

บทที่ 3 วิธีการดำเนินงาน

3.1 อุปกรณ์ในการวิจัย

3.1.1 อุปกรณ์ที่มีอยู่แล้ว

- 1) เครื่องคอมพิวเตอร์
- 2) เครื่องพิมพ์เลเซอร์
- 3) โปรแกรมสำเร็จรูป Weka version 3.9
- 4) โปรแกรมสำเร็จรูป Microsoft Excel 2007

3.1.2 อุปกรณ์ที่ต้องการเพิ่ม

- 1) แบบสอบถาม

3.2 ขั้นตอนการดำเนินงานวิจัย

งานวิจัยนี้ได้ทำการศึกษาเรื่องการเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดเกมของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร จากแบบสอบถามภาวะการติดเกมของเด็กและวัยรุ่นระหว่างวิธีความใกล้เคียงกันมากที่สุด วิธีแผนภาพต้นไม้เพื่อการตัดสินใจ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีฐานกฎ วิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม และวิธีนาอิว เบย์ โดยมีขั้นตอนการดำเนินงานดังนี้

3.2.1 การเก็บรวบรวมข้อมูล

ทำการเก็บรวบรวมข้อมูลจากแบบสอบถามภาวะการติดเกมของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร 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

- 1) เพศ
 - 1.1 เพศชาย
 - 1.2 เพศหญิง
- 2) อายุ (หน่วย : ปี)
- 3) ระดับการศึกษา
 - 3.1 ประถมศึกษา
 - 3.2 มัธยมศึกษาตอนต้น
 - 3.3 มัธยมศึกษาตอนปลาย
 - 3.4 ปวช.
 - 3.5 ปวส.
 - 3.6 อนุปริญญา
 - 3.7 ปริญญาตรี
 - 3.8 ปริญญาโท
 - 3.9 ปริญญาเอก

- 4) รายได้ต่อเดือน
 - 4.1 ต่ำกว่า 10,000 บาท
 - 4.2 10,000 - 20,000 บาท
 - 4.3 20,001 - 30,000 บาท
 - 4.4 30,001 - 40,000 บาท
 - 4.5 40,001 - 50,000 บาท
 - 4.6 มากกว่า 50,000 บาท
- 5) การใช้สื่อสังคมออนไลน์
 - 5.1 ใช้เป็นประจำ
 - 5.2 ไม่ได้ใช้เป็นประจำ
- 6) ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันจันทร์-ศุกร์ (หน่วย : ชั่วโมง)
- 7) ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันเสาร์-อาทิตย์/วันหยุด (หน่วย : ชั่วโมง)
- 8) ความคิดเห็นของผู้ปกครองเกี่ยวกับการติตเกมของลูก
 - 8.1 ติตเกม
 - 8.2 เกือบติตเกม
 - 8.2 ไม่ติตเกม
- 9) การหมกมุ่นในเกม
 - 9.1 หมกมุ่นในเกม
 - 9.2 ไม่ได้หมกมุ่นในเกม
- 10) ความสูญเสียหน้าที่ที่รับผิดชอบ
 - 10.1 ไม่มีความรับผิดชอบ
 - 10.2 มีความรับผิดชอบ
- 11) การสูญเสียในการควบคุมตัวเอง อารมณ์ พฤติกรรม และเวลา
 - 11.1 ควบคุมตัวเองไม่ได้
 - 11.2 ควบคุมตัวเองได้
- 12) ภาวะการติตเกม
 - 12.1 ติตเกม
 - 12.2 ไม่ได้ติตเกม

3.2.2 การบันทึกข้อมูล

ทำการบันทึกข้อมูลผลการสอบถามการติตเกมของเด็กและวัยรุ่น จำนวน 1,800 ชุด แต่จากการตรวจสอบแบบสอบถามเบื้องต้นพบว่าแบบสอบถามไม่สมบูรณ์จำนวน 42 ชุด ทำการตัดทิ้ง จึงเหลือแบบสอบถามจำนวน 1,758 ชุด แล้วทำการบันทึกลงในโปรแกรม Microsoft Excel โดยให้แนวคอลัมน์เป็นเพศ อายุ ระดับการศึกษา รายได้ต่อเดือน การใช้สื่อสังคมออนไลน์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันจันทร์-ศุกร์ ระยะเวลาเฉลี่ยต่อวันที่เล่นเกมในช่วงวันเสาร์-อาทิตย์/วันหยุด ความคิดเห็นของผู้ปกครองเกี่ยวกับการติตเกมของลูก การหมกมุ่นในเกม ความสูญเสียหน้าที่ที่รับผิดชอบ การสูญเสียในการควบคุมตัวเอง อารมณ์ พฤติกรรมและเวลา และ ภาวะการติตเกม ส่วนแนวแถวเป็นลำดับที่ของการติตเกมของเด็กและวัยรุ่น

3.2.3 การแบ่งข้อมูล

นำข้อมูลทั้งหมดมาแบ่งออกเป็น 3 ส่วน จากแบบสอบถามภาวะการติดเกมของเด็กและวัยรุ่น จำนวน 1,758 ชุด สำหรับการวิเคราะห์ข้อมูล โดยแบ่งข้อมูลออกเป็นสัดส่วนดังนี้ (พยุณ พาณิชยกุล, 2548)

ส่วนที่ 1 : ข้อมูลชุดฝึกหัด (Training Data Set) เพื่อนำไปสร้างตัวแบบ

มีข้อมูล 80% ของข้อมูลทั้งหมด คิดเป็น 1,406 ชุด แต่ยังคงตรวจพบแบบสอบถามที่ไม่สมบูรณ์จำนวน 13 ชุด ทำการตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ในส่วนที่ 1 จำนวน 1,393 ชุด

ส่วนที่ 2 : ข้อมูลชุดทดสอบประสิทธิภาพ (Evaluation Data Set) เพื่อนำไปทดสอบความถูกต้องของตัวแบบ

มีข้อมูล 15% ของข้อมูลทั้งหมด คิดเป็น 264 ชุด แต่ยังคงตรวจพบแบบสอบถามที่ไม่สมบูรณ์จำนวน 14 ชุด ทำการตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ในส่วนที่ 2 จำนวน 250 ชุด

ส่วนที่ 3 : ข้อมูลชุดทดสอบ (Testing Data Set) เพื่อนำไปทำนายตัวแบบ

มีข้อมูล 5% ของข้อมูลทั้งหมด คิดเป็น 88 ชุด แต่ยังคงตรวจพบแบบสอบถามที่ไม่สมบูรณ์จำนวน 5 ชุด ทำการตัดทิ้ง จึงเหลือข้อมูลที่สมบูรณ์ 83 ชุด ทำการเก็บข้อมูลเพิ่มเติมอีก 32 ชุด เพื่อให้ครบ 1,758 ชุด ซึ่งจะได้ข้อมูลที่สมบูรณ์ในส่วนที่ 3 จำนวน 115 ชุด

3.2.4 เครื่องมือที่ใช้ในการวิจัย

ในการวิเคราะห์ข้อมูลครั้งนี้ใช้โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) เวอร์ชัน 3.9 ซึ่งเป็นโปรแกรมที่สามารถดาวน์โหลดได้จากเว็บไซต์ ซึ่งอยู่ภายใต้การควบคุมของ GPL License ซึ่งโปรแกรม WEKA ได้ถูกพัฒนามาจากภาษาจาวาทั้งหมด ซึ่งเป็นที่นิยมในการใช้งานด้านการทำเหมืองข้อมูล

3.2.5 การวิเคราะห์ข้อมูล

ในการทำวิจัยครั้งนี้ ผู้วิจัยได้จัดข้อมูลแต่ละชุดออกเป็น 3 ส่วน โดยส่วนที่ 1 ใช้ข้อมูล 80 เปอร์เซ็นต์ของข้อมูลทั้งหมดในการสร้างตัวแบบ ส่วนที่ 2 ใช้ข้อมูล 15 เปอร์เซ็นต์ของข้อมูลทั้งหมดในการทดสอบความถูกต้องของ ตัวแบบ ส่วนที่ 3 ใช้ข้อมูล 5 เปอร์เซ็นต์ของข้อมูลทั้งหมดในการทำนายตัวแบบ แปลงไฟล์ข้อมูลให้เป็นนามสกุล *.csv เพื่อใช้วิเคราะห์ประสิทธิภาพการจำแนกกลุ่มข้อมูลในโปรแกรม WEKA ซึ่งเป็นโปรแกรมที่สามารถนำมาทดสอบอัลกอริทึมของวิธีการจำแนกกลุ่มได้เนื่องจากมีอัลกอริทึมที่ถูกระบุไว้ให้เลือกใช้ในโปรแกรมครบตามที่กำหนด ผู้วิจัยได้กำหนดวิธีการจำแนกกลุ่มเพื่อนำมาทดสอบดังนี้

1) วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor : KNN) ใช้อัลกอริทึมชนิด IBk เนื่องจากเป็นฟังก์ชันหลักที่สนใจ ซึ่งเป็นพื้นฐานของอัลกอริทึม 8.1 อัลกอริทึม IBk ยังสามารถกำหนดน้ำหนักระยะห่างและทางเลือก (option) เพื่อกำหนดค่า k โดยใช้ cross-validation (Kumar, V. and Wu, X., 2009)

2) วิธีแผนภาพต้นไม้เพื่อการตัดสินใจ (Decision Tree) ใช้อัลกอริทึมชนิด J48 ซึ่งพัฒนามาจาก ID3 สามารถใช้ได้กับข้อมูลแบบไม่ต่อเนื่องและแบบต่อเนื่อง ต่างจาก ID3 ที่ใช้ได้เพียงข้อมูลแบบไม่ต่อเนื่องเท่านั้น (รุจิรา ธรรมสมบัติ, 2555)

3) วิธีโครงข่ายประสาทเทียม (Neural Network) ใช้อัลกอริทึมชนิดเพอร์เซปตรอนหลายชั้น (Multilayer Perceptron) โดยกำหนดค่าอัตราการเรียนรู้ (Learning Rate) เป็น 0.1 ค่าโมเมนตัม (Momentum) เป็น 0.9 (Berson, A. and Smith, S. J., 1997) จำนวนรอบการสอน (Training Time) 20,000 รอบ การวิจัยครั้งนี้ใช้อัลกอริทึมของวิธีโครงข่ายประสาทเทียมชนิดเพอร์เซปตรอนหลายชั้นที่มีชั้นซ่อน (Hidden Layer) 1 ชั้น แม้ว่าโครงสร้างโครงข่ายประสาทเทียมที่ซับซ้อนสามารถมีชั้นซ่อนมากกว่า 1 ชั้น แต่ในทางปฏิบัติ การกำหนดชั้นซ่อน 1 ชั้น ก็เพียงพอต่อการวิเคราะห์ข้อมูล (Berson, A. and Smith, S. J., 1997)

4) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ใช้อัลกอริทึม SMO ชนิดโพลีโนเมียลเคอร์เนล (Polynomial Kernel) เนื่องจากงานวิจัยที่อ้างอิงจากวาทีน น้อยเพียร และคณะ (2553) ได้ผลว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนที่ใช้อัลกอริทึมชนิดโพลีโนเมียลเคอร์เนลดีที่สุด

5) วิธีฐานกฎ (Rules-Based Method) ใช้ชุดลำดับของกฎมาสร้างรูปแบบการแยกประเภทข้อมูล โดยส่วนใหญ่แล้วจะใช้กฎที่เป็น If ... then ซึ่งเป็นกฎอย่างง่าย (Murti, S. and Mahantappa, M., 2012) ใช้อัลกอริทึม Decision Table เป็นเครื่องมือที่ใช้แสดงเงื่อนไขการตัดสินใจและเลือกการทำงานหรือกระทำกิจกรรมภายใต้เหตุการณ์ของเงื่อนไขที่ระบุ วิธีการตัดสินใจแบบ Decision Table จะเป็นตาราง 2 มิติ

6) วิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม (Binary Logistic Regression Method) เป็นการวิเคราะห์การถดถอยแบบหนึ่งโดยที่ตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่มีค่าได้เพียง 2 ค่า ส่วนตัวแปรอิสระอาจจะเป็นตัวแปรเชิงปริมาณหรือเชิงคุณภาพ หรืออาจจะมีทั้งตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพก็ได้ (กัลยา วานิชย์บัญชา, 2552)

7) วิธีนาอิว เบย์ (Naïve Bayes Method) คืออัลกอริทึมที่ใช้หลักการของความน่าจะเป็นในการคัดกรองแต่ละคำตอบ (Class) โดยมีคำตอบ 2 คำตอบ (Larose, D. T., 2006)

การนำผลการวิเคราะห์ข้อมูลมาประเมินผลเพื่อเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดเกมของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร ใช้วิธีการจำแนกกลุ่ม 7 วิธี โดยนำข้อมูลที่ได้จากการวิเคราะห์มาเปรียบเทียบว่าวิธีในการจำแนกกลุ่มข้อมูลวิธีใดระหว่างวิธีความใกล้เคียงกันมากที่สุด วิธีแผนภาพต้นไม้เพื่อการตัดสินใจ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีฐานกฎ วิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม และวิธีนาอิว เบย์ ที่มีค่าความถูกต้อง (Accuracy) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-Measure) มากที่สุด ส่วนค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error : MSE) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) ที่มีค่าน้อยที่สุด จะทำให้มีประสิทธิภาพในการทำนายผลดีที่สุด โดยที่ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเป็นมาตรวัดการประเมินค่าได้ดี เนื่องจากค่าความคลาดเคลื่อนกำลังสองเฉลี่ยประกอบด้วยทั้งความแปรปรวนและความเอนเอียง (Larose, D. T., 2005) ส่วนค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยก็มีในบางงานวิจัยนำมาเปรียบเทียบกันด้วย ดังนั้นในงานวิจัยนี้จึงใช้ทั้งค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยและค่าความคลาดเคลื่อนกำลังสองเฉลี่ย