การเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร

An Efficiency Comparison in Prediction of Child and Adolescence Game Addition in Bangkok

สายชล สินสมบูรณ์ทอง*

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ถนนฉลองกรุง เขตลาดกระบัง กรุงเทพมหานคร 10520

Saichon Sinsomboonthong*

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang,
Chalongkrung Road, Ladkrabang, Bangkok 10520

บทคัดย่อ

งานวิจัยนี้ได้ศึกษาและเปรียบเทียบประสิทธิภาพในการทำนาย การติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี ได้แก่ วิธีเพื่อนบ้านใกล้สุด k ตัว โดยใช้อัลกอริทึมชนิด IBk วิธีต้นไม้การตัดสินใจโดยใช้อัลกอริทึมชนิด J48 วิธีโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมชนิดเพอร์เซปตรอนแบบหลายชั้น วิธีซัพพอร์ตเวกเตอร์แมชชีนโดยใช้อัลกอริทึม SMO ชนิดโพลิโนเมียลเคอร์เนล วิธีฐานกฎโดยใช้อัลกอริทึมชนิดตารางการตัดสินใจ วิธีการถดถอยลอจิสติกทวิภาค และวิธีนาอีฟเบส์ โดยข้อมูลที่นำมาศึกษาได้จากการเก็บรวบรวมด้วยแบบสอบถาม เกณฑ์ในการเปรียบเทียบพิจารณาจากค่าความแม่น ความเที่ยง ความระลึก ค่าวัดเอฟ ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าคลาดเคลื่อนกำลังสองเฉลี่ย โดยวิธีการจำแนกกลุ่มที่มีประสิทธิภาพมากที่สุดจะให้ค่าความแม่น ความเที่ยง ความระลึก และค่าวัดเอฟที่มีค่ามากที่สุด แต่จะให้ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด ผลการศึกษาพบว่าวิธีต้นไม้การตัดสินใจมีค่าความแม่น ความระลึกของชั้นไม่ได้ติดเกม ค่าวัดเอฟของชั้นไม่ได้ติดเกม ความเที่ยงของชั้นติดเกมและค่าวัดเอฟของชั้นติดเกมมากที่สุดเท่ากับ 92.17 %, 0.976, 0.948, 0.923 และ 0.842 ตามลำดับ วิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความระลึกของชั้นไม่ได้ติดเกมมากที่สุดเท่ากับ 0.976 และค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดเท่ากับ 0.0870 วิธีฐานกฎมีค่าความเที่ยงของชั้นไม่ได้ติดเกมและความระลึกของชั้นติดเกมมากที่สุดเท่ากับ 0.930 และ 0.806 ตามลำดับ ส่วนวิธีการถดถอยลอจิสติกทวิภาคมีค่าความระลึกของชั้นไม่ได้ติดเกมมากที่สุดเท่ากับ 0.976 และค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดเท่ากับ 0.0612 แต่เนื่องจากวิธีต้นไม้การตัดสินใจมีประสิทธิภาพในการทำนายดีที่สุด 5 ใน 9 ค่า ดังนั้นวิธีต้นไม้การตัดสินใจเป็นวิธีที่สามารถทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานครที่มีประสิทธิภาพมากที่สุด

คำสำคัญ : วิธีเพื่อนบ้านใกล้สุด k ตัว; วิธีต้นไม้การตัดสินใจ; วิธีโครงข่ายประสาทเทียม; วิธีซัพพอร์ตเวกเตอร์แมชชีน; วิธีฐานกฎ; วิธีการถดถอยลอจิสติกทวิภาค; วิธีนาอีฟเบส์

***ผู้รับผิดชอบบทความ :** saichon.ss49@gmail.com

doi: 10.14456/tstj.2018.36

Abstract

In this research to study and compare efficiency in prediction of child and adolescence game addition in Bangkok. The seven classification methods to efficiency comparison in prediction are k-nearest neighbor method by IBk algorithm, decision tree method by J48 algorithm, artificial neural network method by multilayer perceptron algorithm, support vector machine method by polynomial kernel and SMO algorithm, rule based method by decision table algorithm, binary logistic regression method and Naïve Bayes method. The data are collected by questionnaire. An efficiency comparison in prediction of seven classification methods by using the maximum of accuracy, precision, recall, F-Measure and the minimum of mean absolute error and mean square error. The result of the study efficiency comparison the classification method demonstrated that decision tree method has the maximum accuracy, recall (class = no), F-Measure (class = no), precision (class = yes) and F-Measure (class = yes), 92.17 %, 0.976, 0.948, 0923 and 0.842 respectively. In addition, support vector machine method has the maximum recall (class = no), 0.976 and the minimum MAE, 0.0870, rule based method has the maximum precision (class = no) and recall (class = yes), 0.930 and 0.806 respectively. Also binary logistic regression method has the maximum recall (class = no), 0.976 and the minimum MSE, 0.0612 but decision tree method has the best an efficiency five in nine values. Then, decision tree method has the maximum efficiency in prediction of child and adolescence game addition in Bangkok.

Keywords: k-nearest neighbor method; decision tree method; artificial neural network method; support vector machine method; rule based method; binary logistic regression method; Naïve Bayes method

1. บทนำ

ปัจจุบันเทคโนโลยีได้เข้ามามีบทบาทสำคัญกับคนทุกเพศทุกวัย การดำเนินชีวิตส่วนใหญ่มักจะเกี่ยวข้องกับเทคโนโลยีอย่างหลีกเลี่ยงไม่ได้ ตั้งแต่กิจวัตรประจำวัน การทำงาน หรือแม้กระทั่งการพักผ่อน จากสภาพสังคมในปัจจุบัน เทคโนโลยีถูกนำมาใช้เพื่อความบันเทิงมากกว่าวัตถุประสงค์อื่น ๆ โดยเฉพาะอย่างยิ่งเกมคอมพิวเตอร์ประเภทออนไลน์และออฟไลน์ที่กำลังแพร่กระจายอยู่ในทุกชุมชนทุกหนทุกแห่ง เกมคอมพิวเตอร์ได้รับความสนใจจากคนทุกเพศทุกวัยตั้งแต่นักเรียน นักศึกษา และวัยทำงาน เป็นต้น เนื่องจากเกมคอมพิวเตอร์ที่มีภาพเหมือนจริงสามารถดึงดูดความสนใจได้มากกว่าเกมคอมพิวเตอร์รุ่นก่อน ๆ ซึ่งสามารถเล่นได้เป็นกลุ่มครั้งละหลายคน

เกมคอมพิวเตอร์ประเภทออนไลน์เป็นของเล่นทางเทคโนโลยีชนิดหนึ่ง ซึ่งถูกดัดแปลงและพัฒนาออกมาในรูปแบบต่าง ๆ ตามยุคสมัยและความก้าวหน้าทางเทคโนโลยี โดยเฉพาะเด็ก ๆ ซึ่งเป็นวัยอยากรู้อยากลอง เมื่อเกมคอมพิวเตอร์เหล่านี้ คือ ของเล่นแปลกใหม่ที่เข้ามาเป็นผู้สร้างจินตนาการและตอบสนองความต้องการ ทำให้เป็นตัวการ์ตูนที่ชื่นชอบ ผู้ชนะจะได้รับการยอมรับและได้ระบายอารมณ์ลงไป

เล่นเกม แต่เด็ก ๆ ที่เล่นเกมคอมพิวเตอร์โดยไม่รู้จักควบคุมตัวเอง ประกอบกับขาดการเอาใจใส่ที่เหมาะสมจากครอบครัวและสังคมที่อาศัยอยู่ ไม่ตระหนักถึงความสำคัญของปัญหาเหล่านี้ก็มีโอกาสที่จะกลายเป็นเด็กติดเกมได้เช่นกัน แม้จะเป็นเพียงบทบาทที่สมมุติขึ้นก็ตาม จึงไม่ใช่เรื่องแปลกถ้าเด็กจะเพลิดเพลินอยู่กับการเล่นเกมคอมพิวเตอร์ จากปัญหาการติดเกมของเด็กและวัยรุ่นที่มีแนวโน้มทวีความรุนแรงขึ้นจนเป็นปัญหาระดับประเทศ การศึกษาปัจจัยที่ส่งผลต่อการติดเกมอาจจะมีประโยชน์ในการคัดกรองและป้องกันก่อนที่เด็กจะติดเกมได้

เนื่องจากมีข้อมูลเป็นจำนวนมากที่ได้จากแบบสอบถาม การทำเหมืองข้อมูล (data mining) เป็นกระบวนการหนึ่งที่กระทำกับข้อมูลจำนวนมาก เพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์ การแพทย์ เศรษฐกิจและสังคม รวมทั้งในด้านการศึกษา

Xhemali และคณะ [1] ได้ศึกษาการจำแนกกลุ่มของการฝึกหัดเว็บเพจ (web page) ด้วยวิธีการ 3 วิธี คือ วิธีนาอีฟเบส์ วิธีต้นไม้การตัดสินใจ และวิธีโครงข่ายประสาทเทียม โดยการจำแนกกลุ่มเว็บใช้เทคโนโลยีที่แตกต่างกันจำนวนมาก ในการศึกษานี้เน้นการเปรียบเทียบวิธีนาอีฟเบส์ วิธีต้นไม้การตัดสินใจและวิธีโครงข่ายประสาทเทียม สำหรับการวิเคราะห์อัตโนมัติและการจำแนกกลุ่มของข้อมูลคุณลักษณะจากหลักสูตรฝึกหัดเว็บเพจ เพื่อหาอัตราความสำเร็จของวิธีการเหล่านี้ในหลักสูตรฝึกหัดเว็บเพจ ผลการวิจัยนี้พบว่าวิธีนาอีฟเบส์ดีที่สุดสำหรับหลักสูตรฝึกหัดเว็บเพจ โดยมีค่าวัดเอฟ (F-measure) มากกว่า 97 % ต่อมา Tuisima และคณะ [2] ศึกษาเรื่องการจัดจำแนกกลุ่มของระดับการติดเกมคอมพิวเตอร์ในนักศึกษาชั้นมัธยมศึกษาปีที่ 1-3 ด้วยโครงข่ายประสาทเทียม จำแนกระดับการติดเกมโดยใช้วิธีโครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับและเปรียบเทียบผลกับวิธีต้นไม้การตัดสินใจ วัดความแม่น (accuracy) ของตัวแบบโดยใช้วิธี 10-fold cross validation งานวิจัยนี้จำแนกเกมคอมพิวเตอร์โดยพิจารณาจากคุณลักษณะ 4 ประเภท คือ เกมระยะยาว เกมที่ไม่มีกฎเกณฑ์แน่นอน เกมระยะเวลาจริง และเกมพื้นฐาน ผลการทดลองพบว่าความแม่นของวิธีโครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับสำหรับเกมระยะยาว เกมที่ไม่มีกฎเกณฑ์แน่นอน เกมระยะเวลาจริง และเกมพื้นฐานเท่ากับ 97.75, 91.35, 90.00 และ 97.73 ตามลำดับ ส่วนความแม่นของวิธีต้นไม้การตัดสินใจเท่ากับ 88.76, 92.63, 87.50 และ 90.91 ตามลำดับ นอกจากนี้ Chou [3] ได้ศึกษาการใช้ Tic-Tac-Toe สำหรับเรียนรู้การจำแนกกลุ่มการทำเหมืองข้อมูลและการประเมินผล วัตถุประสงค์ของการศึกษานี้เพื่อตรวจสอบว่าการเรียนรู้กลไกลวิธีไหนที่ประสบความสำเร็จในการจำแนกกลุ่มของเกมนี้ ผลการทดลองพบว่าวิธีเพื่อนบ้านใกล้สุด 3 ตัว (k = 3) มีอัตราความแม่นในการจำแนกกลุ่มของเกม Tic-Tac-Toe มากที่สุดเท่ากับ 99 % วิธีการจำแนกกลุ่มทั้ง 4 วิธี คือ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีการถดถอยลอจิสติก และวิธีเพื่อนบ้านใกล้สุด 3 ตัว มีอัตราความแม่นมากกว่า 98 % ซึ่งวิธีการทั้ง 4 วิธี มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

งานวิจัยนี้ผู้วิจัยสนใจเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร โดยใช้ปัจจัยต่าง ๆ ได้แก่ เพศ อายุ ระดับการศึกษา รายได้ต่อเดือน การใช้สื่อสังคมออนไลน์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันจันทร์-ศุกร์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวัน

เสาร์-อาทิตย์/วันหยุด ความคิดเห็นของผู้ปกครองเกี่ยวกับการติดเกมของลูก การหมกมุ่นในเกม ความสูญเสียหน้าที่รับผิดชอบ และการสูญเสียในการควบคุมตัวเอง อารมณ์ พฤติกรรมและเวลา เป็นต้น โดยใช้วิธีการทำนาย 7 วิธี คือ วิธีเพื่อนบ้านใกล้สุด k ตัว (K-nearest neighbor) วิธีต้นไม้การตัดสินใจ (decision tree) วิธีโครงข่ายประสาทเทียม (artificial neural network) วิธีซัพพอร์ตเวกเตอร์แมชชีน (support vector machine) วิธีฐานกฎ (rule based) วิธีการถดถอยลอจิสติกทวิภาค (binary logistic regression) และวิธีนาอีฟเบส์ (Naïve Bayes) เพื่อหาวิธีที่มีประสิทธิภาพทำนายการติดเกมที่เหมาะสมต่อไป

2. วิธีการวิจัย

2.1 เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในการวิจัยครั้งนี้ คือ แบบสอบถามการติดเกมของเด็กและวัยรุ่น ซึ่งประกอบด้วยคำถามเกี่ยวกับเพศ อายุ ระดับการศึกษา รายได้ต่อเดือน การใช้สื่อสังคมออนไลน์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันจันทร์-ศุกร์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันเสาร์-อาทิตย์/วันหยุด ความคิดเห็นของผู้ปกครองเกี่ยวกับการติดเกมของลูก การหมกมุ่นในเกม ความสูญเสียหน้าที่รับผิดชอบ การสูญเสียในการควบคุมตัวเอง อารมณ์ พฤติกรรมและเวลา และการติดเกม และในการวิเคราะห์ข้อมูลใช้โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) เวอร์ชั่น 3.9 โปรแกรม WEKA เป็นโปรแกรมที่สามารถดาวน์โหลดได้จากเว็บไซต์ โดยอยู่ภายใต้การควบคุมของ GPL License ซึ่งเป็นโปรแกรมที่พัฒนามาจากภาษาจาวา (Java) ทั้งหมดและเป็นโปรแกรมที่นิยมในการใช้งานด้านการทำเหมืองข้อมูล

2.2 การเก็บรวบรวมข้อมูลและการแบ่งข้อมูล

2.2.1 การเก็บรวบรวมข้อมูล โดยเก็บรวบรวมข้อมูลจากแบบสอบถามการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ตั้งแต่เดือนมกราคมถึงเดือนเมษายน พ.ศ. 2560 ในเขตลาดกระบัง มีนบุรี และบางเขน เขตละ 500 ชุด (คน) รวม 1,500 ชุด เก็บสำรองเพิ่มเติมอีก 20 % คิดเป็น 300 ชุด รวมเป็นทั้งหมด 1,800 ชุด โดยข้อมูลประกอบด้วยคุณลักษณะ (attribute) ต่าง ๆ ดังนี้ เพศ อายุ ระดับการศึกษา รายได้ต่อเดือน การใช้สื่อสังคมออนไลน์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันจันทร์-ศุกร์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันเสาร์-อาทิตย์/วันหยุด ความคิดเห็นของผู้ปกครองเกี่ยวกับการติดเกมของลูก การหมกมุ่นในเกม ความสูญเสียหน้าที่รับผิดชอบ การสูญเสียในการควบคุมตัวเอง อารมณ์ พฤติกรรม และเวลา และการติดเกม

2.2.2 การแบ่งข้อมูล นำข้อมูลทั้งหมดมาแบ่งออกเป็น 3 ส่วน จากแบบสอบถามการติดเกมของเด็กและวัยรุ่น จำนวน 1,800 ชุด แต่จากการตรวจสอบแบบสอบถามเบื้องต้นพบว่าแบบสอบถามไม่สมบูรณ์ จำนวน 42 ชุด ซึ่งได้ตัดทิ้ง จึงเหลือแบบสอบถาม จำนวน 1,758 ชุด สำหรับการวิเคราะห์ข้อมูล โดยแบ่งข้อมูลออกเป็นสัดส่วนดังนี้ [4]

ส่วนที่ 1 : ชุดข้อมูลฝึกหัด (training data set) เพื่อนำไปสร้างตัวแบบ มีข้อมูล 80 % ของข้อมูลทั้งหมด คิดเป็น 1,406 ชุด แต่ยังตรวจพบแบบสอบถามที่ไม่สมบูรณ์จำนวน 13 ชุด ซึ่งได้ตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ในส่วนที่ 1 จำนวน 1,393 ชุด

ส่วนที่ 2 : ชุดข้อมูลประเมินค่า (evaluation data set) เพื่อนำไปทดสอบความถูกต้องของตัวแบบ มีข้อมูล 15 % ของข้อมูลทั้งหมด คิดเป็น 264 ชุด แต่ยังตรวจพบแบบสอบถามที่ไม่สมบูรณ์ จำนวน 14 ชุด ซึ่งได้ตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ในส่วนที่ 2 จำนวน 250 ชุด

ส่วนที่ 3 : ชุดข้อมูลทดสอบ (testing data set) เพื่อนำไปทำนายตัวแบบ มีข้อมูล 5 % ของข้อมูลทั้งหมด คิดเป็น 88 ชุด แต่ยังตรวจพบแบบสอบถามที่ไม่สมบูรณ์จำนวน 5 ชุด ซึ่งได้ตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ 83 ชุด เก็บข้อมูลเพิ่มเติมอีก 32 ชุด เพื่อให้ครบ 1,758 ชุด ซึ่งจะได้ข้อมูลที่สมบูรณ์ในส่วนที่ 3 จำนวน 115 ชุด

2.3 การวิเคราะห์ข้อมูล

การทำวิจัยครั้งนี้ ผู้วิจัยได้จัดข้อมูลแต่ละชุดออกเป็น 3 ส่วน โดยส่วนที่ 1 ใช้ข้อมูล 80 เปอร์เซ็นต์ ของข้อมูลทั้งหมดในการสร้างตัวแบบ ส่วนที่ 2 ใช้ข้อมูล 15 เปอร์เซ็นต์ ของข้อมูลทั้งหมดในการทดสอบความถูกต้องของตัวแบบ ส่วนที่ 3 ใช้ข้อมูล 5 เปอร์เซ็นต์ ของข้อมูลทั้งหมดในการทำนายตัวแบบ แปลงไฟล์ข้อมูลให้เป็นนามสกุล *.csv เพื่อใช้วิเคราะห์ประสิทธิภาพการจำแนกกลุ่มข้อมูลในโปรแกรม WEKA ซึ่งเป็นโปรแกรมที่สามารถนำมาทดสอบอัลกอริทึมของวิธีการจำแนกกลุ่มได้ เนื่องจากมีอัลกอริทึมที่ได้ระบุไว้ให้เลือกใช้ในโปรแกรมครบตามที่กำหนด ผู้วิจัยได้กำหนดวิธีการจำแนกกลุ่มเพื่อนำมาทดสอบดังนี้

2.3.1 วิธีเพื่อนบ้านใกล้สุด k ตัว (K-nearest neighbor) ใช้อัลกอริทึมชนิด IBk เนื่องจากเป็นฟังก์ชันหลักที่สนใจ ซึ่งเป็นพื้นฐานของอัลกอริทึม 8.1 อัลกอริทึม IBk ยังสามารถกำหนดน้ำหนัก ระยะห่าง และทางเลือก (option) เพื่อกำหนดค่า k โดยใช้ cross-validation [5]

2.3.2 วิธีต้นไม้การตัดสินใจ (decision tree) ใช้อัลกอริทึมชนิด J48 ซึ่งพัฒนามาจาก ID3 สามารถใช้ได้กับข้อมูลแบบไม่ต่อเนื่องและแบบต่อเนื่อง ต่างจาก ID3 ที่ใช้ได้เพียงข้อมูลแบบไม่ต่อเนื่องเท่านั้น [6]

2.3.3 วิธีโครงข่ายประสาทเทียม (artificial neural network) ใช้อัลกอริทึมชนิดเพอร์เซปตรอนหลายชั้น (multilayer perceptron) โดยกำหนดค่าอัตราการเรียนรู้ (learning rate) เป็น 0.1 ค่าโมเมนตัม (momentum) เป็น 0.9 จำนวนรอบการสอน (training time) 20,000 รอบ การวิจัยครั้งนี้ใช้อัลกอริทึมของวิธีโครงข่ายประสาทเทียมชนิดเพอร์เซปตรอนหลายชั้นที่มีชั้นซ่อน (hidden layer) 1 ชั้น แม้ว่าโครงสร้างโครงข่ายประสาทที่ซับซ้อนสามารถมีชั้นซ่อนมากกว่า 1 ชั้น แต่ในทางปฏิบัติการกำหนดชั้นซ่อน 1 ชั้น ก็เพียงพอต่อการวิเคราะห์ข้อมูล [7]

2.3.4 วิธีซัพพอร์ตเวกเตอร์แมชชีน (support vector machine) ใช้อัลกอริทึม SMO ชนิดโพลิโนเมียลเคอร์เนล (polynomial Kernel) เนื่องจากงานวิจัยที่อ้างอิงจาก วาทินี และคณะ [8] ได้ผลว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนที่ใช้อัลกอริทึมชนิดโพลิโน-เมียลเคอร์เนลดีที่สุด

2.3.5 วิธีฐานกฎ (rule based) ใช้ชุดลำดับของกฎมาสร้างรูปแบบการแยกประเภทข้อมูล โดยส่วนใหญ่แล้วจะใช้กฎที่เป็น If ... then ซึ่งเป็นกฎอย่างง่าย [9] ใช้อัลกอริทึมตารางการตัดสินใจเป็นเครื่องมือที่ใช้แสดงเงื่อนไขการตัดสินใจ และเลือกการทำงานหรือกระทำกิจกรรมภายใต้เหตุการณ์ของเงื่อนไขที่ระบุ วิธีการตัดสินใจแบบตารางการตัดสินใจจะเป็นตาราง 2 มิติ

2.3.6 วิธีการถดถอยลอจิสติกทวิภาค (binary logistic regression) เป็นการวิเคราะห์การถดถอยแบบหนึ่งโดยที่ตัวแปรตามเป็นตัวแปรเชิงคุณภาพ ส่วนตัวแปรอิสระอาจจะเป็นตัวแปรเชิงปริมาณหรือเชิงคุณภาพ หรืออาจมีทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพก็ได้ [10]

2.3.7 วิธีนาอีฟเบส์ (Naïve Bayes) คืออัลกอริทึมที่ใช้หลักการของความน่าจะเป็นในการคัดกรองแต่ละชั้น (class) โดยมี 2 ชั้น [11]

การนำผลการวิเคราะห์ข้อมูลมาประเมินผลเพื่อเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่น ใช้วิธีการจำแนกกลุ่ม 7 วิธี โดยนำข้อมูลที่ได้จากการวิเคราะห์มาเปรียบเทียบว่าวิธีในการจำแนกกลุ่มข้อมูล วิธีใดที่มีความแม่น (accuracy) ความเที่ยง (precision) ความระลึก (recall) และค่าวัดเอฟ (F-measure) มากที่สุด และมีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (mean absolute error, MSE) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error, MSE) น้อยที่สุด จะเป็นวิธีที่มีประสิทธิภาพในการทำนายดีที่สุด โดยที่ค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็นมาตรวัดการประเมินค่าได้ดี เนื่องจากค่าคลาดเคลื่อนกำลังสองเฉลี่ยประกอบด้วยทั้งความแปรปรวนและความเอนเอียง [12] ส่วนค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยมีบางงานวิจัยนำมาเป็นเกณฑ์ในการเปรียบเทียบด้วย ดังนั้นในงานวิจัยนี้จึงใช้ทั้งค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าคลาดเคลื่อนกำลังสองเฉลี่ยเป็นเกณฑ์ในการเปรียบเทียบ

ชั้นของค่าจริง	ชั้นของค่าทำนาย	
	A	B
A	TP	FN
B	FP	TN

รูปที่ 1 เมตริกซ์ความสับสน [13]

จากรูปที่ 1 แสดงเมตริกซ์ความสับสน (confusion matrix) ซึ่งใช้เป็นค่าพื้นฐานในการคำนวณค่าวัดประสิทธิภาพในการทำนาย โดยที่ TP คือ จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น A; TN คือ จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น B; FP คือ จำนวนข้อมูลที่จำแนกผิดว่าเป็นชั้น A ซึ่งชั้นที่แท้จริงเป็นชั้น B; FN คือ จำนวนข้อมูลที่จำแนกผิดว่าเป็นชั้น B ซึ่งชั้นที่แท้จริงเป็นชั้น A

ค่าวัดประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นมีสูตรในการคำนวณดังนี้

1) ความแม่น (accuracy)

$$= \frac{\text{จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น A และ B}}{\text{จำนวนข้อมูลทั้งหมด}}$$

$$= \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

2) ความระลึก (recall)

$$= \text{TP Rate}$$

$$= \frac{TP}{TP + FN} \quad (2)$$

3) ค่าวัดเอฟ (F-measure)

$$= \frac{2\,(\text{ความเที่ยง})\,(\text{ความระลึก})}{\text{ความเที่ยง} + \text{ความระลึก}} \quad (3)$$

โดยที่ ความเที่ยง (precision)

$$= \frac{\text{จำนวนข้อมูลที่จำแนกถูกว่าเป็นชั้น A}}{\text{จำนวนข้อมูลที่ทำนายได้ในชั้น A}} \quad (4)$$

$$= \frac{TP}{TP + FN}$$

4) ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (mean absolute error, MAE)

$$\text{MAE} = \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{n} = \sum_{i=1}^{n} \frac{|e_i|}{n} \quad (5)$$

5) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (mean square error, MSE)

$$\text{MSE} = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n} = \sum_{i=1}^{n} \frac{e_i^2}{n} \quad (6)$$

โดยที่ y_i แทนค่าจริงที่ i และ $\hat{y}_i$ แทนค่าทำนายที่ i

3. ผลการวิจัย

3.1 ผลการทำนายค่าความถูกต้องในแต่ละคำตอบจากการทำนาย และเมตริกซ์ความสับสนจากการทำนายการติดเกมของเด็กและวัยรุ่น

ผลการทำนายการติดเกมของเด็กและวัยรุ่นดังตารางที่ 1 ค่าความถูกต้องในแต่ละคำตอบจากการทำนายการติดเกมของเด็กและวัยรุ่นดังตารางที่ 2 และเมตริกซ์ความสับสนจากการทำนายการติดเกมของเด็กและวัยรุ่นดังตารางที่ 3

ตารางที่ 1 ผลการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี

วิธีการจำแนกกลุ่ม	ค่าความแม่น (ร้อยละ)	ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย	ค่าคลาดเคลื่อนกำลังสองเฉลี่ย
วิธีเพื่อนบ้านใกล้สุด k ตัว	86.09	0.1482	0.1433
วิธีต้นไม้การตัดสินใจ	**92.17**	0.1299	0.0638
วิธีโครงข่ายประสาทเทียม	89.57	0.1177	0.0919
วิธีซัพพอร์ตเวกเตอร์แมชชีน	91.30	**0.0870**	0.0870
วิธีฐานกฎ	91.30	0.1446	0.0640
วิธีการถดถอยลอจิสติกทวิภาค	91.30	0.1371	**0.0612**
วิธีนาอีฟเบส์	87.83	0.1461	0.0949

ตารางที่ 2 ค่าความถูกต้องในแต่ละคำตอบจากการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี

วิธีการจำแนกกลุ่ม	ความเที่ยง (ไม่ได้ติดเกม)	ความระลึก (ไม่ได้ติดเกม)	ค่าวัดเอฟ (ไม่ได้ติดเกม)	ความเที่ยง (ติดเกม)	ความระลึก (ติดเกม)	ค่าวัดเอฟ (ติดเกม)
วิธีเพื่อนบ้านใกล้สุด k ตัว	0.905	0.905	0.905	0.742	0.742	0.742
วิธีต้นไม้การตัดสินใจ	0.921	**0.976**	**0.948**	**0.923**	0.774	**0.842**
วิธีโครงข่ายประสาทเทียม	0.909	0.952	0.930	0.852	0.742	0.793
วิธีซัพพอร์ตเวกเตอร์แมชชีน	0.911	**0.976**	0.943	0.920	0.742	0.821
วิธีฐานกฎ	**0.930**	0.952	0.941	0.862	**0.806**	0.833
วิธีการถดถอยลอจิสติกทวิภาค	0.911	**0.976**	0.943	0.920	0.742	0.821
วิธีนาอีฟเบส์	0.898	0.940	0.919	0.815	0.710	0.759

ตารางที่ 3 เมตริกซ์ความสับสนจากการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี

วิธีการจำแนกกลุ่ม	สถานการณ์จริง	การจำแนกการติดเกมของเด็กและวัยรุ่น	
		ไม่ได้ติดเกม	ติดเกม
วิธีเพื่อนบ้านใกล้สุด k ตัว	ไม่ได้ติดเกม	76	8
	ติดเกม	8	23
วิธีต้นไม้การตัดสินใจ	ไม่ได้ติดเกม	**82**	2
	ติดเกม	7	**24**
วิธีโครงข่ายประสาทเทียม	ไม่ได้ติดเกม	80	4
	ติดเกม	8	23
วิธีซัพพอร์ตเวกเตอร์แมชชีน	ไม่ได้ติดเกม	82	2
	ติดเกม	8	23
วิธีฐานกฎ	ไม่ได้ติดเกม	80	4
	ติดเกม	6	25
วิธีการถดถอยลอจิสติกทวิภาค	ไม่ได้ติดเกม	82	2
	ติดเกม	8	23
วิธีนาอีฟเบส์	ไม่ได้ติดเกม	79	5
	ติดเกม	9	22

ตารางที่ 1 พบว่าวิธีต้นไม้การตัดสินใจมีค่าความแม่นมากที่สุดร้อยละ 92.17 ส่วนวิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดเท่ากับ 0.0870 และวิธีการถดถอยลอจิสติกทวิภาคมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดเท่ากับ 0.0612 แต่เนื่องจากวิธีต้นไม้การตัดสินใจมีความแม่นมากที่สุดและมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.0638 ใกล้เคียงกับวิธีการถดถอย ลอจิสติกทวิภาค ซึ่งมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุดเท่ากับ 0.0612 ดังนั้นวิธีต้นไม้การตัดสินใจมีประสิทธิภาพในการทำนายดีที่สุด

ตารางที่ 2 พบว่าวิธีต้นไม้การตัดสินใจมีค่าความระลึกของชั้นไม่ได้ติดเกม (0.976) ค่าวัดเอฟของชั้นไม่ได้ติดเกม (0.948) ค่าความเที่ยงของชั้นติดเกม (0.923) และค่าวัดเอฟของชั้นติดเกม (0.842) มากที่สุด วิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความระลึกของชั้นไม่ได้ติดเกม (0.976) มากที่สุด วิธีฐานกฎมีค่าความเที่ยงของชั้นไม่ได้ติดเกม (0.930) และค่าความระลึกของชั้นติดเกม (0.806) มากที่สุด และวิธีการถดถอยลอจิสติกทวิภาคมีค่าความระลึกของชั้นไม่ได้ติดเกม (0.976) มากที่สุด แต่เนื่องจากวิธีต้นไม้การตัดสินใจมีประสิทธิภาพในการทำนายดีที่สุด 4 ใน 6 ค่า ดังนั้นวิธีต้นไม้การตัดสินใจมีประสิทธิภาพในการทำนายดีที่สุด

ตารางที่ 3 พบว่าจากจำนวนเด็กและวัยรุ่น 115 คน วิธีต้นไม้การตัดสินใจ ตัวแบบสามารถทำนายจำนวนเด็กและวัยรุ่นได้ถูกต้องมากที่สุด 106 คน

(82 + 24) โดยมีจำนวนเด็กและวัยรุ่นที่จำแนกถูกว่าไม่ได้ติดเกม 82 คน และจำนวนเด็กและวัยรุ่นที่จำแนกถูกว่าติดเกม 24 คน และตัวแบบไม่สามารถทำนายจำนวนเด็กและวัยรุ่นได้ถูกต้องน้อยที่สุด 9 คน (2 + 7) โดยมีจำนวนเด็กและวัยรุ่นที่จำแนกผิดว่าติดเกมซึ่งค่าจริง ๆ แล้วไม่ได้ติดเกม 2 คน และจำนวนเด็กและวัยรุ่นที่จำแนกผิดว่าไม่ได้ติดเกมซึ่งค่าจริง ๆ แล้วติดเกม 7 คน ดังนั้นวิธีต้นไม้การตัดสินใจมีประสิทธิภาพในการทำนายดีที่สุด

4. สรุปผลการวิจัย อภิปรายผลการวิจัย และข้อเสนอแนะ

4.1 สรุปผลการวิจัย

การเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร ด้วยวิธีการจำแนกกลุ่ม 7 วิธี ได้ผลสรุปดังนี้ คือ วิธีต้นไม้การตัดสินใจมีค่าความแม่น ค่าความระลึกของชั้นไม่ได้ติดเกม ค่าวัดเอฟของชั้นไม่ได้ติดเกม ค่าความเที่ยงของชั้นติดเกมและค่าวัดเอฟของชั้นติดเกมมากที่สุดเท่ากับ 92.17 %, 0.976, 0.948, 0.923 และ 0.842 ตามลำดับ วิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความระลึกของชั้นไม่ได้ติดเกมมากที่สุด และค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดเท่ากับ 0.976 และ 0.0870 ตามลำดับ วิธีฐานกฎมีค่าความเที่ยงของชั้นไม่ได้ติดเกมและค่าความระลึกของชั้นติดเกมมากที่สุดเท่ากับ 0.930 และ 0.806 ตามลำดับ ส่วนวิธีการถดถอยลอจิสติกทวิภาคมีค่าความระลึกของชั้นไม่ได้ติดเกมมากที่สุดและค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดเท่ากับ 0.976 และ 0.0612 ตามลำดับ แต่เนื่องจากวิธีต้นไม้การตัดสินใจมีค่าวัดประสิทธิภาพในการทำนายดีที่สุด 5 ใน 9 ค่า ดังนั้นวิธีต้นไม้การตัดสินใจเป็นวิธีที่สามารถทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานครที่มีประสิทธิภาพมากที่สุด

4.2 อภิปรายผลการวิจัย

ผลการศึกษาการเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานคร โดยพิจารณาจากค่าความแม่น ความเที่ยง ความระลึก และค่าวัดเอฟ พบว่าวิธีต้นไม้การตัดสินใจเป็นวิธีที่สามารถทำนายการติดเกมของเด็กและวัยรุ่นในกรุงเทพมหานครที่มีประสิทธิภาพมากที่สุด รองลงมา คือ วิธีฐานกฎ วิธีซัพพอร์ตเวกเตอร์แมชชีน และวิธีการถดถอยลอจิสติกทวิภาค ตามลำดับ ซึ่งให้ผลสอดคล้องกับงานวิจัยของ ณัฐวุฒิ และคณะ [13] ซึ่งศึกษาเปรียบเทียบประสิทธิภาพการจำแนกประเภทโดยใช้วิธีเพื่อนบ้านใกล้สุด k ตัว วิธีต้นไม้การตัดสินใจ วิธีโครงข่ายประสาทเทียม และวิธีซัพพอร์ตเวกเตอร์แมชชีน ที่พบว่าวิธีต้นไม้การตัดสินใจดีที่สุดและให้ผลใกล้เคียงกับงานวิจัยของ วาทินี และคณะ [8] ซึ่งได้ศึกษาเพื่อเปรียบเทียบวิธีการจำแนกข้อมูลโดยเลือกใช้อัลกอริทึมโครงข่ายประสาทเทียมแบบมัลติเลเยอร์เปอร์เซ็ปตรอน ซัพพอร์ตเวกเตอร์แมชชีน นาอีฟเบส์ และเพื่อนบ้านใกล้สุด k ตัว เพื่อประเมินประสิทธิภาพค่าความแม่น ความเที่ยง ความระลึก และค่าวัดเอฟ โดยใช้ข้อมูลจาก UCI ประกอบด้วย ozone days และ adult จากผลการวิจัยอัลกอริทึมที่ดีที่สุดของข้อมูล ozone days คือ ซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบ Rbf มีค่าความแม่น 94.83 % ความเที่ยง 96 % ความระลึก 96 % และค่าวัดเอฟ 96 % ส่วนข้อมูล adult คือ ซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบโพลิโนเมียล มีค่าความแม่น 79.66 % ความเที่ยง 80 % ความระลึก 80 % และค่าวัดเอฟ 80 % แต่ให้ผลแตกต่างจากงานวิจัยอื่น ๆ อยู่บ้าง เช่น งานวิจัยของ Xhemali และคณะ [1] ได้ศึกษาการจำแนกกลุ่มของการฝึกหัดเว็บเพจโดยใช้วิธีนาอีฟเบส์ วิธีต้นไม้การตัดสินใจ และวิธีโครงข่ายประสาทเทียม ในการจำแนกกลุ่มของการฝึกหัดเว็บ-

เพจเพื่อหาอัตราความสำเร็จของวิธีการเหล่านี้ ในหลักสูตรฝึกหัดเว็บเพจ ผลการวิจัยนี้พบว่าวิธีนาอีฟ-เบส์ดีที่สุด โดยมีค่าวัดเอฟมากกว่า 97 % ส่วนงานวิจัยของ Tuisima และคณะ [2] ได้ศึกษาเรื่องการจัดจำแนกกลุ่มของระดับการติดเกมคอมพิวเตอร์ในนักศึกษาชั้นมัธยมศึกษาปีที่ 1-3 โดยใช้โครงข่ายประสาทเทียม โดยจำแนกระดับการติดเกมโดยใช้โครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับและเปรียบเทียบผลกับต้นไม้การตัดสินใจ วัดความแม่นของตัวแบบโดยใช้วิธี 10-fold cross validation ผลการทดลองพบว่าความแม่นของวิธีโครงข่ายประสาทเทียมด้วยอัลกอริทึมการแพร่แบบย้อนกลับสำหรับเกมระยะยาว เกมที่ไม่มีกฎเกณฑ์แน่นอน เกมระยะเวลาจริง และเกมพื้นฐานเท่ากับ 97.75, 91.35, 90.00 และ 97.73 ตามลำดับ ส่วนความแม่นของอัลกอริทึมต้นไม้การตัดสินใจเท่ากับ 88.76, 92.63, 87.50 และ 90.91 ตามลำดับ และงานวิจัยของ Chou [3] ศึกษาการใช้ Tic-Tac-Toe สำหรับเรียนรู้การจำแนกกลุ่มการทำเหมืองข้อมูลและการประเมินผล ผลการทดลองพบว่าวิธีเพื่อนบ้านใกล้สุด 3 ตัว มีอัตราความแม่นในการจำแนกกลุ่มของเกม Tic-Tac-Toe มากที่สุดเท่ากับ 99 % วิธีการจำแนกกลุ่มทั้ง 4 วิธี คือ วิธีโครงข่ายประสาทเทียม วิธีซัพ-พอร์ตเวกเตอร์แมชชีน วิธีการถดถอยลอจิสติก และวิธีเพื่อนบ้านใกล้สุด 3 ตัว มีอัตราความแม่นมากกว่า 98 % โดยทั้ง 4 วิธี มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ

ส่วนการเปรียบเทียบประสิทธิภาพในการทำนายการติดเกมของเด็กและวัยรุ่นโดยพิจารณาจากค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าคลาดเคลื่อนกำลังสองเฉลี่ย พบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดเท่ากับ 0.0870 รองลงมา คือ วิธีโครงข่ายประสาทเทียมมีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยเท่ากับ 0.1177 ส่วนวิธีการถดถอยลอจิสติกทวิภาคมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดเท่ากับ 0.0612 รองลงมา คือ วิธีต้นไม้ตัดสินใจมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.0638 ซึ่งให้ผลคล้ายกับงานวิจัยของ สายชล [14] ในการเปรียบเทียบประสิทธิภาพในการทำนายการเป็นโรคเบาหวาน ซึ่งพบว่าวิธีโครงข่ายประสาทเทียมมีค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุด นอกจากนี้ยังให้ผลใกล้เคียงกับงานวิจัยของ ณัฐวุฒิ และคณะ [13] ในการเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มโดยใช้วิธีเพื่อนบ้านใกล้สุด k ตัว วิธีต้นไม้การตัดสินใจ วิธีโครงข่ายประสาทเทียม และวิธีซัพพอร์ตเวกเตอร์แมชชีน และการเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศอินเดีย ของ สุรวัชร [15] ที่พบว่าวิธีการจำแนกกลุ่มที่มีประสิทธิภาพในการทำนายดีที่สุดโดยเปรียบเทียบจากค่าคลาดเคลื่อนกำลังสองเฉลี่ย คือ วิธีต้นไม้การตัดสินใจ แต่ต่างจากงานวิจัยของ สายชล [14] ในการเปรียบเทียบประสิทธิภาพในการทำนายการเป็นโรคเบาหวาน ซึ่งพบว่าวิธีโครงข่ายประสาทเทียมมีค่าคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด และงานวิจัยของ Vijayarani และ Dhayanand [16] ซึ่งศึกษาข้อมูลโรคไตเรื้อรังโดยใช้วิธีซัพพอร์ตเวกเตอร์แมชชีนและวิธีโครงข่ายประสาทเทียม จากการศึกษาพบว่าวิธีโครงข่ายประสาทเทียมมีประสิทธิ-ภาพในการทำนายได้ดีกว่าวิธีซัพพอร์ตเวกเตอร์แมชชีน สาเหตุที่การเปรียบเทียบประสิทธิภาพให้ผลแตกต่างกันอาจเนื่องมาจากเรื่องที่ศึกษาอยู่ในสาขาต่างกันจึงให้ประสิทธิภาพในการทำนายที่แตกต่างกัน

4.3 ข้อเสนอแนะ

4.3.1 ตัวแปรที่นำมาใช้ในการงานวิจัยนี้เป็นเพียงส่วนหนึ่งของการพิจารณาการติดเกมของเด็กและวัยรุ่นเท่านั้น เพื่อให้การทำนายมีประสิทธิภาพ

มากขึ้น ควรเพิ่มตัวแปรที่เกี่ยวข้องอื่น ๆ อีก ได้แก่ ความคิดเห็นของเพื่อน ๆ เกี่ยวกับการติดเกมของผู้เล่นเกม

4.3.2 เพื่อให้ผลสรุปครอบคลุมกว้างขวางเพิ่มขึ้น ควรจะศึกษาวิธีอื่น ๆ ที่เป็นเทคนิคการทำเหมืองข้อมูลในการจำแนกกลุ่มเหมือนกัน เช่น วิธีโครงข่ายความเชื่อของเบส์เซียน (Bayesian belief network method)

4.3.3 เพื่อให้ได้ข้อสรุปของผลการวิเคราะห์ข้อมูลที่มีความสมบูรณ์มากขึ้น อาจใช้อัลกอริทึมประเภทอื่น ๆ โดยวิธีเพื่อนบ้านใกล้สุด k ตัว ยังมีอัลกอริทึม KStar และ LWL วิธีต้นไม้การตัดสินใจมีอัลกอริทึม 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 เป็นต้น

5. รายการอ้างอิง

[1] Xhemali, D., Hinde, C.J. and Stone, R.G., 2009, Naïve Bayes vs. decision trees vs. neural networks in the classification of training web pages, Int. J. Comp. Sci. 4(1): 16-23.

[2] Tuisima, S., Vachirapan, K. and Sinthupinyo, S., 2012, Classification of computer game addiction level in students in secondary education (M. 1-3) using neural network, pp. 19-23, The 2nd International Conference on Management and Artificial Intelligence IPEDR, Vol. 35 IACSIT Press, Singapore.

[3] Chou, C.H., 2013, Using Tic-Tac-Toe for learning data mining classifications and evaluations, Int. J. Inform. Edu. Technol. 3: 437-441.

[4] พยูน พาณิชย์กุล, 2548, การพัฒนาระบบดาต้าไมน์นิ่งโดยใช้ Decision Tree, โครงการพัฒนาระบบงาน, วิทยานิพนธ์ปริญญาโท, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง , กรุงเทพฯ, 62 น.

[5] Wu, X. and Kumar, V., 2009, The Top Ten Algorithms in Data Mining, University of Minnesota Department of Computer Science and Engineering, Minneapolis, CRC Press, Boca Raton, 215 p.

[6] รุจิรา ธรรมสมบัติ, 2554, ระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือโดยใช้ต้นไม้ตัดสินใจ, สาขาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์, กรุงเทพฯ, 68 น.

[7] Berson, A. and Smith, S.J., 1997, Data Warehousing, Data Mining, and OLAP, McGraw-Hill, New York, 612 p.

[8] วาทินี นุ้ยเพียร, พยุง มีสัจจะ และเดช ธรรมศิริ, 2553, การเปรียบเทียบประสิทธิภาพและวิเคราะห์การจำแนกข้อมูลโดยใช้โครงข่ายประสาทเทียม ซัพพอร์ตเวกเตอร์แมชชีน นาอีฟเบย์ และเคเนียรเรสต์เนเบอร์, น. 131-138, การประชุมวิชาการระดับชาติทางด้านคอมพิวเตอร์

และเทคโนโลยีสารสนเทศครั้งที่ 5, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, กรุงเทพฯ.

[9] Murti, S. and Mahantappa, M., 2012, Using rule based classifiers for the predictive analysis of breast cancer recurrence, J. Inform. Eng. Appl. 2(2): 12-19.

[10] กัลยา วานิชย์บัญชา, 2552, การวิเคราะห์ข้อมูลหลายตัวแปร, บริษัทธรรมสารจำกัด, กรุงเทพฯ, 589 น.

[11] Larose, D.T., 2006, Data Mining: Methods and Models, John Wiley and Sons, New Jersey, 322 p.

[12] Larose, D.T., 2005, Discovering Knowledge in Data: An Introduction to Data Mining, John Wiley and Sons, Hoboken, N.J., 222 p.

[13] ณัฐวุฒิ ศิริกุลรุ่งโรจน์, รัฐธชัย บุญวิเศษ, สหัสนัยน์ ไพสานต์ และสุรวัชร ศรีเปารยะ, 2556, การเปรียบเทียบประสิทธิภาพการจำแนกประเภทโดยใช้วิธีความใกล้เคียงกันมากที่สุด วิธีแผนภาพต้นไม้เพื่อการตัดสินใจ วิธีโครงข่ายประสาทเทียมและวิธีซัพพอร์ตเวกเตอร์แมชชีน, ปัญหาพิเศษปริญญาตรี, สาขาวิชาสถิติ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ, 132 น.

[14] สายชล สินสมบูรณ์ทอง, 2560, การเปรียบเทียบประสิทธิภาพในการทำนายผลการเป็นโรคเบาหวาน, ว.วิทยาศาสตร์และเทคโนโลยี 26(5): 839-853.

[15] สุรวัชร ศรีเปารยะ, 2559, การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศอินเดีย, ว.วิทยาศาสตร์และเทคโนโลยี 25(5): 839-853.

[16] Vijayarani, S. and Dhayanand, S., 2015, Kidney disease prediction using SVM and ANN algorithms, Int. J. Cybernet. Inform. 4(4): 13-25.