การประมาณกฎจากวิธีการโปรแกรมตรรกะเชิงอุปนัยด้วยวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก

นายสุกรี สินธุภิญโญ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2544

ISBN 974-03-0472-9

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

I20388317 26 มี.ค. 2546

APPROXIMATE MATCH OF INDUCTIVE LOGIC PROGRAMMING RULES BY BACKPROPAGATION NEURAL NETWORKS

MR. SUKREE SINTHUPINYO

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Computer Engineering

Department of Computer Engineering

Faculty of Engineering

Chulalongkorn University

Academic year 2001

ISBN 974-03-0472-9

หัวข้อวิทยานิพนธ์	การประมาณกฎจากวิธีการโปรแกรมตรรกะเชิงอุปนัยด้วยวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก
โดย	นายสุกรี สินธุภิญโญ
สาขาวิชา	วิศวกรรมคอมพิวเตอร์
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาดุษฎีบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา)

..... อาจารย์ที่ปรึกษา
(ผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล)

..... กรรมการ
(รองศาสตราจารย์ ดร.วันชัย ริ้วไพบูลย์)

..... กรรมการ
(รองศาสตราจารย์ ดร.สมชาย ประสิทธิ์จูตระกูล)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ธนารักษ์ ธีระมั่นคง)

สุกรี สินธุภิญโญ : การประมาณกฎจากวิธีการโปรแกรมตรรกะเชิงอุปนัยด้วยวิธีการแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก. (APPROXIMATE MATCH OF INDUCTIVE LOGIC PROGRAMMING RULES BY BACKPROPAGATION NEURAL NETWORKS)
อ. ที่ปรึกษา : ผศ. ดร.บุญเสริม กิจศิริกุล, 56 หน้า. ISBN 974-03-0472-9.

ปัญหาประการหนึ่งของวิธีการโปรแกรมตรรกะเชิงอุปนัย (Inductive Logic Programming) หรือไอแอลพี คือ กฎที่ได้จากกระบวนการไอแอลพีไม่สามารถจำแนกตัวอย่างใหม่ หรือตัวอย่างที่มีสัญญาณรบกวนได้ถูกต้อง เนื่องจากระบบไอแอลพีจะเลือกกฎที่ตรงกับตัวอย่างพอดีแล้วจำแนกตามกฎนั้น เมื่อไม่มีกฎข้อใดที่ครอบคลุมหรือตรงพอดีกับตัวอย่างนั้น ระบบไอแอลพีจะไม่สามารถจำแนกตัวอย่างได้ เพื่อแก้ปัญหาในกรณีดังกล่าว ผู้วิจัยจึงได้สร้างขั้นตอนวิธีดึงลักษณะสำคัญ (feature extraction algorithm) เพื่อใช้ร่วมกับแบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก (Backpropagation Neural Network) ในการประมาณกฎใกล้เคียงสำหรับตัวอย่าง ในการวิจัยได้ทดลองเปรียบเทียบวิธีการนี้โดยใช้ชุดข้อมูลที่เป็นตรรกะลำดับที่หนึ่งจำนวน 4 ชุดข้อมูล ผลการทดลองแสดงถึงเปอร์เซ็นต์ความถูกต้องที่เพิ่มขึ้นจากการใช้กฎจากระบบไอแอลพีเพียงอย่างเดียวจำแนกตัวอย่าง โดยเฉพาะปัญหาที่มีลักษณะเป็นหลายกลุ่ม (multi-class problem) และเปอร์เซ็นต์ความถูกต้องที่ได้ยังสูงกว่าระบบอื่นที่ได้ทดลองเปรียบเทียบไว้ในงานวิจัยนี้ นอกจากนี้ยังได้ทดลองเพื่อแสดงให้เห็นถึงความทนทานต่อสัญญาณรบกวนในชุดข้อมูลที่มีลักษณะเป็นสองกลุ่ม ผลการทดลองปรากฏว่าเปอร์เซ็นต์ความถูกต้องของวิธีการที่สร้างขึ้นในงานวิจัยนี้ ลดลงช้ากว่าเปอร์เซ็นต์ความถูกต้องของกฎที่ได้จากระบบไอแอลพี

ภาควิชา..วิศวกรรมคอมพิวเตอร์..... ลายมือชื่อนิสิต.....
สาขาวิชา..วิศวกรรมคอมพิวเตอร์..... ลายมือชื่ออาจารย์ที่ปรึกษา.....
ปีการศึกษา....2544..... ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4271827721 : MAJOR COMPUTER ENGINEERING

KEY WORD: INDUCTIVE LOGIC PROGRAMMING / FEATURE EXTRACTION / BACKPROPAGATION NEURAL NETWORK

SUKREE SINTHUPINYO : APPROXIMATE MATCH OF INDUCTIVE LOGIC PROGRAMMING RULES BY BACKPROPAGATION NEURAL NETWORKS. THESIS ADVISOR : ASST. PROF. BOONSERM KIJSIRIKUL Ph.D., 56 pp. ISBN 974-03-0472-9.

One of the disadvantages of Inductive Logic Programming is that the constructed rules may not be able to correctly classify unseen or noisy examples. This is because ILP requires the exact match between the example and the rules. Therefore it could not predict the class of the example that is not covered or does not exactly match with the rules. In order to deal with this kind of problem, we developed a feature extraction algorithm and applied Backpropagation Neural Networks to approximate match between the rules and the example. Our method has been evaluated on four domains of first-order learning problems. The experimental results showed the improvements of our method over the use of ILP's rules, especially in multi-class problems, and the achieved accuracy of our method is higher than other methods tested in the experiments. Furthermore, we have tested the robustness of our method on a noisy two-class domain. The results showed that the accuracy of our method decreased much slower than that of the original rules obtained from ILP.

Department .....COMPUTER ENGINEERING... Student's signature.....

Field of study....COMPUTER ENGINEERING.... Advisor's signature.....

Academic year....2001..... Co-advisor's signature.....

กิตติกรรมประกาศ

การวิจัยและวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้เป็นอย่างดีด้วยความอนุเคราะห์อย่างยิ่งของผู้ช่วยศาสตราจารย์ ดร.บุญเสริม กิจศิริกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ท่านได้ให้ความรู้ คำแนะนำ ข้อคิดเห็น และข้อเสนอแนะต่างๆ ทั้งก่อนวิจัย ขณะทำการวิจัย รวมทั้งหลังทำการวิจัย จนทำให้มีวิทยานิพนธ์ฉบับนี้เกิดขึ้น นอกจากตัววิทยานิพนธ์ฉบับนี้แล้วยังมีผลงานทางวิชาการที่ตีพิมพ์ในที่ต่างๆ ที่อาจารย์เป็นผู้สนับสนุน ให้คำแนะนำ เป็นที่ปรึกษา รวมทั้งข้อคิดที่ควรปรับปรุงในขณะทำการวิจัย ข้าพเจ้าขอขอบคุณผู้สนับสนุนในด้านเงินทุนวิจัย จากทุนพัฒนาอาจารย์ของทบวงมหาวิทยาลัย และทุนผู้ช่วยสอนจากจุฬาลงกรณ์มหาวิทยาลัย ต้นสังกัด ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ที่อนุญาตให้ลาศึกษาต่อได้เต็มเวลา และข้าพเจ้าขอขอบพระคุณ อาจารย์ประจำภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ทุกท่านที่ได้ให้ความเมตตา อบรมสั่งสอน ดูแล ในระหว่างที่ข้าพเจ้าศึกษาอยู่ตลอด 7 ปีการศึกษา ทั้งระดับปริญญาตรี โท และเอก

ท้ายที่สุดข้าพเจ้าขอกราบขอบพระคุณบิดามารดา ครูอาจารย์ทุกท่าน ที่อบรมสั่งสอนทั้งในด้านวิชาการและคุณธรรมเรื่อยมาจนข้าพเจ้าสำเร็จการศึกษา

สารบัญ

หน้า

บทคัดย่อวิทยานิพนธ์ ..... ง
บทคัดย่อวิทยานิพนธ์ (ภาษาอังกฤษ) ..... จ
กิตติกรรมประกาศ ..... ฉ
สารบัญ ..... ช
สารบัญตาราง ..... ฌ
สารบัญภาพ ..... ญ
บทที่ 1 บทนำ 1
ความเป็นมาและความสำคัญของปัญหา 1
วัตถุประสงค์ของงานวิจัย 3
ขอบเขตของงานวิจัย 3
ขั้นตอนและวิธีดำเนินการวิจัย 3
ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย 3
ผลงานที่ตีพิมพ์จากงานวิจัย 4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 5
ทฤษฎีที่เกี่ยวข้อง 5
การเรียนรู้ของเครื่อง 5
การโปรแกรมตรรกะเชิงอุปนัย 7
แบ็กพรอพาเกชันนิวรอลเน็ตเวิร์ก 16
งานวิจัยที่เกี่ยวข้อง 19
บทที่ 3 วิธีดำเนินการวิจัย 23
โครงสร้างของระบบ 23
กระบวนการดึงลักษณะสำคัญ 24
การสร้างนิวรอลเน็ตเวิร์กจากลักษณะสำคัญ 30
การเรียนรู้ของนิวรอลเน็ตเวิร์ก 32
บทที่ 4 ผลการทดลอง 35
เครื่องมือที่ใช้ในการทดลอง 35
ชุดข้อมูลที่ใช้ในการทดลอง 37
ผลการทดลอง 39

สารบัญ (ต่อ)

หน้า

บทที่ 5 สรุปการวิจัยและข้อเสนอแนะ 49
สรุปการวิจัย 49
ข้อเสนอแนะ 51
รายการอ้างอิง 53
ประวัติผู้เขียนวิทยานิพนธ์ 56

สารบัญตาราง

ตาราง หน้า

ตารางที่ 4.1 ตารางแสดงจำนวนลักษณะสำคัญ จำนวนกฎ และเวลาที่ใช้ในการเรียนรู้โดยระบบ BANNAR เทียบกับระบบ PROGOL หรือ GOLEM 40

ตารางที่ 4.2 ตารางแสดงผลการทดลองเปรียบเทียบระหว่าง BANNAR กับระบบอื่น 40

ตารางที่ 4.3 ตารางแสดงจำนวนตัวอย่างที่ถูกจำแนกถูกต้องเปรียบเทียบระหว่าง BANNAR กับ PROGOL หรือ GOLEM แยกกรณีที่ตัวอย่างถูกครอบคลุมด้วยกฎ และไม่ถูกครอบคลุมด้วยกฎ 42

ตารางที่ 4.4 ตารางแสดงเปอร์เซ็นต์ความถูกต้องของชุดข้อมูล KRK เมื่อมีการเพิ่มสัญญาณรบกวนที่ระดับ 5%, 10% และ 15% 43

ตารางที่ 4.5 ตารางแสดงเปอร์เซ็นต์ความถูกต้องของค่าความจริงของลักษณะสำคัญที่สร้างได้จากระบบ BANNAR และค่าคุณลักษณะจากระบบ LINUS โดยใช้นิวรอลเน็ตเวิร์กและ C4.5 ในการทดสอบ 45

ตารางที่ 4.6 ตารางแสดงเปอร์เซ็นต์ความถูกต้องเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เป็นตัวอย่างที่ใช้เรียนรู้และทดสอบของระบบ TILDE, 1BC และ C4.5 46

ตารางที่ 4.7 ตารางแสดงจำนวนตัวอย่างที่ถูกจำแนกถูกต้องเมื่อใช้ค่าความจริงของลักษณะสำคัญจากระบบ BANNAR เพียงอย่างเดียว เปรียบเทียบกับวิธีการ PROGOL หรือ GOLEM 48

สารบัญภาพ

ภาพประกอบ		หน้า
รูปที่ 2.1	ตัวอย่างของการใช้กฎรีโซลูชันในการหาข้อสรุปจากอนุประโยค 2 อนุประโยค	10
รูปที่ 2.2	ตัวอย่างของการใช้ตัวดำเนินการไออาร์ในการสร้างสมมติฐานจากอนุประโยค 2 อนุประโยค	11
รูปที่ 2.3	ตัวอย่างของการใช้วิธีการไออาร์ในการสร้างกฎสำหรับเพรดิเคตเป้าหมาย grandchild(Y,X)	13
รูปที่ 2.4	องค์ประกอบของไบโพลาร์ (รูปบน) และองค์ประกอบซิกมอยด์ (รูปล่าง)	18
รูปที่ 3.1	ขั้นตอนในการเรียนรู้	23
รูปที่ 3.2	ขั้นตอนในการรู้จำ	24
รูปที่ 3.3	ตัวอย่างของกราฟซึ่งใช้แทนกฎลำดับที่หนึ่ง	28
รูปที่ 3.4	ตัวอย่างของลักษณะสำคัญที่อยู่ภายในกฎ	29
รูปที่ 3.5	โครงสร้างของนิวรอลเน็ตเวิร์ก	32
รูปที่ 4.1	น้ำหนักและไบแอสของเส้นเชื่อมของกฎซึ่งได้จากกระบวนการเรียนรู้ของนิวรอลเน็ตเวิร์ก	44