาร์กิวเมนต์ของสัญพจน์สามารถเป็นได้ทั้งตัวแปรอินพุต ตัวแปรเอาต์พุต และค่าคงที่ นอกจากหน้าที่ของอาร์กิวเมนต์ต่างๆ แล้ว การประกาศลักษณะการทำงานของสัญพจน์นี้ยังสามารถระบุได้ถึงจำนวนครั้งที่ต้องการให้สัญพจน์นี้ปรากฏในกฎที่สร้างได้จากระบบอีกด้วย การสร้างกฎของระบบ PROGOL จะเริ่มตั้งแต่สุ่มตัวอย่าง สร้างอนุประโยคที่เฉพาะเจาะจงที่สุด (most-specific) สำหรับตัวอย่างนั้น แล้วทำการค้นหาด้วยวิธีการแบบ A* (A*-like search) [Nilsson, 1980] โดยมีจุดมุ่งหมายเพื่อให้ได้การบีบอัด (compression) สูงสุดในการค้นหานั้น ทำให้ระบบ PROGOL ใช้เวลาและเนื้อที่หน่วยความจำมากในการสร้างกฎเมื่อเทียบกับระบบอื่นๆ ที่ใช้ในการทดลองครั้งนี้

เนื่องจากระบบ PROGOL เป็นระบบไอแอลพีที่ทำงานกับปัญหาที่มีลักษณะเป็นสองกลุ่ม คือ กลุ่มตัวอย่างบวก และกลุ่มตัวอย่างลบ ดังนั้นในการสร้างกฎสำหรับปัญหาที่มีลักษณะเป็นหลายกลุ่ม จึงสร้างกฎโดยสร้างครั้งละกลุ่ม กำหนดให้ตัวอย่างบวกในการเรียนรู้คือตัวอย่างในกลุ่มนั้น และตัวอย่างลบคือ ตัวอย่างของกลุ่มอื่นๆ ทำซ้ำจนครบทุกกลุ่ม จะได้กฎสำหรับทุกๆ กลุ่ม

2. GOLEM

ระบบ GOLEM [Muggleton และ Feng, 1990] เป็นระบบไอแอลพีที่ทำงานกับปัญหาที่มีลักษณะเป็นสองกลุ่มเช่นเดียวกับระบบ PROGOL แต่แตกต่างกันในวิธีการค้นหาอนุประโยค โดยระบบ GOLEM รับอินพุตเป็นตัวอย่างบวก ตัวอย่างลบ และกลุ่มความรู้ภูมิหลัง ระบบเริ่มสร้างอนุประโยคด้วยการสุ่มเลือกตัวอย่างขึ้นมาเป็นคู่ หาอาร์แอลจีจีของตัวอย่างคู่นั้น เลือกผลการทำอาร์แอลจีจีที่ครอบคลุมตัวอย่างบวกมากที่สุด จากนั้นจึงวางนัยทั่วไปโดยนำอนุประโยคที่ได้มาทำอาร์แอลจีจีกับตัวอย่างบวกตัวใหม่ซึ่งสุ่มเลือกมาจากเซตของตัวอย่างบวกจนกระทั่งไม่สามารถหาอนุประโยคที่ครอบคลุมตัวอย่างได้มากขึ้น

ระบบ GOLEM อนุญาตให้ผู้ใช้ระบุจำนวนคู่ของตัวอย่างที่จะสุ่มมาทำอาร์แอลจีจีเพื่อหาอนุประโยคที่ครอบคลุมตัวอย่างมากที่สุดได้ ซึ่งจำนวนคู่ที่ระบบกำหนดไว้คือ 8 คู่ ถ้ากำหนดให้มีค่ามากขึ้น โอกาสที่ระบบจะพบอนุประโยคที่ครอบคลุมตัวอย่างได้มากจะสูงขึ้นตามไปด้วย ซึ่งในการทดลองครั้งนี้ได้ใช้ค่าที่ระบบกำหนดไว้ให้ คือ 8 คู่ในการสร้างกฎ และเนื่องจากระบบ GOLEM เป็นระบบที่ถูกสร้างขึ้นมาให้ใช้กับปัญหาที่เป็นสองกลุ่ม ดังนั้นในการสร้างกฎสำหรับกลุ่มใดๆ จะใช้ตัวอย่างของกลุ่มนั้นเป็นตัวอย่างบวกและใช้ตัวอย่างของกลุ่มอื่นๆ ที่เหลือเป็นตัวอย่างลบ เช่นเดียวกับขั้นตอนการสร้างกฎของระบบ PROGOL

3. TILDE

ระบบ TILDE จะทำการเรียนรู้เพื่อสร้างต้นไม้ตัดสินใจ (decision tree) ที่สามารถแทนกฎลำดับที่หนึ่งได้ หลักการทำงานของระบบจะคล้ายกับขั้นตอนวิธีการสร้างต้นไม้ตัดสินใจแบบทั่วไปที่ใช้กับค่าตรรกศาสตร์ประพจน์ โดยเริ่มจากสร้างโนดขึ้นมาทดสอบเพื่อเปรียบเทียบค่าฮิวริสติก (heuristic) ของแต่ละโนด ระบบจะเลือกโนดทดสอบที่มีค่าฮิวริสติกดีที่สุดในขณะนั้นมาสร้างเป็นโนดจริงในต้นไม้ จากนั้นจะนำตัวอย่างทั้งหมดมาทำการทดสอบกับโนดนี้เพื่อแบ่งตัวอย่างออกเป็นสองกลุ่ม คือ กลุ่มที่ให้ค่าความจริงเป็นจริงและให้ค่าความจริงเป็นเท็จเมื่อเทียบกับโนดปัจจุบัน จากนั้นจึงสร้างโนดทดสอบขึ้นมาใหม่สำหรับทดสอบกับตัวอย่างทั้งสองกลุ่มแล้วทำการทดสอบไปเรื่อยๆ จนกระทั่งมีตัวอย่างเพียงกลุ่มเดียวเท่านั้นที่ตรงกับโนดนั้น ระบบ TILDE นี้จะต่างจากระบบ PROGOL และ GOLEM ตรงที่ระบบ TILDE สามารถรู้จำตัวอย่างที่มีลักษณะเป็นหลายกลุ่มได้ เนื่องจากที่โนดใบ (leaf node) ของต้นไม้ตัดสินใจที่สร้างขึ้นจะแทนตัวอย่างแต่ละกลุ่ม เมื่อตัวอย่างไปตกอยู่ที่กลุ่มใด ก็จะถูกจำแนกเป็นกลุ่มนั้น

4. 1BC

1BC [Flach และ Lachiche, 1999] เป็นระบบที่รู้จำแบบเบย์ (Bayesian Classifier) สามารถใช้กับตรรกะลำดับที่หนึ่งได้ ระบบนี้ใช้แนวคิดจากวิธีการของเบย์กับตรรกศาสตร์ประพจน์ ซึ่งใช้ค่าทางสถิติของตัวอย่างมาจำแนกตัวอย่างใหม่ เมื่อนำแนวคิดของเบย์มาใช้กับตรรกะลำดับที่หนึ่งจึงต้องมีการปรับเปลี่ยนวิธีการรู้จำ ระบบ 1BC แบ่งลักษณะของสัญพจน์ออกเป็นสองกลุ่มด้วยกันคือ กลุ่มที่เป็นสัญพจน์โครงสร้าง และกลุ่มที่เป็นสัญพจน์ที่บอกลักษณะ ซึ่งสัญพจน์ทั้งสองกลุ่มนี้ผู้ใช้ต้องเป็นผู้กำหนดให้ในตอนเริ่มเรียนรู้ จากนั้น

ระบบจะสร้างลักษณะสำคัญจากสัญพจน์ทั้งสองกลุ่ม แล้วจึงเรียนรู้ด้วยขั้นตอนวิธีนาอีฟเบย์ (naive Bayes algorithm) นอกจากลักษณะของสัญพจน์ที่ผู้ใช้เป็นผู้กำหนดแล้ว ระบบ 1BC ยังอนุญาตให้ผู้ใช้กำหนดลักษณะของลักษณะสำคัญที่สร้างขึ้นได้ โดยระบุจำนวนสัญพจน์และจำนวนตัวแปรที่สามารถปรากฏได้ในลักษณะสำคัญข้อหนึ่งๆ ซึ่งในการทดลองครั้งนี้ใช้ค่าปกติของระบบคืออนุญาตให้มีจำนวนสัญพจน์และจำนวนตัวแปรเท่ากับ 3 ทั้งสองค่า แต่ในการทดลองกับปัญหาการวิเคราะห์ไฟไนต์เอลิเมนต์เกิดปัญหาหน่วยความจำไม่พอ เนื่องจากมีสัญพจน์ที่มีค่าทั้งสองเป็น 3 อยู่เป็นจำนวนมาก จึงกำหนดให้ค่าทั้งสองมีค่าเป็น 2

5. LINUS

ระบบ LINUS [Lavrac และคณะ, 1991] เป็นระบบไอแอลพีที่สามารถใช้กับปัญหาที่มีลักษณะเป็นหลายกลุ่มและมีเป็นตรรกะลำดับที่หนึ่งได้ โดยระบบจะเปลี่ยนรูปแบบของปัญหาจากตรรกะลำดับที่หนึ่งให้เป็นค่าคุณสมบัติ (attribute value) การทำงานของระบบจะเริ่มโดยการสร้างลักษณะสำคัญจากตัวอย่างและความรู้ภูมิหลังซึ่งอยู่ในรูปตรรกะลำดับที่หนึ่ง ใช้ลักษณะสำคัญเหล่านั้นเป็นค่าคุณสมบัติของตัวอย่างแต่ละตัว จากนั้นจึงใช้ระบบที่สามารถรู้จำตัวอย่างจากค่าคุณสมบัติได้ เช่น ASSISTANT, CN2, NEWGEM และ C4.5 ฯลฯ ในการทดลองเปรียบเทียบครั้งนี้ได้สร้างลักษณะสำคัญโดยระบบ LINUS แล้วนำลักษณะสำคัญซึ่งอยู่ในรูปค่าคุณสมบัติไปทำการรู้จำด้วยระบบ C4.5 เนื่องจากระบบ C4.5 สามารถจำแนกตัวอย่างซึ่งมีลักษณะเป็นหลายกลุ่มได้

6. FOSSIL

FOSSIL [Furnkranz, 1994] เป็นระบบไอแอลพีระบบหนึ่งที่สามารถใช้กับปัญหาที่มีสัญญาณรบกวน (noisy domain) ได้โดยใช้ค่าฮิวริสติกซึ่งวัดจากค่าสหสัมพันธ์ทางสถิติ (statistical correlation) ซึ่งเป็นวิธีการที่แตกต่างจากระบบโดยทั่วไปที่ใช้แนวคิดจากขั้นตอนวิธีการเรียนรู้ต้นไม้ตัดสินใจ (decision tree learning algorithm) มาจัดการกับสัญญาณรบกวนในเซตของตัวอย่าง ในการทดลองครั้งนี้ได้นำระบบ FOSSIL มาทำการทดลองเปรียบเทียบกับระบบ BANNAR ในกลุ่มปัญหาที่มีสัญญาณรบกวน

ชุดข้อมูลที่ใช้ในการทดลอง

ชุดข้อมูลหรือกลุ่มของปัญหาที่ใช้ในการทดลองครั้งนี้มี 4 กลุ่ม คือ การรู้จำภาพตัวพิมพ์อักษรไทย การวิเคราะห์ไฟไนต์เอลิเมนต์ การวิเคราะห์ความสามารถก่อกลายพันธุ์ และการวิเคราะห์ตำแหน่งตัวหมากรุก รายละเอียดของกลุ่มปัญหาทั้งหมดเป็นดังนี้

1. การรู้จำภาพตัวพิมพ์อักษรไทย (Thai Optical Character Recognition: TCR)

กลุ่มตัวอย่างในข้อมูลชุดนี้ประกอบด้วย พยัญชนะ สระ วรรณยุกต์ และตัวเลขไทย รวมทั้งหมดเป็นตัวพิมพ์อักษรไทย 77 ตัวอักษร ใช้ตัวอักษร 2 แบบ 7 ขนาด รวมเป็นชุดตัวอย่างที่ใช้เรียนรู้ทั้งสิ้น 1,078 ตัว ภาพตัวอย่างถูกสแกนด้วยความละเอียด 300 dpi จากนั้นนำมาผ่านกระบวนการกำจัดสัญญาณรบกวน

กระบวนการทำภาพให้บางเพื่อหาเวกเตอร์พื้นฐานซึ่งประกอบกันเป็นภาพตัวอักษรนั้น นำเวกเตอร์พื้นฐาน บริเวณที่เป็นจุดร่วมของเส้น บริเวณที่เป็นรอยหยักขึ้นและรอยหยักลง จัดเป็นตัวอย่างที่ใช้เรียนรู้ กลุ่มความรู้ภูมิหลังที่ใช้ในการทดลองประกอบด้วยกฎซึ่งอยู่ในรูปกฎลำดับที่หนึ่งทั้งหมด 55 ข้อ กฎแต่ละข้อเป็นลักษณะของเส้นที่ใช้แบ่งแยกตัวอักษร เช่น `headzone(A,B)` คือ ตัวอักษรที่มีเวกเตอร์พื้นฐาน A มีส่วนหัวอยู่บริเวณ B หรือ `headprim(A,B)` หมายถึง ส่วนหัวของภาพตัวอักษรที่มีเวกเตอร์พื้นฐาน A มีเวกเตอร์พื้นฐาน B เป็นส่วนหัวของตัวอักษร ฯลฯ ซึ่งรายละเอียดของตัวอย่างที่ใช้เรียนรู้และความรู้ภูมิหลังปรากฏอยู่ใน [Kijsirikul, B. และ Sinthupinyo, S., 1999] กลุ่มตัวอย่างที่ใช้ทดสอบเป็นภาพตัวอักษรกลุ่มเดียวกับตัวอย่างที่ใช้เรียนรู้ ซึ่งได้รับการเพิ่มสัญญาณรบกวนโดยนำไปถ่ายเอกสารแบบเข้มและแบบจาง จากนั้นจึงนำไปผ่านกระบวนการเช่นเดียวกับตัวอย่างที่ใช้เรียนรู้ นำเวกเตอร์พื้นฐาน บริเวณที่เป็นจุดร่วมของเส้น บริเวณที่เป็นรอยหยักขึ้นและรอยหยักลง มาจัดเป็นตัวอย่างเพื่อใช้ทดสอบ

2. การวิเคราะห์ไฟไนต์เอลิเมนต์ (Finite Element Mesh Design: FEM)

จุดมุ่งหมายของปัญหา FEM [Dolsak และ Muggleton, 1992] คือ การสร้างกฎเพื่อวิเคราะห์ไฟไนต์เอลิเมนต์ในโครงสร้าง โดยอาศัยกลุ่มความรู้ภูมิหลังเป็นลักษณะต่างๆ ของโครงสร้าง ประกอบด้วยลักษณะของเส้นเชื่อม เช่น `long`, `short` และ `usual` ฯลฯ เงื่อนไขขอบเขต เช่น `free` และ `one_side_fixed` ฯลฯ โลด เช่น `cont_loaded` และ `one_sideloaded` ฯลฯ กลุ่มตัวอย่างประกอบด้วยโครงสร้าง 5 แบบ แบ่งตัวอย่างออกเป็น 13 กลุ่ม แต่ละกลุ่มคือ จำนวนองค์ประกอบ (element) ที่เหมาะสมของโครงสร้างนั้น โดยตัวอย่างแต่ละตัวอย่างถูกจัดอยู่ในรูปแบบ `mesh(Edge, Number)` เมื่อ `Edge` คือ ชื่อโครงสร้าง และ `Number` คือ จำนวนองค์ประกอบภายในโครงสร้างนั้น รวมจำนวนตัวอย่างทั้งหมด 278 ตัวอย่าง

3. การวิเคราะห์ความสามารถก่อกลายพันธุ์ (Mutagenesis: MUTA)

จุดมุ่งหมายของปัญหา MUTA [Srinivasan และ คณะ, 1996] คือ การสร้างกฎเพื่อวิเคราะห์ความสามารถก่อกลายพันธุ์ของโมเลกุล โดยอาศัยกลุ่มความรู้ภูมิหลังเป็นลักษณะต่างๆ ภายในโมเลกุลนั้นประกอบด้วยลักษณะของอะตอมและลักษณะโครงสร้างของโมเลกุล เช่น `benzene`, `carbon_6_ring`, `carbon_5_aromatic_ring`, `hetero_aromatic_6_ring`, `hetero_aromatic_5_ring`, `ring_size_6`, `ring_size_5`, `nitro`, `methyl`, `anthracene`, `phenanthrene` และ `ball3` ฯลฯ กลุ่มตัวอย่างประกอบด้วยตัวอย่างที่เป็นข้อมูลของโมเลกุลทั้งหมด 188 โมเลกุล แบ่งเป็น 125 โมเลกุลอยู่ในกลุ่มตัวอย่างบวก และ 63 โมเลกุลอยู่ในกลุ่มตัวอย่างลบ

4. การวิเคราะห์ตำแหน่งตัวหมากรุกสากล (King-Rook-King Chess Endgame: KRK)

ในข้อมูลชุด KRK [Muggleton และคณะ, 1989] เป็นการวิเคราะห์ตำแหน่งของตัวหมากบนกระดานหมากรุกสากลที่ไม่สามารถเกิดขึ้นได้ในขณะที่ฝ่ายสีขาวเป็นผู้เดินและมีหมากเหลืออยู่บนกระดาน 3 ตัว คือ ขุนฝ่ายขาว เรือฝ่ายขาว และ ขุนฝ่ายดำ ตัวอย่างที่ใช้อยู่ในรูป `illegal(WKf, WKr, WRf,`

WRr, BKf, BKr) เมื่อ WKf, WKr, WRf, WRr, BKf และ BKr คือ แถว (file) ของขุนฝ่ายขาว หลัก (rank) ของขุนฝ่ายขาว แถวของเรือฝ่ายขาว หลักของเรือฝ่ายขาว แถวของขุนฝ่ายดำ และหลักของขุนฝ่ายดำ กลุ่มความรู้ภูมิหลังคือความสัมพันธ์ของตำแหน่งบนกระดาน ในชุดข้อมูลนี้ใช้ความสัมพันธ์ 2 แบบ คือ adj(X,Y) และ lt(X,Y) ซึ่งแสดงถึงตำแหน่ง X และ Y ที่อยู่ติดกัน และตำแหน่งที่มีค่าน้อยกว่ากัน ตามลำดับ จำนวนตัวอย่างในข้อมูลชุดนี้ประกอบด้วยตัวอย่างทั้งหมด 10,000 ตัวอย่าง แบ่งเป็นตัวอย่างบวก 3,361 ตัว และตัวอย่างลบ 6,639 ตัว

ผลการทดลอง

1. การทดลองเปรียบเทียบกับระบบอื่น

การทดลองกับกลุ่มตัวอย่าง FEM, MUTA และ KRK ใช้การทดลองโดยแบ่งกลุ่มตัวอย่างออกเป็น 3 กลุ่มแบบสุ่ม จากนั้นทำการทดลองโดยใช้กลุ่มตัวอย่างหนึ่งกลุ่มเป็นกลุ่มตัวอย่างที่ใช้ทดสอบและใช้กลุ่มตัวอย่างที่เหลืออีกสองกลุ่มเป็นกลุ่มตัวอย่างที่ใช้เรียนรู้ (3-fold cross-validation: 3CV) สำหรับในกลุ่มตัวอย่าง TCR ใช้การทดสอบโดยสร้างกลุ่มตัวอย่างที่ใช้ทดสอบจากกลุ่มตัวอย่างที่ใช้เรียนรู้ แต่เพิ่มสัญญาณรบกวนโดยผ่านการถ่ายเอกสารแบบเข้มและแบบจาง การทดลองเริ่มด้วยสร้างกฎจากตัวอย่างที่ใช้เรียนรู้ด้วยระบบ PROGOL หรือ GOLEM ตามกลุ่มตัวอย่าง โดยในกลุ่ม FEM และ KRK ใช้ระบบ GOLEM ในการสร้างกฎ ส่วนในกลุ่ม TCR และ MUTA ใช้ระบบ PROGOL ในการสร้างกฎ จากนั้นนำกฎที่ได้มาหาลักษณะสำคัญเพื่อนำไปเทียบกับตัวอย่างสำหรับเรียนรู้และทดสอบแล้วสร้างเป็นอินพุตเวกเตอร์และเอาต์พุตเวกเตอร์นำอินพุตเวกเตอร์และเอาต์พุตเวกเตอร์มาเรียนรู้และทดสอบ เวลาที่ใช้ในการเรียนรู้โดยระบบ BANNAR เมื่อเทียบกับระบบ PROGOL หรือ GOLEM ซึ่งใช้ในการสร้างกฎเป็นดังตารางที่ 4.1 และผลการทดลองเปรียบเทียบระหว่างระบบ BANNAR กับระบบอื่น เป็นดังตารางที่ 4.2

จากตารางที่ 4.1 พบว่าเวลาที่ใช้ในกระบวนการเรียนรู้ทั้งหมดของระบบ BANNAR น้อยมากเมื่อเทียบกับเวลาที่ใช้ในกระบวนการเรียนรู้ของ PROGOL และ GOLEM โดยเฉพาะอย่างยิ่งในชุดข้อมูล TCR เนื่องจาก TCR จะต้องสร้างกฎเพื่อให้ครอบคลุมตัวอย่างทั้งหมด 77 กลุ่ม จึงทำให้ PROGOL ต้องใช้เวลาในการสร้างกฎนานมาก กระบวนการเรียนรู้ของ BANNAR จะแบ่งออกเป็น 2 ช่วง คือ ช่วงที่เป็นการดึงลักษณะสำคัญจากกฎแล้วนำมาเทียบกับตัวอย่างที่ใช้เรียนรู้ กับช่วงที่ใช้ในกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์ก เวลาที่ใช้ในการเทียบตัวอย่างกับลักษณะสำคัญจะแปรผันตรงกับจำนวนลักษณะสำคัญและจำนวนตัวอย่างที่ใช้เรียนรู้ ดังนั้นเวลาที่ใช้ในการเรียนรู้สำหรับชุดข้อมูล TCR และ KRK จึงมากกว่าเวลาที่ใช้ในชุดข้อมูลอื่น

เนื่องจากระบบ PROGOL และ GOLEM เป็นระบบไอแอลพีที่ใช้กับปัญหาที่มีลักษณะเป็นสองกลุ่ม ดังนั้นเมื่อนำไปใช้กับปัญหาที่มีลักษณะเป็นหลายกลุ่ม และตัวอย่างที่ใช้ทดสอบไม่ตรงกับกฎข้อใดในเซตของ

ตารางที่ 4.1 ตารางแสดงจำนวนลักษณะสำคัญ จำนวนกฎ และเวลาที่ใช้ในการเรียนรู้โดยระบบ BANNAR เทียบกับระบบ PROGOL หรือ GOLEM

ชุดข้อมูล	BANNAR		PROGOL หรือ GOLEM	
	จำนวนลักษณะสำคัญ	เวลา (วินาที)	จำนวนกฎ	เวลา (วินาที)
TCR	467.0	3224.0	77.0	134112.9
FEM	115.0	72.0	38.0	3607.9
MUTA	18.3	57.0	6.3	6014.3
KRK	39.0	191.0	15.3	371.1

หมายเหตุ จำนวนลักษณะสำคัญ จำนวนกฎและเวลาที่ใช้ในชุดข้อมูล FEM, MUTA และ KRK เป็นค่าเฉลี่ยในการเรียนรู้ของทั้งสามกลุ่ม ส่วนในชุดข้อมูล TCR เป็นจำนวนลักษณะสำคัญ จำนวนกฎ และเวลาที่ใช้เรียนรู้ของกลุ่มตัวอย่างที่ใช้เรียนรู้ เวลาที่ใช้วัดจากเครื่องคอมพิวเตอร์ซึ่งมีซีพียู (CPU) Pentium II ความเร็ว 400 MHz ระบบ BANNAR ทำงานบนระบบปฏิบัติการ Windows 98 ส่วนระบบ PROGOL หรือ GOLEM ทำงานบนระบบปฏิบัติการ Linux

ตารางที่ 4.2 ตารางแสดงผลการทดลองเปรียบเทียบระหว่าง BANNAR กับระบบไอแอลพีอื่นๆ

ชุดข้อมูล	จำนวนตัวอย่างที่ใช้เรียนรู้	จำนวนตัวอย่างที่ใช้ทดสอบ	จำนวนกลุ่ม	BANNAR	PROGOL หรือ GOLEM	TILDE	1BC	LINUS
TCR	1,074	2,143	77	94.40	72.00[3]	88.57[3]	77.23[3]	66.54[3]
FEM	278	3CV	13	64.45	57.80[2]	58.02[2]	46.73[2]	60.45[1]
MUTA	188	3CV	2	83.58	82.01[0]	68.94[1]	77.72[0]	74.41[1]
KRK	10,000	3CV	2	99.93	99.89[0]	69.80[3]	87.12[3]	99.89[0]

หมายเหตุ ใช้การทดสอบค่าทีแบบทางเดียว (one-tailed paired t-test) ตัวเลขยกแสดงระดับความเชื่อมั่น (confidence level) ตัวเลข 0, 1, 2 และ 3 แสดงถึงระดับความเชื่อมั่นที่ ต่ำกว่า 90.0%, 90.0%, 99.0% และ 99.5% ตามลำดับ

กฎ PROGOL และ GOLEM ไม่สามารถทำการจำแนกตัวอย่างนั้นได้ จึงต้องอาศัยวิธีการอื่นมาใช้จำแนกตัวอย่างเหล่านั้น ในการทดลองครั้งนี้ได้เลือกใช้วิธีการจำแนกตามกลุ่มหลักมาใช้ในกรณีดังกล่าว ผลการทดลองที่ปรากฏในตารางที่ 4.2 ในส่วนของ PROGOL และ GOLEM เป็นผลการทดลองที่ใช้วิธีการจำแนกตามกลุ่มหลักมาใช้ โดยในกรณีของปัญหาที่เป็นสองกลุ่ม คือ MUTA และ KRK ตัวอย่างที่ไม่ตรงกับกฎข้อใดเลย จะถูกจำแนกเป็นกลุ่มลบ ตัวอย่างที่ตรงกับกฎตั้งแต่หนึ่งข้อขึ้นไป จะถูกจำแนกว่าเป็นตัวอย่างบวก แต่ในกรณีปัญหาที่เป็นหลายกลุ่ม จะเลือกกลุ่มโดยใช้วิธีจำแนกตามกลุ่มหลัก กล่าวคือ เมื่อตัวอย่างที่ใช้

ทดสอบไม่ตรงกับกฎข้อใดในเซตของกลุ่ม จะทำการจำแนกเป็นกลุ่มที่มีจำนวนมากที่สุดในกลุ่มตัวอย่างที่ใช้เรียนรู้

ในส่วนของระบบ LINUS ที่นำมาใช้ทดลองเปรียบเทียบในการทดลองครั้งนี้ ใช้ระบบ LINUS สร้างลักษณะสำคัญของชุดตัวอย่าง จากนั้นนำลักษณะสำคัญที่ได้ซึ่งอยู่ในรูปตรรกศาสตร์ประพจน์ไปเรียนรู้และทดสอบด้วยระบบ C4.5

ผลการทดลองในตารางที่ 4.2 แสดงให้เห็นว่าเปอร์เซ็นต์ความถูกต้องของระบบ PROGOL หรือ GOLEM เมื่อนำไปใช้กับปัญหาที่มีลักษณะเป็นสองกลุ่ม คือ MUTA และ KRK ให้ผลการรู้จำสูงกว่า TILDE ในขณะที่เมื่อนำไปใช้กับปัญหาที่มีลักษณะเป็นหลายกลุ่ม คือ TCR และ FEM TILDE ซึ่งโดยปกติจะใช้งานกับปัญหาที่มีลักษณะเป็นหลายกลุ่มอยู่แล้ว ให้ผลการรู้จำสูงกว่า PROGOL หรือ GOLEM เมื่อเทียบผลการรู้จำของ PROGOL หรือ GOLEM กับ 1BC พบว่า PROGOL หรือ GOLEM ให้ผลการรู้จำสูงกว่า 1BC ในชุดข้อมูล FEM, MUTA และ KRK ให้ผลการรู้จำต่ำกว่า 1BC ในชุดข้อมูล TCR เทียบผลการรู้จำของ PROGOL หรือ GOLEM กับระบบ LINUS ปรากฏว่า PROGOL หรือ GOLEM ให้ผลการรู้จำสูงกว่า LINUS ในชุดข้อมูล TCR และ MUTA ให้ผลการรู้จำเท่ากันในชุดข้อมูล KRK และสูงกว่าในชุดข้อมูล TCR

จากผลการทดลองเมื่อเทียบเปอร์เซ็นต์ความถูกต้องของ PROGOL หรือ GOLEM กับระบบอื่น 3 ระบบ (TILDE, 1BC และ LINUS) ใน 4 ชุดข้อมูล รวมเปรียบเทียบ 12 การทดลอง พบว่า ในกลุ่มปัญหาที่มีลักษณะเป็นหลายกลุ่ม (TCR และ FEM) PROGOL หรือ GOLEM ให้ผลการรู้จำสูงกว่าระบบอื่น 2 จาก 6 การทดลอง ต่ำกว่า 4 จาก 6 การทดลอง และเมื่อเปรียบเทียบในกลุ่มปัญหาที่มีลักษณะเป็นสองกลุ่ม (MUTA และ KRK) PROGOL หรือ GOLEM ให้ผลการรู้จำสูงกว่าระบบอื่น 5 จาก 6 การทดลอง และให้ผลการรู้จำเท่ากัน 1 จาก 6 การทดลอง จะเห็นได้ว่า กฎที่ได้จากระบบ PROGOL หรือ GOLEM จะสามารถใช้ได้ดีในกลุ่มปัญหาที่มีลักษณะเป็นสองกลุ่ม และใช้ได้ไม่ดีในกลุ่มปัญหาที่มีลักษณะเป็นหลายกลุ่ม เมื่อนำวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์กมาใช้ในระบบ BANNAR ทำให้เปอร์เซ็นต์ความถูกต้องสูงขึ้น และเมื่อเทียบเปอร์เซ็นต์ความถูกต้องของ BANNAR กับระบบอื่น พบว่า BANNAR ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าระบบอื่นทุกระบบในทุกชุดข้อมูล และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าแบบมีนัยสำคัญทางสถิติเมื่อเทียบกับระบบ PROGOL หรือ GOLEM ใน 2 ชุดข้อมูล (TCR และ FEM) ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าแบบมีนัยสำคัญทางสถิติเมื่อเทียบกับ TILDE ในทุกชุดข้อมูล ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าแบบมีนัยสำคัญทางสถิติเมื่อเทียบกับ 1BC ใน 3 ชุดข้อมูล (TCR, FEM และ KRK) และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าแบบมีนัยสำคัญทางสถิติเมื่อเทียบกับ LINUS ใน 3 ชุดข้อมูล (TCR, FEM และ MUTA)

ตัวเลขในตารางที่ 4.3 ของระบบ PROGOL หรือ GOLEM ในกรณีที่ตัวอย่างไม่ถูกครอบคลุมด้วยกฎในชุดข้อมูลที่เป็นหลายกลุ่ม จะใช้วิธีการจำแนกตามกลุ่มหลัก ส่วนกรณีที่ตัวอย่างไม่ถูกครอบคลุมด้วยกฎ

ตารางที่ 4.3 ตารางแสดงจำนวนตัวอย่างที่ถูกจำแนกถูกต้องเปรียบเทียบระหว่าง BANNAR กับ PROGOL หรือ GOLEM แยกกรณีที่ตัวอย่างถูกครอบคลุมด้วยกฎ และไม่ถูกครอบคลุมด้วยกฎ

ชุดข้อมูล	จำนวนตัวอย่างทดสอบ	BANNAR		PROGOL หรือ GOLEM	
		ครอบคลุม	ไม่ครอบคลุม	ครอบคลุม	ไม่ครอบคลุม
TCR	2,143	1570/1611	453/532	1540/1611	3/532
FEM	278	145/200	34/78	146/200	14/78
MUTA	188	102/109	55/79	100/109	54/79
KRK	10,000	3356/3360	6637/6640	3355/3360	6634/6640

หมายเหตุ ตัวเลขในตารางแสดงถึง จำนวนตัวอย่างที่ถูกจำแนกถูกต้อง/จำนวนตัวอย่างทั้งหมดในกลุ่มนั้น

ในชุดข้อมูลที่เป็นสองกลุ่ม จะจำแนกเป็นกลุ่มตัวอย่างลบ ผลการทดลองเปรียบเทียบระหว่าง PROGOL หรือ GOLEM กับ BANNAR พบว่าในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม BANNAR สามารถจำแนกตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎได้ถูกต้องกว่า PROGOL หรือ BANNAR ในชุดข้อมูล TCR BANNAR สามารถจำแนกตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎได้ถูกต้อง 453 ตัวอย่างจากตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎทั้งหมด 532 ตัวอย่าง ในขณะที่ใช้วิธีการจำแนกตามกลุ่มหลักสามารถจำแนกตัวอย่างได้ถูกต้องเพียง 3 ตัวอย่าง จาก 532 ตัวอย่าง ในส่วนของตัวอย่างที่ถูกครอบคลุมด้วยกฎ BANNAR สามารถจำแนกได้ถูกต้อง 1570 ตัวอย่าง จาก 1,611 ตัวอย่าง ในขณะที่กฎจาก PROGOL สามารถจำแนกได้ถูกต้อง 1,540 ตัวอย่าง จาก 1,611 ตัวอย่าง เช่นเดียวกับในชุดข้อมูล FEM ซึ่งมีลักษณะเป็นหลายกลุ่มเช่นกัน BANNAR สามารถจำแนกตัวอย่างในกรณีที่ตัวอย่างไม่ถูกครอบคลุมด้วยกฎถูกต้อง 34 ตัวอย่างจาก 78 ตัวอย่าง ในขณะที่เมื่อใช้วิธีการจำแนกตามกลุ่มหลักสามารถจำแนกได้ถูกต้อง 14 ตัวอย่าง จาก 78 ตัวอย่าง และทั้ง BANNAR และกฎจาก GOLEM สามารถจำแนกตัวอย่างได้ถูกต้องใกล้เคียงกัน คือ BANNAR จำแนกได้ถูกต้อง 145 ตัวอย่าง และกฎจาก GOLEM สามารถจำแนกได้ถูกต้อง 146 ตัวอย่าง จาก 200 ตัวอย่าง สำหรับในชุดข้อมูลที่มีลักษณะเป็นสองกลุ่ม คือ ชุดข้อมูล MUTA และ KRK ทั้ง BANNAR และ PROGOL หรือ GOLEM ให้ผลการรู้จำใกล้เคียงกัน จะเห็นได้จากตารางที่ 4.3 ว่า BANNAR สามารถเพิ่มเปอร์เซ็นต์ความถูกต้องได้ในกรณีที่ตัวอย่างไม่ถูกครอบคลุมด้วยกฎในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม ซึ่งระบบ PROGOL หรือ GOLEM ไม่สามารถจำแนกตัวอย่างเหล่านี้ได้

2. ผลการทดลองในชุดข้อมูล KRK ที่ได้รับการเพิ่มสัญญาณรบกวน

จากตารางที่ 4.3 พบว่าในชุดข้อมูลที่มีลักษณะเป็นสองกลุ่ม BANNAR จะให้ผลการรู้จำใกล้เคียงกับ PROGOL หรือ GOLEM จึงได้ทำการทดลองในชุดข้อมูล KRK โดยเพิ่มสัญญาณรบกวนที่ระดับ 5%, 10% และ 15% ซึ่งเมื่อเพิ่มสัญญาณรบกวนเป็น 20% ไม่สามารถทำการทดลองเปรียบเทียบได้ เนื่องจากระบบ GOLEM สร้างกฎไม่ได้ จึงรายงานผลการทดลองที่ระดับสัญญาณรบกวนสูงสุด 15% ทำการทดลอง

เปรียบเทียบระหว่าง BANNAR กับ GOLEM และทำการทดลองเปรียบเทียบกับระบบ FOSSIL ซึ่งสามารถแก้ปัญหาในลักษณะที่มีสัญญาณรบกวนได้ ผลการทดลองเป็นดังตารางที่ 4.4

ตารางที่ 4.4 ตารางแสดงเปอร์เซ็นต์ความถูกต้องของชุดข้อมูล KRK เมื่อมีการเพิ่มสัญญาณรบกวนที่ระดับ 5%, 10% และ 15%

ระดับสัญญาณรบกวน	GOLEM		BANNAR		FOSSIL
	จำนวนกฎ	เปอร์เซ็นต์ความถูกต้อง	จำนวนลักษณะสำคัญ	เปอร์เซ็นต์ความถูกต้อง	เปอร์เซ็นต์ความถูกต้อง
0%	15.3	99.89	39.0	99.93	99.97
5%	104.0	91.89	451.0	99.28	98.93
10%	99.7	81.98	475.0	98.33	97.10
15%	95.0	76.86	487.0	95.87	94.68

หมายเหตุ ทุกระบบทำการทดลองแบบ 3CV จำนวนกฎ จำนวนลักษณะสำคัญ และเปอร์เซ็นต์ความถูกต้องเป็นค่าเฉลี่ยจากทั้งสามกลุ่ม

จากตารางที่ 4.4 การเพิ่มสัญญาณรบกวนในชุดข้อมูล KRK ที่ระดับ n% หมายความว่า จากกลุ่มตัวอย่างที่ใช้เรียนรู้ทั้งหมด ทำการสุ่มเลือกมา n% จากนั้นเปลี่ยนกลุ่มของตัวอย่างที่เลือกได้ ให้เป็นตรงข้าม จากกลุ่มตัวอย่างบวกเปลี่ยนเป็นกลุ่มตัวอย่างลบ และจากกลุ่มตัวอย่างลบเปลี่ยนเป็นกลุ่มตัวอย่างบวก จากนั้นทำการทดลองแบบ 3CV ในทุกระบบ ผลการทดลองที่ได้จะเห็นว่า ระบบ GOLEM ให้เปอร์เซ็นต์ความถูกต้องลดลงเร็วที่สุดเมื่อระดับสัญญาณรบกวนเพิ่มขึ้น ในขณะที่ระบบ BANNAR และ FOSSIL ให้เปอร์เซ็นต์ความถูกต้องลดลงช้ากว่า เมื่อเปรียบเทียบผลการทดลองในแต่ระดับสัญญาณรบกวนพบว่า ระบบ GOLEM ให้เปอร์เซ็นต์ความถูกต้องต่ำกว่าระบบ BANNAR ทุกระดับสัญญาณรบกวน โดยมีระดับความเชื่อมั่น 97.5% ในระดับสัญญาณรบกวน 5% และระดับความเชื่อมั่น 99% ในระดับสัญญาณรบกวน 10% และ 15% ในขณะที่เมื่อเปรียบเทียบระหว่างระบบ BANNAR กับระบบ FOSSIL พบว่าระบบ BANNAR มีอัตราการรู้จำสูงกว่าในระดับสัญญาณรบกวน 5%, 10% และ 15% โดยมีระดับความเชื่อมั่น 99% ในระดับสัญญาณรบกวน 10% ระดับความเชื่อมั่นต่ำกว่า 90% ในระดับสัญญาณรบกวน 5% และ 15% และระบบ BANNAR มีเปอร์เซ็นต์ความถูกต้องต่ำกว่าในกลุ่มที่ไม่มีการเพิ่มสัญญาณรบกวน จากผลการทดลองจะเห็นว่าระบบ BANNAR สามารถเพิ่มความสามารถในการจัดการกับสัญญาณรบกวนให้กับกฎที่ได้จากระบบ GOLEM ได้ และให้เปอร์เซ็นต์ความถูกต้องใกล้เคียงกับระบบ FOSSIL ซึ่งมีความสามารถในการแก้ปัญหาในกลุ่มตัวอย่างซึ่งมีสัญญาณรบกวนอยู่แล้ว

เมื่อพิจารณาถึงวิธีการที่ใช้แก้ปัญหาสัญญาณรบกวนของระบบ BANNAR กับระบบ FOSSIL พบว่า ระบบ FOSSIL จะใช้วิธีฮิวริสติกเพื่อป้องกันไม่ให้กฎที่ได้เฉพาะเจาะจง (specific) กับตัวอย่างที่ใช้เรียนรู้มากเกินไป ในขณะที่ BANNAR ใช้ค่าน้ำหนักของเส้นเชื่อมภายในนิวรอลเน็ตเวิร์กเพื่อบ่งบอกความสำคัญของลักษณะสำคัญแต่ละข้อของกฎ จากผลการทดลองในตารางที่ 4.4 จะเห็นได้ว่า ถึงแม้ในระดับสัญญาณ

รบกวนต่างๆ ที่ระบบ GOLEM สร้างกฎที่เฉพาะเจาะจงมากเกินไป และให้เปอร์เซ็นต์ความถูกต้องลดลง แต่ระบบ BANNAR สามารถสร้างนิวรอลเน็ตเวิร์กที่ไม่เฉพาะเจาะจงกับตัวอย่างที่ใช้เรียนรู้ และให้เปอร์เซ็นต์ความถูกต้องลดลงช้ากว่าระบบ GOLEM ซึ่งตัวอย่างของค่าน้ำหนักของลักษณะสำคัญที่ได้จากกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์กในการทดลองกลุ่มแรกของระดับสัญญาณรบกวน 5% เป็นดังนี้

จากกฎ

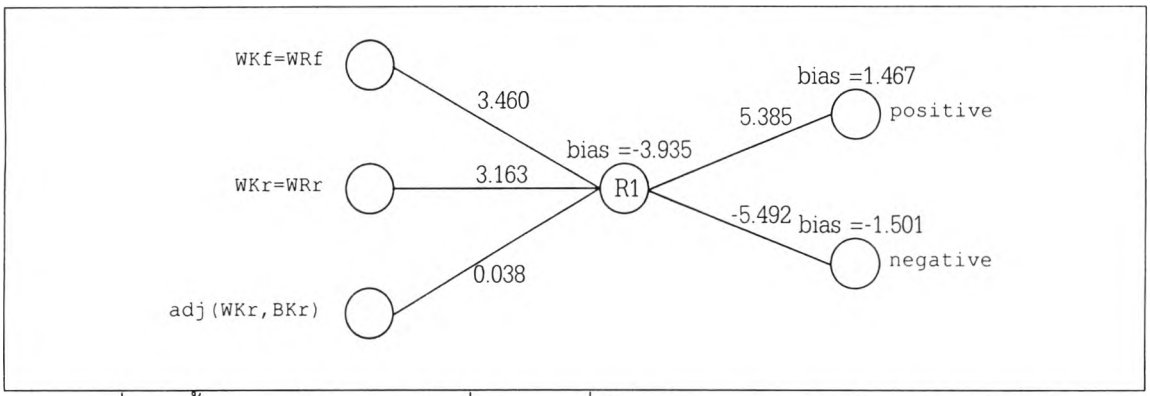
```
illegal(WKf,WKr,WRf,WRr,BKf,BKr) :- WKf=WRf, WKr=WRr, adj(WKr,BKr).
```

เมื่อผ่านกระบวนการดึงลักษณะสำคัญจะได้ลักษณะสำคัญ 3 ข้อ ดังนี้

(1) `WKr=WRf`

(2) `WKr=WRr`

(3) `adj(WKr,BKr)`



รูปที่ 4.1 น้ำหนักและไบแอสของเส้นเชื่อมของกฎซึ่งได้จากกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์ก

หลังจากผ่านกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์กจะได้ค่าน้ำหนักของลักษณะสำคัญแต่ละข้อได้ดังรูปที่ 4.1 ค่าน้ำหนักของเส้นเชื่อมจากโนดซึ่งแทนลักษณะสำคัญ `WKf=WRf`, `WKr=WRr` และ `adj(WKr,BKr)` มีค่าเป็น 3.460, 3.163 และ 0.038 ตามลำดับ และค่าไบแอสของโนดซึ่งแทนกฎข้อนั้นมีค่าเท่ากับ -3.935 จะเห็นว่าในขณะที่ลักษณะสำคัญ `WKf=WRf` และ `WKr=WRr` มีค่าความจริงเป็นจริง จะทำให้ได้ผลรวมของลักษณะสำคัญทั้งสองข้อนี้ที่โนด R1 มีค่าเท่ากับ -3.935+(3.460+3.163) = 2.688 เมื่อนำไปรวมกับลักษณะสำคัญ `adj(WKr,BKr)` ในกรณีที่ลักษณะสำคัญ `adj(WKr,BKr)` มีค่าความจริงเป็นจริง จะทำให้ผลรวมของโนด R1 มีค่าเท่ากับ 2.688+0.038 = 2.726 และในกรณีที่ ลักษณะสำคัญ `adj(WKr,BKr)` มีค่าความจริงเป็นเท็จ จะทำให้ผลรวมของโนด R1 มีค่าเท่ากับ 2.688+(-0.038) = 2.650 ซึ่งเมื่อผ่านฟังก์ชันซิกมอยด์จะให้ค่าเอาต์พุตที่โนด R1 เท่ากับ 0.939 และ 0.934 ตามลำดับ ค่าที่ผ่านฟังก์ชันซิกมอยด์ทั้งสองค่ามีค่าใกล้เคียงกัน จะเห็นได้ว่าค่าน้ำหนักที่ได้จากนิวรอลเน็ตเวิร์กจะเป็นตัวกำหนดความสำคัญให้กับลักษณะสำคัญแต่ละข้อ จากตัวอย่างนี้พบว่า ลักษณะสำคัญ `WKf=WRf` และ `WKr=WRr` มีความสำคัญมากกว่าลักษณะสำคัญ `adj(WKr,BKr)` เมื่อพิจารณาที่ความหมายของกฎ และเมื่อสัญพจน์ `WKf=WRf` และ `WKr=WRr` มีค่าความจริงเป็นจริง หมายถึง ตำแหน่งของขุนฝ่ายขาวตรงกับตำแหน่งของเรือ

ฝ่ายขาวทั้งในแนวแถวและหลัก ซึ่งเพียงพอแล้วที่จะจำแนกตัวอย่างนี้เป็นกลุ่มบวกได้โดยไม่ต้องตรวจสอบค่าความจริงของสัญพจน์ `adj(wKr,BKr)` เนื่องจากหมากสองตัวไม่สามารถอยู่ในตำแหน่งเดียวกันได้

3. ผลการทดลองเปรียบเทียบระหว่างลักษณะสำคัญจากระบบ BANNAR กับระบบ LINUS

เพื่อให้เห็นถึงประสิทธิภาพของลักษณะสำคัญที่สร้างได้จากระบบ BANNAR ในการวิจัยครั้งนี้จึงได้ทดลองเปรียบเทียบลักษณะสำคัญจากระบบ BANNAR กับระบบ LINUS ซึ่งสามารถสร้างลักษณะสำคัญจากความรู้ภูมิหลังได้ โดยได้เปรียบเทียบประสิทธิภาพของลักษณะสำคัญจากทั้งสองระบบในสองกรณีด้วยกัน คือ เปรียบเทียบโดยใช้นิวรอลเน็ตเวิร์กตามระบบ BANNAR และเปรียบเทียบโดยใช้ระบบ C4.5 มาทำการเรียนรู้ เนื่องจากลักษณะสำคัญจากทั้ง BANNAR และ LINUS จัดอยู่ในรูปค่าคุณลักษณะ จึงสามารถเปรียบเทียบด้วยระบบ C4.5 ได้ ผลการทดลองเป็นดังตารางที่ 4.5

ตารางที่ 4.5 ตารางแสดงเปอร์เซ็นต์ความถูกต้องของค่าความจริงของลักษณะสำคัญที่สร้างได้จากระบบ BANNAR และค่าคุณลักษณะจากระบบ LINUS โดยใช้นิวรอลเน็ตเวิร์กและ C4.5 ในการทดสอบ

ชุดข้อมูล	จำนวนลักษณะสำคัญ		เปอร์เซ็นต์ความถูกต้องของนิวรอลเน็ตเวิร์ก		เปอร์เซ็นต์ความถูกต้องของ C4.5	
	BANNAR	LINUS	BANNAR	LINUS	BANNAR	LINUS
TCR	467.0	36.0	94.40	67.10	89.20	66.54
FEM	115.0	209.0	64.45	57.97	60.46	60.45
MUTA	18.3	16.0	83.58	78.53	81.44	74.41
KRK	39.0	72.0	99.93	99.82	99.88	99.89

หมายเหตุ ระบบทุกระบบทดลองแบบ 3CV จำนวนกฎ จำนวนลักษณะสำคัญ และเปอร์เซ็นต์ความถูกต้องเป็นค่าเฉลี่ยจากทั้งสามกลุ่ม

จากตารางที่ 4.5 พบว่า เปอร์เซ็นต์ความถูกต้องเมื่อใช้ลักษณะสำคัญของระบบ BANNAR และระบบ LINUS เป็นค่าคุณลักษณะเพื่อป้อนให้กับนิวรอลเน็ตเวิร์กและ C4.5 ทำการเรียนรู้ พบว่าค่าความจริงของลักษณะสำคัญของระบบ BANNAR ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าลักษณะสำคัญจากระบบ LINUS ในทุกชุดข้อมูลเมื่อทดสอบด้วยนิวรอลเน็ตเวิร์ก และให้เปอร์เซ็นต์ความถูกต้องสูงกว่าในชุดข้อมูล TCR, FEM และ MUTA ให้เปอร์เซ็นต์ความถูกต้องต่ำกว่าในชุดข้อมูล KRK เมื่อทดสอบด้วยระบบ C4.5

4. ผลการทดลองเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เป็นตัวอย่างที่ใช้เรียนรู้ในระบบ TILDE, 1BC และ C4.5

เพื่อแสดงให้เห็นถึงประสิทธิภาพของลักษณะสำคัญที่ได้จากระบบ BANNAR ในการวิจัยนี้จึงได้ทดลองเปรียบเทียบโดยการใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เป็นตัวอย่างในการเรียนรู้

ของระบบอื่นที่ทำการทดลองเปรียบเทียบในการวิจัยครั้งนี้ โดยเลือกระบบที่สามารถจำแนกตัวอย่างออกเป็นหลายกลุ่มได้ คือ ระบบ TILDE, 1BC และ C4.5 โดยระบบ TILDE และ 1BC สามารถเรียนรู้และทดสอบได้โดยใช้โครงสร้างภายในระบบเอง ส่วนระบบ C4.5 ได้ทดสอบโดยใช้ลักษณะสำคัญซึ่งเป็นค่าคุณลักษณะจากระบบ LINUS แล้วจึงทำการเรียนรู้และทดสอบโดยระบบ C4.5 ผลการทดลองเปรียบเทียบเมื่อใช้ค่าความจริงจากระบบ BANNAR เป็นตัวอย่างที่ใช้เรียนรู้กับใช้ระบบต่างๆ เรียนรู้เพียงอย่างเดียวเป็นดังตารางที่ 4.6

ตารางที่ 4.6 ตารางแสดงเปอร์เซ็นต์ความถูกต้องเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เป็นตัวอย่างที่ใช้เรียนรู้และทดสอบของระบบ TILDE, 1BC และ C4.5

ชุดข้อมูล	TILDE		1BC		C4.5	
	TILDE	FEATURE+TILDE	1BC	FEATURE +1BC	C4.5	FEATURE +C4.5
TCR	88.57	89.13	77.23	91.55	66.54	89.22
FEM	58.02	60.33	46.73	52.36	60.45	60.46
MUTA	68.94	81.44	77.72	83.02	74.41	81.44
KRK	69.80	97.60	87.12	92.00	99.89	99.88

หมายเหตุ เปอร์เซ็นต์ความถูกต้องของระบบ C4.5 ใช้ค่าคุณลักษณะจากระบบ LINUS แล้วจึงใช้ระบบ C4.5 ทำการเรียนรู้และทดสอบ และ FEATURE+X คือ การใช้ลักษณะสำคัญที่ระบบ BANNAR สร้างขึ้นให้ระบบ X เรียนรู้

ผลการทดลองที่ได้แสดงถึงประสิทธิภาพของลักษณะสำคัญที่ระบบ BANNAR สร้างขึ้น เมื่อใช้ค่าความจริงจากลักษณะสำคัญเหล่านี้แทนค่าอินพุตปกติของระบบ TILDE และระบบ 1BC และใช้แทนค่าคุณลักษณะจากระบบ LINUS ในตารางช่อง C4.5 พบว่าค่าความจริงของลักษณะสำคัญสามารถทำให้เปอร์เซ็นต์ความถูกต้องของทุกระบบสูงขึ้นในทุกชุดข้อมูล ยกเว้นในชุดข้อมูล KRK เมื่อทำการทดสอบด้วยระบบ C4.5 พบว่า ค่าคุณลักษณะจากระบบ LINUS ให้เปอร์เซ็นต์ความถูกต้องสูงกว่าค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เล็กน้อย

5. ผลการทดลองเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เพียงอย่างเดียว

นอกจากการนำค่าความจริงของลักษณะสำคัญจากระบบ BANNAR มาทดสอบกับระบบต่างๆ แล้ว เพื่อทดสอบประสิทธิภาพของลักษณะสำคัญที่ได้จากระบบ BANNAR ในการวิจัยนี้ จึงได้ทำการทดสอบโดยใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เพียงอย่างเดียวมาใช้จำแนกตัวอย่างที่ใช้ทดสอบ ซึ่งมีวิธีการทดสอบดังต่อไปนี้

1. หาอัตราส่วนของลักษณะสำคัญที่ตรงกับตัวอย่างต่อจำนวนลักษณะสำคัญทั้งหมดในกฎข้อนั้น ดังสมการ

$$matchRatio(r,e) = \frac{trueFeature(r,e)}{allFeature(r)}$$

เมื่อ *trueFeature(r,e)* คือ จำนวนของลักษณะสำคัญในกฎข้อ r ที่ตรงกับตัวอย่าง e

allFeature(r) คือ จำนวนลักษณะสำคัญทั้งหมดในกฎข้อ r

2. จำแนกตัวอย่างโดยแบ่งออกเป็น 2 กรณี คือ

2.1 ในกรณีปัญหาที่มีลักษณะเป็นหลายกลุ่ม จะจำแนกตัวอย่างโดยเลือกกลุ่มซึ่งตรงกับกฎที่ให้ค่า *matchRatio* สูงที่สุด

2.2 ในกรณีปัญหาที่มีลักษณะเป็นสองกลุ่ม เนื่องจากในปัญหาที่มีลักษณะเป็นสองกลุ่มนี้ไม่มีกฎสำหรับกลุ่มตัวอย่างลบ ดังนั้นจึงต้องใช้วิธีการหาค่า *matchRatio* ที่มากที่สุด (θ_t) ในกลุ่มตัวอย่างลบของกลุ่มตัวอย่างที่ใช้เรียนรู้ หากตัวอย่างที่ใช้ทดสอบให้ค่า *matchRatio* สูงกว่า θ_t จะถูกจำแนกเป็นตัวอย่างบวก และหากน้อยกว่าหรือเท่ากับค่า θ_t จะถูกจำแนกเป็นตัวอย่างลบ ซึ่งสามารถเขียนเป็นเงื่อนไขสำหรับตัวอย่างที่จะถูกจำแนกเป็นกลุ่มตัวอย่างบวกได้ดังนี้

$$\exists r \; matchRatio(r,e) > \theta_t$$

เมื่อ

$$\theta_t = \max_{n \in NE} matchRatio(r,n)$$

และ NE คือ เซตของตัวอย่างลบที่ใช้ทำการเรียนรู้

ผลการทดลองในตารางที่ 4.7 แสดงจำนวนตัวอย่างที่ถูกจำแนกถูกต้องเมื่อทดสอบโดยใช้ค่าความจริงจากลักษณะสำคัญของระบบ BANNAR เพียงอย่างเดียว โดยในกรณีที่ปัญหามีลักษณะหลายกลุ่ม (TCR และ FEM) จะจำแนกโดยเลือกกลุ่มซึ่งตรงกับกฎที่มีค่า *matchRatio* สูงที่สุด ส่วนในกรณีปัญหาที่มีลักษณะเป็นสองกลุ่ม (MUTA และ KRK) ใช้วิธีการจำแนกโดยการหาค่า θ_t สำหรับกฎแต่ละข้อจากกลุ่มตัวอย่างที่ใช้เรียนรู้ก่อน จากนั้นจึงนำค่า θ_t ที่ได้ไปจำแนกตัวอย่างที่ใช้ทดสอบ หากตัวอย่างนั้นเมื่อนำมาเทียบกับลักษณะสำคัญของกฎแต่ละข้อแล้วได้ค่า *matchRatio* ในกฎข้อใดก็ตามสูงกว่าค่า θ_t สำหรับกฎข้อนั้นจะจำแนกตัวอย่างที่ทำการทดสอบนี้เป็นกลุ่มตัวอย่างบวก และหากไม่มีกฎข้อใดเลยที่เมื่อนำตัวอย่างมาเทียบกับลักษณะสำคัญแล้วให้ค่า *matchRatio* สูงกว่า θ_t จะจำแนกตัวอย่างนั้นเป็นตัวอย่างลบ

จากตารางที่ 4.7 แสดงให้เห็นว่าเมื่อใช้ลักษณะสำคัญเพียงอย่างเดียว สามารถจำแนกตัวอย่างได้ถูกต้องมากขึ้นกว่ากฎจากระบบ PROGOL หรือ GOLEM โดยเฉพาะในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม สามารถเพิ่มความถูกต้องในส่วนของตัวอย่างที่ไม่ครอบคลุมด้วยกฎข้อใดข้อหนึ่งภายในเซตของกฎได้ จากชุด

ตารางที่ 4.7 ตารางแสดงจำนวนตัวอย่างที่ถูกจำแนกถูกต้องเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เพียงอย่างเดียว เปรียบเทียบกับวิธีการ PROGOL หรือ GOLEM

ชุดข้อมูล	จำนวนตัวอย่างทดสอบ	BANNAR		ลักษณะสำคัญอย่างเดียว		PROGOL หรือ GOLEM	
		ครอบคลุม	ไม่ครอบคลุม	ครอบคลุม	ไม่ครอบคลุม	ครอบคลุม	ไม่ครอบคลุม
TCR	2,143	1570/1611	453/532	1540/1611	222/532	1540/1611	3/532
FEM	278	145/200	34/78	146/200	24/78	146/200	14/78
MUTA	188	102/109	55/79	100/109	54/79	100/109	54/79
KRK	10,000	3356/3360	6637/6640	3355/3360	6634/6640	3355/3360	6634/6640

หมายเหตุ ตัวเลขในตารางแสดงถึง จำนวนตัวอย่างที่ถูกจำแนกถูกต้อง/จำนวนตัวอย่างทั้งหมดในกลุ่มนั้น

ข้อมูล TCR พบว่า มีตัวอย่างที่ไม่ครอบคลุมด้วยกฎจำนวน 532 ตัวอย่าง เมื่อใช้วิธีการเลือกตามกลุ่มหลักให้ความถูกต้องเพียง 3 ตัวอย่าง ในขณะที่เมื่อใช้ค่าความจริงของลักษณะสำคัญจาก BANNAR เพียงอย่างเดียว สามารถจำแนกได้ถูกต้อง 222 ตัวอย่าง ชุดข้อมูล FEM ซึ่งเป็นชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่มเช่นกัน เมื่อใช้วิธีการจำแนกตามกลุ่มหลัก สามารถจำแนกได้ถูกต้อง 14 ตัวอย่าง จากตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎทั้งหมด 78 ตัวอย่าง และเมื่อใช้ค่าความจริงของลักษณะสำคัญจาก BANNAR เพียงอย่างเดียว สามารถจำแนกได้ถูกต้อง 24 ตัวอย่าง สำหรับในชุดข้อมูลที่มีลักษณะเป็นสองกลุ่ม คือ MUTA และ KRK ลักษณะสำคัญแต่เพียงอย่างเดียว ให้จำนวนตัวอย่างที่ถูกจำแนกได้ถูกต้องเท่ากับระบบ PROGOL หรือ GOLEM ทั้งในกลุ่มตัวอย่างที่ถูกครอบคลุมและไม่ถูกครอบคลุมด้วยกฎ จะเห็นได้ว่า ลักษณะสำคัญแต่เพียงอย่างเดียวสามารถเพิ่มประสิทธิภาพให้กับกฎที่ได้จากระบบ PROGOL หรือ GOLEM ได้ โดยเฉพาะอย่างยิ่งในกลุ่มตัวอย่างที่ไม่ถูกครอบคลุมด้วยกฎในชุดข้อมูลที่มีลักษณะเป็นหลายกลุ่ม

เมื่อเปรียบเทียบความถูกต้องของระบบ BANNAR กับการใช้ลักษณะสำคัญเพียงอย่างเดียว พบว่าความถูกต้องที่เพิ่มขึ้นเมื่อใช้นิวรอลเน็ตเวิร์กทำการเรียนรู้และทดสอบในระบบ BANNAR สามารถทำให้จำนวนตัวอย่างที่ถูกจำแนกถูกต้องในชุดข้อมูล TCR เพิ่มขึ้นจาก 222 ตัวอย่างเมื่อใช้ค่าความจริงของลักษณะสำคัญเพียงอย่างเดียว เป็น 453 ตัวอย่าง และสามารถจำแนกตัวอย่างได้ถูกต้องในชุดข้อมูล FEM เพิ่มขึ้นจาก 24 ตัวอย่างเมื่อใช้ค่าความจริงของลักษณะสำคัญเพียงอย่างเดียว เป็น 34 ตัวอย่างเมื่อใช้นิวรอลเน็ตเวิร์ก จะเห็นว่านิวรอลเน็ตเวิร์กสามารถเพิ่มความถูกต้องให้กับลักษณะสำคัญจากระบบ BANNAR โดยการให้ค่าน้ำหนักต่างๆ กันกับลักษณะสำคัญที่มีความสำคัญไม่เท่ากัน ผลการทดลองในตารางที่ 4.7 แสดงให้เห็นว่าเปอร์เซ็นต์ความถูกต้องที่เพิ่มขึ้นในระบบ BANNAR เป็นผลมาจากประสิทธิภาพของนิวรอลเน็ตเวิร์กส่วนหนึ่ง และอีกส่วนหนึ่งคือประสิทธิภาพของลักษณะสำคัญที่ได้จากระบบ BANNAR