บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การพัฒนาของศาสตร์การเรียนรู้ของเครื่อง (Machine Learning) เพิ่มคุณภาพชีวิตของมนุษย์และลดความเหลื่อมล้ำทางโอกาสในสังคม เทคนิคทางด้านการรู้จำตัวอักษรช่วยแปลงเอกสารจากกระดาษให้อยู่ในรูปที่แบบแฟ้มข้อความที่ค้นหาได้ เทคนิคทางด้านการรู้จำเสียงช่วยบุคคลที่เสียแขนไปสามารถสั่งการด้วยเสียงให้อุปกรณ์ต่างๆทำงาน การพยากรณ์อากาศและพายุที่แม่นยำด้วยระบบการเรียนรู้ของเครื่องที่ช่วยให้ชาวประมงเสี่ยงชีวิตน้อยลง สำหรับในทางธุรกิจศาสตร์สาขาการเรียนรู้ของเครื่อง นอกจากจะช่วยลดความเสี่ยงในทางการดำเนินธุรกิจแล้วยังช่วยให้องค์กรสามารถเรียนรู้พฤติกรรมและความต้องการของลูกค้าได้ดียิ่งขึ้น เช่นการทำเหมืองข้อมูลในระบบการบริการความสัมพันธ์ลูกค้า ประโยชน์เหล่านี้เป็นแรงจูงใจให้เกิดการพัฒนาด้านการเรียนรู้ของเครื่องอย่างกว้างขวาง ด้วยหวังให้เกิดสังคมที่ดีขึ้น

ศาสตร์การเรียนรู้ของเครื่องยกระดับความสามารถของคอมพิวเตอร์ ให้มีความสามารถใกล้เคียงมนุษย์มากขึ้น โดยหลักคือความสามารถในการเรียนรู้และการตัดสินใจ ในอดีตคอมพิวเตอร์ทำงานได้เพียงตามที่มนุษย์โปรแกรมไว้ การเรียนรู้ของเครื่อง เป็นการศึกษาและพัฒนาให้คอมพิวเตอร์สามารถเรียนรู้ ทำงานและตัดสินใจได้เอง โดยลดการควบคุมจากมนุษย์เพื่อให้คอมพิวเตอร์ทำงานได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

ในปัจจุบันศาสตร์สาขานี้มีการพัฒนาและนำไปประยุกต์ใช้อย่างต่อเนื่อง แต่ในทางปฏิบัติความสามารถของคอมพิวเตอร์ยังจำกัดอยู่ สองสิ่งที่ท้าทายนักวิจัยอยู่คือความถูกต้องและความรวดเร็วในการเรียนรู้และการทำงาน โครงข่ายประสาทเทียมเคยมีบทบาทสูงในการวิจัยด้านปัญญาประดิษฐ์ ต่อมาตั้งแต่ช่วงปี 1990 เทคนิคการเรียนรู้แบบซัพพอร์ตเวกเตอร์แมชชีนเริ่มมีบทบาทมากขึ้นจากผลงานของ Vapnik [10] โดยมีนักวิจัยจำนวนหนึ่งได้นำไปทดลองใช้จัดการกับปัญหาการรู้จำแบบและพบว่าให้ผลที่ดีมาก

ซัพพอร์ตเวกเตอร์แมชชีน (เอสวีเอ็ม) เป็นเทคนิคการเรียนรู้ที่มีศักยภาพสูงในการใช้งานจริง Schölkopf [8] กล่าวไว้ว่า ในทางทฤษฎี เอสวีเอ็มมีพื้นฐานมาจากแนวคิดที่เรียบง่ายและช่วยให้เข้าใจว่าการเรียนรู้จากตัวอย่างคืออะไรได้อย่างชัดเจน นอกจากนี้ยังให้ผลที่ดีมากเมื่อนำไปใช้ในทางปฏิบัติ

เอสวีเอ็มใช้คุณสมบัติเชิงเรขาคณิตในการคำนวณหาระนาบหลายมิติ (Hyperplane) ที่ดีที่สุดในการแบ่งแยกข้อมูลสอนออกจากกัน นอกจากนี้ยังสามารถจัดการกับ

กรณีปัญหาที่ไม่สามารถแบ่งแยกแบบเชิงเส้นได้อีกด้วย เช่นกรณีที่มีข้อมูลที่ถูกสัญญาณรบกวนปะปนอยู่ในข้อมูลสอน คุณสมบัติเหล่านี้ทำให้เอสวีเอ็มมีจุดเด่นที่สามารถแก้ปัญหาโอเวอร์ฟิตติง (Overfitting) ที่พบในการเรียนรู้โดยใช้โครงข่ายประสาทเทียมได้ [6] อย่างไรก็ดีเอสวีเอ็มมีข้อจำกัดคือสามารถจำแนกกลุ่มได้เพียง 2 กลุ่มเท่านั้น คือใช่กลุ่มที่สนใจหรือไม่ใช่เท่านั้น ด้วยเหตุนี้การนำเอสวีเอ็มมาประยุกต์ใช้ เช่น การรู้จำตัวอักษร ที่ต้องจำแนกตัวอักษรที่แตกต่างกันเกือบร้อยตัว จำต้องอาศัยกระบวนการแยกประเภทแบบหลายกลุ่มที่มีประสิทธิภาพ

มีผู้เสนอวิธีการแก้ปัญหาแบบหลายกลุ่มนี้หลายแนวทางด้วยกัน วิธีมาตรฐานใช้การเปรียบเทียบกลุ่มแต่ละกลุ่มกับกลุ่มอื่นทั้งหมด (1-v-r) อีกวิธีหนึ่งเสนอโดย Friedman [5] ซึ่งเสนอให้เปรียบเทียบกลุ่มแต่ละกลุ่มกับกลุ่มอื่นทีละกลุ่ม แล้วเลือกกลุ่มที่ได้รับการเลือกมากครั้งที่สุด (1-v-1, Max Wins) วิธีนี้ให้ความถูกต้องใกล้เคียงกับผลของวิธีแรก และแม้จะใช้เวลาสอนสั้นกว่า แต่ก็ใช้เวลาจำแนกนานกว่า Dietterich และ Bakiri [4] เสนอกรอบงานทั่วไปที่มีพื้นฐานจากรหัสแก้ความผิดพลาด (Error Correcting Output Code) ซึ่งตัวอย่างจากกลุ่มแต่ละกลุ่มจะถูกจัดสรรสู่เซตย่อยที่ตรงข้ามกันแบบสุ่ม โดยใช้เทคนิคจากทฤษฎีข่าวสารเพื่อเพิ่มความถูกต้องของการจำแนก ในปีพ.ศ. 2543 Platt และคณะ [7] เสนอวิธีกราฟไม่มีวงแบบมีทิศทางสำหรับการตัดสินใจ (Decision Directed Acyclic Graphs : DDAGs) โดยวิธีนี้สามารถลดเวลาจำแนกกลุ่มลงได้ รวมทั้งยังแสดงการคำนวณหาขอบเขตของความผิดพลาดไว้ด้วย อย่างไรก็ดีงานวิจัยในด้านเอสวีเอ็มแบบหลายกลุ่มยังเปิดกว้างสำหรับการพัฒนาอีกมาก

วิทยานิพนธ์ฉบับนี้เสนอการพัฒนาเอสวีเอ็มให้สามารถจัดการกับปัญหาแบบหลายกลุ่ม โดยการวิเคราะห์ข้อบกพร่องของขั้นตอนวิธีดีดีเอจี (DDAG) จากนั้นจึงพัฒนาขั้นตอนวิธีที่แก้ไขข้อบกพร่องของดีดีเอจี พร้อมเสนอทฤษฎีค่าคาดหวังความถูกต้องของทั้งสองขั้นตอนวิธี พร้อมบทพิสูจน์ทฤษฎี เพื่อเป็นแนวทางในการวิจัยด้านนี้ต่อไป รวมถึงสนับสนุนให้สามารถนำเอสวีเอ็มไปประยุกต์ใช้กับปัญหาจริง

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและเปรียบเทียบขั้นตอนวิธีที่ใช้จำแนกแบบหลายกลุ่มที่มีในปัจจุบัน แล้วพัฒนาขั้นตอนวิธีการจำแนกแบบหลายกลุ่มสำหรับตัวจำแนกแบบเอสวีเอ็ม

1.3 ขอบเขตของการวิจัย

1. งานวิจัยทำในส่วนการจำแนกประเภทแบบหลายกลุ่มสำหรับเอสวีเอ็ม
2. ดำเนินการทดลองภายใต้ระบบปฏิบัติการไมโครซอฟต์วินโดวส์ (Microsoft Windows)
3. ทำการทดสอบกับชุดข้อมูลมาตรฐาน ที่อยู่ในรูปแบบแฟ้มข้อความที่แบ่งเป็นลักษณะสำคัญแล้ว
4. การทดลองเปรียบเทียบจะทำการเปรียบเทียบกับเอสวีเอ็มแบบ 1-v-1 เดิม

1.4 ขั้นตอนการดำเนินการ

1. ศึกษาทฤษฎีพื้นฐานของเอสวีเอ็มรวมทั้งงานวิจัยอื่นที่เกี่ยวข้อง
2. ศึกษาเปรียบเทียบประสิทธิภาพขั้นตอนวิธีที่ใช้ในปัจจุบัน
3. ออกแบบขั้นตอนวิธีที่ใช้ในการจำแนกแบบหลายกลุ่ม
4. ทดสอบขั้นตอนวิธีที่นำเสนอ
5. วิเคราะห์ผลการทดลอง
6. สรุปผลและเรียบเรียงวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้แนวทางในการจำแนกแบบหลายกลุ่มสำหรับเอสวีเอ็ม เพื่อนำไปใช้ในทางปฏิบัติ
2. เป็นข้อมูลส่งเสริมการทำวิจัยด้านเอสวีเอ็มในประเทศไทย
3. เป็นส่วนประกอบช่วยให้เอสวีเอ็มมีประสิทธิภาพสูงขึ้น นำไปใช้ทดแทนวิธีการเรียนรู้แบบเดิมได้
4. ลดความผิดพลาดจากการจำแนกแบบเดิม

1.6 ผลงานที่ตีพิมพ์จากงานวิจัย

วิทยานิพนธ์นี้ได้ตีพิมพ์และนำเสนอในงานประชุมวิชาการนานาชาติดังต่อไปนี้

1. N. Ussivakul and B. Kijsirikul. 2001. Adaptive DAG: Another Approach for Multiclass Classification. The Second International Conference on Intelligent Technologies. pp. 334-339. Thailand.
2. B. Kijsirikul and N. Ussivakul. 2002. Multiclass Support Vector Machines Using Adaptive Directed Acyclic Graph. The IEEE/INNS International Joint Conference on Neural Networks. Hawaii.

1.7 ลำดับขั้นตอนในการเสนอผลวิทยานิพนธ์

เนื้อหาของวิทยานิพนธ์ฉบับนี้ถูกแบ่งออกเป็น 5 บทดังนี้ บทที่ 1 เป็นบทนำ บทถัดไปคือบทที่ 2 จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้อง เช่น เอสวีเอ็ม การจำแนกแบบหลายกลุ่ม และดีเอจีเอสวีเอ็ม เป็นต้น บทที่ 3 กล่าวถึงการวิเคราะห์ขั้นตอนวิธีดีดีเอจีและขั้นตอนวิธีใหม่คือเอดีเอจี พร้อมทั้งเสนอทฤษฎีที่พิสูจน์ความแตกต่างระหว่างขั้นตอนวิธีทั้งสองขั้นตอนวิธี หลังจากนั้นจะกล่าวถึงการทดลองเพื่อเปรียบเทียบขั้นตอนวิธีในบทที่ 4 และบทที่ 5 ซึ่งเป็นบทสุดท้ายจะเป็นบทสรุปผลการวิจัย รวมทั้งข้อเสนอแนะที่เป็นประโยชน์