บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของงานวิจัย

การจำแนกข้อมูลเป็นส่วนหนึ่งของงานหลายชนิดในสาขาวิชาปัญญาประดิษฐ์ ซึ่งมีการศึกษากันอย่างกว้างขวางและมีหลากหลายแนวทางและวิธีที่ใช้ในการทำงานเพื่อที่จะจำแนกข้อมูลให้ได้อย่างถูกต้อง โปรแกรมตรรกะเชิงอุปนัย (Inductive Logic Programming) นั้นเป็นเครื่องมือการเรียนรู้หนึ่งซึ่งสามารถนำมาใช้จำแนกข้อมูลได้ โดยลักษณะการเรียนรู้จะแบ่งกลุ่มของตัวอย่างออกเป็นตัวอย่างบวก (Positive example) และตัวอย่างลบ (Negative Example) และใช้ความรู้เบื้องต้นพื้นฐานซึ่งเป็นองค์ประกอบของตัวอย่างทั้งสองประเภท (Background Knowledge) มาเป็นข้อมูลในการเรียนรู้ และสร้างกฎขึ้นเพื่อแยกข้อมูลตัวอย่างออกจากกัน

การเรียนรู้ของโปรแกรมตรรกะเชิงอุปนัยนั้นมีหลักการสำคัญอยู่ที่การค้นหากฎที่ครอบคลุมตัวอย่างต่างๆ โดยการค้นหานั้นจะทำโดยนำความรู้ภูมิหลังมาเรียงต่อกันตามความสัมพันธ์ที่กำหนดไว้ และในกฎนั้นไปทดสอบกับตัวอย่างบวกและตัวอย่างลบว่ากฎที่สร้างขึ้นมาในขณะนั้นครอบคลุมตัวอย่างบวกหรือลบอย่างไร ถ้ากฎที่สร้างขึ้นมายังครอบคลุมตัวอย่างบวกไม่ครบ หรือยังครอบคลุมตัวอย่างลบอยู่ ก็จะปรับเปลี่ยนสมมติฐานโดยการเลือกความรู้ภูมิหลังใหม่ขึ้นมา และสร้างเป็นกฎใหม่และนำไปทดสอบกับตัวอย่าง ทำอย่างนี้ซ้ำไปเรื่อยๆ โดยเป้าหมายของการเรียนรู้โดย คือค้นหากฎที่สามารถจำแนกตัวอย่างบวกและลบออกจากกันได้ กระบวนการค้นหาจะหยุดลงเมื่อสามารถค้นหากฎที่ครอบคลุมตัวอย่างบวกทั้งหมด และไม่ครอบคลุมตัวอย่างลบเลย

แต่อย่างไรก็ดีการค้นหากฎที่ครอบคลุมตัวอย่างตามที่ต้องการนั้นถึงแม้การเรียนรู้ของโปรแกรมตรรกะเชิงอุปนัยจะสามารถทำได้อย่างถูกต้องและมีประสิทธิภาพ แต่ก็ยังมีข้อจำกัดดังนี้คือ

1. กระบวนการเรียนรู้ของโปรแกรมตรรกะเชิงอุปนัย เพื่อสร้างกฎนั้นใช้เวลานาน เนื่องจากจะต้องนำกฎที่สร้างได้มาทดสอบกับตัวอย่างบวก และตัวอย่างลบทั้งหมดเพื่อ

2. ค้นหากฎที่ดีที่สุดที่สามารถอธิบายครอบคลุมตัวอย่างบวกได้ทั้งหมด และไม่สามารถอธิบายครอบคลุมตัวอย่างลบเลย

3. ในกรณีที่ต้องการจำแนกตัวอย่างเป็นหลายกลุ่ม กฎที่ได้จากการเรียนรู้โปรแกรมตรรกะเชิงอุปนัยอาจไม่สามารถจำแนกข้อมูลตัวอย่างที่มีลักษณะต่างจากข้อมูลที่เรียนรู้ได้ เช่น การจำแนกข้อมูลใหม่ๆ หรือจำแนกข้อมูลที่มีสัญญาณรบกวน

มีงานวิจัยหลายงานที่สนใจการเพิ่มประสิทธิภาพให้กับโปรแกรมตรรกะเชิงอุปนัย เพื่อจะลดปัญหาทั้งสองและเพิ่มประสิทธิภาพให้กับการใช้โปรแกรมตรรกะเชิงอุปนัย ในงานวิจัย (บุญเสริม กิจศิริกุล และ สุกรี สินธุภิญโญ,2542, หน้า 162-173) ได้ศึกษาการจำแนกตัวอักษรไทยด้วยกฎจากโปรแกรมตรรกะเชิงอุปนัย โดยมุ่งเน้นที่การแก้ปัญหาที่กฎที่ได้จากการเรียนรู้ของโปรแกรมตรรกะเชิงอุปนัยมีความเฉพาะเจาะจงมากเกินไป ทำให้ไม่สามารถจะจำแนกตัวอักษรที่มีลักษณะไม่ตรงกับกลุ่มอักษรที่ใช้ในการเรียนรู้ หรือตัวอักษรที่มีสัญญาณรบกวนได้ งานวิจัยนี้ได้แก้ไขปัญหาความเฉพาะเจาะจงของกฎโดยอาศัยแนวคิดว่ากฎทั้งกฎอาจจะมีความเฉพาะเจาะจงมากเกินไปทำให้ความสามารถในการครอบคลุมอธิบายตัวอย่างน้อยลง แต่ถ้าใช้เพียงบางส่วนของกฎมาอธิบายตัวอย่างจะทำให้ความเฉพาะเจาะจงลดลง จะทำให้สามารถครอบคลุมตัวอย่างได้มากขึ้น จึงเกิดวิธีการจำแนกข้อมูลด้วยการแบ่งกฎเป็นกฎย่อยขึ้น โดยในงานวิจัยนี้ ได้ใช้วิธีแบ่งกฎออกเป็นกฎย่อยและใช้การถ่วงน้ำหนักกฎด้วยแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก

ในงานวิจัย (สุกรี สินธุภิญโญ และ ชลวิช นัทธี และ Masayuki Numao and Takashi Okada และ บุญเสริม กิจศิริกุล, 2004, หน้า 183-195) เป็นการเพิ่มขีดความสามารถให้กับโปรแกรมตรรกะเชิงอุปนัย โดยสามารถจำแนกข้อมูลแบบหลายกลุ่ม (Multi-class) ได้โดยอาศัยการกฎย่อยรวมกับการถ่วงน้ำหนักโดยให้คะแนนแบบ Winnow

ถึงแม้ว่าในงานวิจัย (บุญเสริม กิจศิริกุล และ สุกรี สินธุภิญโญ, 2547, หน้า 162-173) จะเพิ่มขีดความสามารถให้กับโปรแกรมตรรกะเชิงอุปนัยดังข้างต้นแล้ว แต่เมื่อมาวิเคราะห์ถึงขั้นตอนการทำงานของทั้งสองงาน แล้วพบว่ากฎย่อยซึ่งถือว่าเป็นหัวใจสำคัญนั้นยังคงได้มาจากการเรียนรู้ที่สมบูรณ์ กล่าวคือกฎย่อยนั้นจะได้มาจากการแบ่งกฎที่เรียนรู้อย่างสมบูรณ์แล้วออกเป็นส่วนๆ ดังนั้นจึงจำเป็นจะต้องมีการเรียนรู้ในครั้งแรกก่อน ซึ่งการเรียนรู้ที่สมบูรณ์เพื่อให้ได้กฎนั้นเมื่อข้อมูลมีขนาดใหญ่จะทำให้ต้องเสียเวลามาก ดังนั้นในเมื่อสิ่งที่เราต้องจะต้องใช้ในการทำงานต่อไปกับนิวรอลเน็ตเวิร์กเป็นกฎย่อย เราอาจจะสร้างกฎย่อยขึ้นมาใช้จำแนกข้อมูลโดยไม่จำเป็นต้องมีการเรียนรู้ที่สมบูรณ์ในรอบแรกได้ เพื่อเป็นการลดเวลาในการเรียนรู้แบบโปรแกรมตรรกะเชิงอุปนัยโดยยังคงประสิทธิภาพในการจำแนกตัวอย่างได้เช่นเดิม

1.2 วัตถุประสงค์

เพื่อศึกษาการเรียนรู้ของการโปรแกรมตรรกะเชิงอุปนัยทั้งแบบปกติ และแบบการใช้งานโดยการแบ่งเป็นกฎย่อย

1.2.1 ศึกษาการใช้นิวรอลเน็ตเวิร์กในการใช้จำแนกข้อมูลรวมกับการใช้โปรแกรมตรรกะเชิงอุปนัย

1.2.2 ศึกษาการจำแนกข้อมูลโดยใช้กฎแบบกึ่งเกียจคร้าน ที่มาจากการพัฒนาการเรียนรู้ของโปรแกรมตรรกะเชิงอุปนัย ซึ่งลดขั้นตอนของการเรียนรู้แบบสมบูรณ์ในรอบแรก

1.2.3 เปรียบเทียบผลลัพธ์และเวลาของการจำแนกตัวอย่างด้วยกฎกึ่งเกียจคร้าน กับการเรียนรู้อย่างสมบูรณ์

1.3 ขอบเขตของงานวิจัย

1.3.1 ชุดข้อมูลที่ใช้ในการทดลองประกอบด้วย ชุดข้อมูลตัวอย่างการหาจำนวนองค์ประกอบจำกัดของโครงข่าย (Finite-Element Mesh), กลุ่มของข้อมูลความสามารถก่อกลายพันธุ์ของโมเลกุล (Mutagenecity of molecules), กลุ่มข้อมูลการก่อมะเร็ง (Carcinogenesis)

1.3.2 ระบบปฏิบัติการที่ใช้เป็น Microsoft Windows XP professional SP2

1.3.3 เครื่องมือในการสร้างกฎด้วยวิธีการโปรแกรมตรรกะเชิงอุปนัย Aleph

1.3.4 เครื่องมือในการสร้างนิวรอลเน็ตเวิร์กใช้แมทแล็ป เวอร์ชัน2006a

1.3.5 ใช้วิธีทดสอบความถูกต้องแบบไขว้ข้ามในการประมาณค่าความผิดพลาด (5-fold Cross-Validation)

1.4 ผลที่คาดว่าจะได้รับ

สามารถพัฒนาวิธีการเรียนรู้โดยโปรแกรมตรรกะเชิงอุปนัยได้ โดยลดเวลาการค้นหากฎที่จะมาสร้างตัวจำแนกข้อมูลได้ โดยการจำแนกข้อมูลที่ได้มาจากกฎเหล่านี้จะให้ผลลัพธ์การจำแนกที่มีความถูกต้องแม่นยำเพิ่มขึ้น

1.5 รายละเอียดของวิทยานิพนธ์

ในส่วนของวิทยานิพนธ์แบ่งออกเป็นส่วนต่างๆ ดังนี้

บทที่1 บทนำ ประกอบด้วย ความเป็นมาและความสำคัญของงานวิจัยนี้ วัตถุประสงค์ ขอบเขตของงานวิจัย ผลที่คาดว่าจะได้รับ และรายละเอียดของโครงร่างวิทยานิพนธ์

บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ประกอบด้วย ความรู้พื้นฐานที่เกี่ยวข้องกับวิทยานิพนธ์ฉบับนี้ และงานวิจัยที่เกี่ยวกับการจัดการกฎของโปรแกรมตรรกะเชิงอุปนัย

บทที่ 3 การดำเนินการวิจัย ประกอบไปด้วย โครงสร้างของระบบ วิธีการทดลองเปรียบเทียบที่ใช้ในวิทยานิพนธ์ฉบับนี้ และ ข้อมูลที่ใช้เป็นตัวอย่างในการทดลองต่างๆ

บทที่ 4 ผลการทดลอง ประกอบไปด้วยวิธีการทำการทดลอง ชุดข้อมูลที่ทำการทดลอง และผลที่ได้จาการทดลอง

บทที่ 5 สรุปผลการศึกษาและข้อเสนอแนะ