

บทที่ 5

สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในการทำวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาทำความเข้าใจและเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่ม รวมทั้งเปรียบเทียบการทำนายผลภาวะการติดเชื้อของเด็กและวัยรุ่น ในเขตกรุงเทพมหานคร วิธีการจำแนกกลุ่มที่นำมาเปรียบเทียบประสิทธิภาพและการทำนายผลคือ วิธีความใกล้เคียงกันมากที่สุดโดยใช้อัลกอริทึมชนิด IBk วิธีแผนภาพต้นไม้เพื่อการตัดสินใจโดยใช้อัลกอริทึมชนิด J48 วิธีโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมชนิดเพอร์เซปตรอนแบบหลายชั้น กำหนดค่าอัตราการเรียนรู้เป็น 0.1 ค่าโมเมนตัมเป็น 0.9 จำนวนรอบการสอน 20,000 รอบ และชั้นซ่อน 1 ชั้น วิธีซัพพอร์ตเวกเตอร์แมชชีนโดยใช้อัลกอริทึม SMO ชนิดโพลีโนเมียลเคอร์เนล วิธีฐานกฎโดยใช้อัลกอริทึม Decision Table วิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม และวิธีนาอ็ฟเบย์ ในการเปรียบเทียบประสิทธิภาพของวิธีการจำแนกกลุ่มทั้ง 7 วิธี จะใช้ค่าความถูกต้อง ค่าความระลึกและค่าความถ่วงดุล โดยจะพิจารณาจากค่าเหล่านี้ที่มีค่ามากที่สุด ส่วนค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) จะพิจารณาจากค่า MAE และ MSE ที่มีค่าน้อยที่สุด

ในการเปรียบเทียบประสิทธิภาพในการทำนายผลของวิธีการจำแนกกลุ่ม 7 วิธี ในการพิจารณาภาวะการติดเชื้อของเด็กและวัยรุ่นได้ผลสรุปดังนี้

1. วิธีต้นไม้ตัดสินใจมีค่าความถูกต้อง ค่าความระลึก ค่าความถ่วงดุล และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยที่ดีที่สุดคือ 92.1739%, 0.976, 0.948 และ 0.0638 ตามลำดับ
2. วิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความระลึกและค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่ดีที่สุดคือ 0.976 และ 0.0870 ตามลำดับ
3. วิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม มีค่าความระลึกที่ดีที่สุดคือ 0.976

เนื่องจากวิธีต้นไม้ตัดสินใจมีประสิทธิภาพในการทำนายผลดีที่สุด 4 ใน 5 ค่า ดังนั้นวิธีต้นไม้ตัดสินใจเป็นวิธีที่ดีที่สุด

5.2 อภิปรายผลการวิจัย

ผลการศึกษาระบบเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดเกมของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร โดยพิจารณาจากค่าความถูกต้อง ค่าความระลึก และค่าความถ่วงดุล พบว่าวิธีต้นไม้ตัดสินใจที่ดีที่สุด รองลงมาคือวิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม ซึ่งให้ผลสอดคล้องกับงานวิจัยของณัฐวุฒิ ศิริกุลรุ่งโรจน์ และคณะ (2556) ศึกษาการเปรียบเทียบประสิทธิภาพการจำแนกประเภทโดยใช้วิธีความใกล้เคียงกันมากที่สุด วิธีต้นไม้เพื่อการตัดสินใจ วิธีโครงข่ายประสาทเทียมและวิธีซัพพอร์ตเวกเตอร์แมชชีน ที่พบว่าวิธีต้นไม้ตัดสินใจที่ดีที่สุด และให้ผลใกล้เคียงกับงานวิจัยของวาทีนีย์ น้อยเพียร และคณะ (2553) ได้ทำการศึกษาเพื่อเปรียบเทียบวิธีการจำแนกข้อมูลโดยเลือกใช้อัลกอริทึมโครงข่ายประสาทเทียมแบบมัลติเลเยอร์เปอร์เซ็ปตรอน ซัพพอร์ตเวกเตอร์แมชชีน นาอ์ฟ เบย์ และความใกล้เคียงกันมากที่สุดเพื่อประเมินประสิทธิภาพค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่าความถ่วงดุล ใช้ข้อมูลจาก UCI ประกอบด้วย Ozone Days และ Adult จากผลการวิจัยอัลกอริทึมที่ดีที่สุดของข้อมูล Ozone Days คือวิธีซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบ Rbf มีค่าความถูกต้อง 94.83% ค่าความแม่นยำ 96% ค่าความระลึก 96% และค่าความถ่วงดุล 96% ส่วนข้อมูล Adult คือวิธีซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบโพลิโนเมียล มีค่าความถูกต้อง 79.66% ค่าความแม่นยำ 80% ค่าความระลึก 80% และค่าความถ่วงดุล 80% แต่ให้ผลแตกต่างจากงานวิจัยอื่นๆ อยู่บ้าง เช่น งานวิจัยของ Xhemali, et. al. (2009) ได้ทำวิจัยเรื่องวิธีนาอ์ฟ เบย์ วิธีต้นไม้ตัดสินใจ และวิธีโครงข่ายประสาทเทียมในการจำแนกกลุ่มของการฝึกหัด หน้าเวป เพื่อหาอัตราความสำเร็จของวิธีการเหล่านี้ในหลักสูตรฝึกหัดหน้าเวป ผลการวิจัยนี้พบว่าวิธีนาอ์ฟ เบย์ดีที่สุดสำหรับหลักสูตรฝึกหัดหน้าเวป มีค่าความถ่วงดุลมากกว่า 97% ส่วนงานวิจัยของ Tuisima, et. al. (2012) ศึกษาเรื่องการจัดจำแนกกลุ่มของระดับการติดเกมคอมพิวเตอร์ในนักศึกษาชั้นมัธยมศึกษาปีที่ 1-3 โดยใช้วิธีโครงข่ายประสาทเทียม โดยจำแนกระดับการติดเกมโดยใช้วิธีโครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับและเปรียบเทียบผลกับวิธีต้นไม้ตัดสินใจ ทำการวัดความถูกต้องของตัวแบบโดยใช้วิธี 10-fold Cross Validation ผลการทดลองพบว่าความถูกต้องของวิธีโครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับสำหรับเกมระยะยาว เกมที่ไม่มีกฎเกณฑ์แน่นอน เกมระยะเวลาจริง และเกมพื้นฐานคือ 97.75, 91.35, 90.00 และ 97.73 ตามลำดับ ส่วนความถูกต้องของอัลกอริทึมต้นไม้ตัดสินใจคือ 88.76, 92.63, 87.50 และ 90.91 ตามลำดับ และงานวิจัยของ Chou, C. H. (2013) ที่ศึกษาการใช้ Tic-Tac-Toe สำหรับเรียนรู้การจำแนกกลุ่มการทำเหมืองข้อมูลและการประเมินผล ผลการทดลองพบว่าวิธี 3-ความใกล้เคียงกันมากที่สุดมีอัตราความถูกต้องในการจำแนกกลุ่มของเกม Tic-Tac-Toe มากที่สุดคือ 99% วิธีการจำแนกกลุ่มทั้ง 4 วิธี คือ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการถดถอยโลจิสติก และวิธี 3-ความใกล้เคียงกันมากที่สุด มีอัตราความถูกต้องมากกว่า 98% ซึ่งวิธีการทั้ง 4 วิธี ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ

ส่วนการเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดเกมของเด็กและวัยรุ่น โดยพิจารณาจากค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยและความคลาดเคลื่อนกำลังสองเฉลี่ย พบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดคือ 0.0870 ส่วนวิธีต้นไม้ตัดสินใจมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดคือ 0.0638 ซึ่งให้ผลตรงกับงานวิจัยในการเปรียบเทียบประสิทธิภาพการจำแนกประเภทโดยใช้วิธีความใกล้เคียงกันมากที่สุด วิธีต้นไม้เพื่อการตัดสินใจ วิธีโครงข่ายประสาทเทียมและวิธีซัพพอร์ตเวกเตอร์แมชชีน (ณัฐวุฒิ ศิริกุลรุ่งโรจน์ และคณะ, 2556) และการเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศอินเดีย (สุรวุฒร์ ศรีเปารยะ, 2559) ที่พบว่าวิธีการจำแนกกลุ่มที่มีประสิทธิภาพในการทำนายผลดีที่สุดโดยเปรียบเทียบจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือวิธีต้นไม้ตัดสินใจ

แต่แตกต่างจากงานวิจัยในการเปรียบเทียบประสิทธิภาพในการทำนายผลการเป็นโรคเบาหวาน ซึ่งพบว่าวิธีโครงข่ายประสาทเทียมมีค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยและความคลาดเคลื่อนกำลังสองเฉลี่ย น้อยที่สุด (สายชล สีนสมบูรณ์ทอง, 2560) และงานวิจัยข้อมูลโรคไตเรื้อรังโดยใช้วิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีโครงข่ายประสาทเทียม จากการศึกษาพบว่าวิธีโครงข่ายประสาทเทียมมีประสิทธิภาพในการทำนายผลได้ดีกว่าวิธีซัพพอร์ตเวกเตอร์แมชชีน (Vijayarani, S. and Dhayanand, S., 2015) สาเหตุที่การเปรียบเทียบประสิทธิภาพให้ผลแตกต่างกันอาจเนื่องมาจากเรื่องที่ทำการศึกษาอยู่ในสาขาที่แตกต่างกัน จึงให้ประสิทธิภาพในการทำนายผลที่แตกต่างกัน

5.3 ข้อเสนอแนะ

1. ตัวแปรที่นำมาใช้ในการงานวิจัยนี้เป็นเพียงส่วนหนึ่งของการพิจารณาภาวะการติดเกมของเด็กและวัยรุ่นเท่านั้น เพื่อให้การทำนายมีประสิทธิภาพมากขึ้น ควรเพิ่มตัวแปรที่เกี่ยวข้องอื่นๆ อีก
2. เพื่อให้ผลสรุปครอบคลุมกว้างขวางเพิ่มขึ้น ควรจะทำการศึกษาวีธีอื่น ๆ ที่เป็นเทคนิคการทำเหมืองข้อมูลในการจำแนกกลุ่มเหมือนกัน เช่น วิธีโครงข่ายความเชื่อของเบย์เซียน (Bayesian belief network method)
3. เพื่อให้ได้ข้อสรุปของผลการวิเคราะห์ข้อมูลที่มีความสมบูรณ์มากขึ้น เราอาจจะใช้อัลกอริทึมประเภทอื่น ๆ โดยวิธีความใกล้เคียงกันมากที่สุดยังมีอัลกอริทึม KStar และ LWL วิธีแผนภาพต้นไม้เพื่อการตัดสินใจมีอัลกอริทึม Decision Stump, LMT, Random Forest, Random Tree, REP Tree และ ID3 วิธีโครงข่ายประสาทเทียมสามารถกำหนดค่าอัตราการเรียนรู้และค่าโมเมนตัมที่ละเอียดมากขึ้นกว่าเดิม อาจกำหนดจำนวนชั้นซ่อนมากกว่า 1 ชั้นได้ และอาจเพิ่มจำนวนรอบการสอนให้มากขึ้น เช่น 100,000 รอบ วิธีซัพพอร์ตเวกเตอร์แมชชีนมีอัลกอริทึม SMO และใช้ฟังก์ชันเคอร์เนลแบบ Normalized Poly Kernel, Radial Basis Function Kernel, Sigmoid Kernel และ Puk และวิธีฐานกฎยังมีอัลกอริทึม JRip, OneR, PART และ ZeroR