

การเปรียบเทียบประสิทธิภาพจำแนกข้อมูลการเล่นเกมด้วยเทคนิคเหมืองข้อมูล

ถิรพร เหมือนอูย์¹ ทวีศักดิ์ เพชรคนชม¹ นิติมา ลักขณานุรักษ์² ไกรุ่ง เสงพะระพรหม^{1*}
และสุพจน์ เสงพะระพรหม¹

¹สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

²สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร

*kairung2011.heng@gmail.com

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพเทคนิคที่เหมาะสมสำหรับการจำแนกข้อมูลการเล่นเกม โดยทำการทดลองด้วยโปรแกรม Weka 3.8.4 ด้วยเทคนิคการจำแนกข้อมูล 2 เทคนิค ได้แก่ J48, JRip โดยทำการทดสอบกับ ข้อมูลการเล่นเกม จำนวน 3,196 ตัวอย่าง โดยใช้คุณลักษณะ 37 คุณลักษณะ

ผลการวิจัยพบว่า เทคนิคที่ให้ประสิทธิภาพในการจำแนกข้อมูลได้ดีที่สุดคือเทคนิคต้นไม้ตัดสินใจหรือ J48 โดยให้ค่า ความแม่นยำในการจำแนกข้อมูลที่ 99.44 % และรองลงมา คือ JRip ซึ่งให้ค่าความแม่นยำในการจำแนกข้อมูลที่ 99.19 % ตามลำดับ แสดงให้เห็นว่าการใช้เทคนิค J48 เป็นแนวทางที่เหมาะสมสำหรับการจำแนกข้อมูลการเล่นเกม

คำสำคัญ: คุณลักษณะ การจำแนกข้อมูล ต้นไม้ตัดสินใจ J48 JRip

A Comparison of Efficiency in Classifying Gaming Data using Data Mining Techniques

Thiraporn Hmeanai¹ Taveesak Pechkonchom¹ Nitima Lukkananuruk²
Kairung Hengpraphorm^{1*} and Supojn Hengpraphorm¹

¹ *Data Science, Faculty of Science and Technology, Nakhon Pathom Rajabhat University*

² *Information Technology and Digital Innovation, Faculty of Science, Silpakorn University*

* *Kairung2011.heng@gmail.com*

Abstract

The aim of this research is to compare the efficiency of the suitable technique for gaming data classification. In the experiment, Weka 3.8.4 software is used with two data classification techniques: J48 and JRip to test with 3,196 gaming data with 37 features.

The experimental results show that J48 gives the best result of the classification accuracy (99.43%). The second technique is JRip with the classification accuracy = 99.19% respectively. Therefore, J48 is the suitable technique for gaming data classification.

Keywords: Feature, Classification, Decision Tree, J48, JRIP

1. บทนำ

เกม เป็นลักษณะของกิจกรรมของมนุษย์เพื่อประโยชน์อย่างใดอย่างหนึ่ง เช่น เพื่อความสนุกสนานบันเทิง เพื่อฝึกทักษะ และเพื่อการเรียนรู้ เป็นต้น และในบางครั้งอาจใช้เพื่อประโยชน์ทางการศึกษาได้

เกมประกอบด้วยเป้าหมาย กฎเกณฑ์ การแข่งขันและปฏิสัมพันธ์ เกมมักจะเป็นการแข่งขันทางจิตใจหรือด้านร่างกาย หรือทั้งสองอย่างรวมกัน ซึ่งส่งผลให้เกิดพัฒนาการของทักษะ ใช้เป็นรูปแบบของการออกกำลังกาย หรือการศึกษา บทบาทสมมุติและจิตศาสตร์ เป็นต้น

วิดีโอเกม (อังกฤษ: Video game) คือ เครื่องเกมอิเล็กทรอนิกส์ที่มนุษย์ใช้ส่วนต่อประสานกับผู้ใช้ (user interface) ส่งผลการกระทำ (input) กลับเข้าไปยังหน่วยประมวลผลกลาง (Central Processor Unit) ในตัวเครื่อง ให้คิดคำนวณแล้วแสดงผลโต้ตอบกลับมาด้วย แสง-เสียง-การสั่น-หรือภาพบนจอภาพ วิดีโอ คำว่า วิดีโอ ในวิดีโอเกม แต่เดิมหมายถึงอุปกรณ์แสดงภาพแบบแรสเตอร์ แต่ปัจจุบันสามารถใช้เรียกอุปกรณ์แสดงภาพใด ๆ ก็ได้ที่สร้างภาพสองมิติหรือสามมิติขึ้นมา ตัวอย่างเช่นคอมพิวเตอร์ส่วนบุคคล และเครื่องเล่นวิดีโอเกม อุปกรณ์เหล่านี้เป็นไปตั้งแต่ตั้งแต่เมนเฟรมคอมพิวเตอร์ ไปจนถึงอุปกรณ์มือถือขนาดเล็ก วิดีโอเกมแบบเฉพาะอย่างเช่น เกมตู้ เคยมีแพร่หลายในอดีต แต่ปัจจุบันค่อย ๆ มีใช้น้อยลง วิดีโอเกมได้พัฒนาไปจนกลายเป็นอุตสาหกรรมและงานศิลปะ

อุปกรณ์นำข้อมูลเข้าเป็นที่ใช้กันในการบังคับวิดีโอเกมเรียกว่า อุปกรณ์ควบคุมเกม (game controller) และแตกต่างกันไปในเครื่องเล่นแบบต่าง ๆ ตัวอย่างเช่น อุปกรณ์ควบคุมอาจประกอบด้วยเพียงแค่ปุ่มกดและก้านควบคุม (joystick) หรืออาจมีปุ่มกดถึงสิบปุ่ม และอาจจะมีมากกว่าหนึ่งก้านควบคุมก็ได้ เกมคอมพิวเตอร์ส่วนบุคคลยุคแรกจำเป็นต้องใช้คีย์บอร์ดในการเล่น หรือต้องการให้ผู้ใช้ซื้อจก้านควบคุมที่มีปุ่มกดอย่างน้อยหนึ่งปุ่มด้วย เกมคอมพิวเตอร์สมัยใหม่จำนวนมากให้ผู้เล่นหรือต้องการให้ผู้เล่นใช้คีย์บอร์ดควบคู่ไปกับเมาส์ อุปกรณ์ควบคุมเกมที่พบได้บ่อยกันคือ เกมแพด เมาส์ คีย์บอร์ด และก้านควบคุม ในหลายปีที่ผ่านมา มีวิธีการนำข้อมูลเข้าเพิ่มเติมเช่น การให้ผู้เล่นสังเกตการณ์ในเครื่องเล่นที่ใช้กล้อง และระบบจอสัมผัสบนโทรศัพท์มือถือ

จากข้อความที่กล่าวมาข้างต้น จึงนำข้อมูลของผู้เล่นเกมมาใช้ในการทดสอบการจำแนกประเภทข้อมูลที่มีผู้เล่นเกม โดยในการทดลองนี้ได้ใช้โปรแกรม Weka-3-8-4 ในการทดสอบประสิทธิภาพในการเปรียบเทียบการจำแนกข้อมูลโดยเทคนิค 2 แบบ คือ ต้นไม้ตัดสินใจ (J48) และ JRip

1. วัตถุประสงค์การวิจัย

เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลผู้เล่นเกมโดยใช้เทคนิค 2 เทคนิค คือ ต้นไม้ตัดสินใจ (J48) และ JRip

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

จิราภา เลหาหะวรรณท์ และคณะ (2556) ทำการวิจัย เรื่อง การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ โดยสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง วิจัยนี้นำเสนอ “ระบบแนะนำแขนงวิชา” เพื่อเป็นเครื่องมือช่วยตัดสินใจเลือกแขนงวิชา โดยโครงงานชิ้นนี้ได้ใช้ข้อมูลผลการเรียนและผลการวัดความสามารถด้านต่างๆ ที่เกี่ยวข้องมาสร้างแบบจำลองพยากรณ์โดยเปรียบเทียบเทคนิคการทำเหมืองข้อมูล 5 เทคนิค และพยากรณ์ผ่านเทคนิค “Ensemble” โดยพบว่าการพยากรณ์มีความแม่นยำอยู่ที่ 72.92% โดยแขนงวิศวกรรมซอฟต์แวร์สามารถทำนายได้แม่นยำถึง 86.67% ซึ่งประโยชน์จากการพัฒนาระบบนี้ทำให้นักศึกษาทราบถึงแขนงวิชาที่เหมาะสมกับตนเองมากที่สุดถึงน้อยที่สุดพร้อมระดับความน่าจะเป็น ซึ่งจะส่งผลให้นักศึกษาตัดสินใจเลือกสาขาที่เหมาะสมกับตนเองได้ และมีโอกาสประสบความสำเร็จในการเรียน

นิตินันท์ มาตา และคณะ (2558) ทำการวิจัย เรื่อง การเปรียบเทียบประสิทธิภาพเทคนิคการลดมิติของข้อมูลสำหรับค้นหาปัจจัย และสร้างโมเดลการจำแนกกลุ่มการระบายน้ำของประตูระบายน้ำ งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบ

ประสิทธิภาพเทคนิคการลดมิติของข้อมูลสำหรับสร้างโมเดลการจำแนกกลุ่มการระบายน้ำของประตูระบายน้ำฝักให้ และหาปัจจัยที่สำคัญสำหรับเป็นข้อมูลในการจำแนกกลุ่มการระบายน้ำ โดยใช้ชุดข้อมูลปัจจัยที่เกี่ยวข้องกับการระบายน้ำของประตูระบายน้ำฝักให้ พบว่าโมเดลการจำแนกกลุ่มที่มีประสิทธิภาพเหมาะสมกับข้อมูลการระบายน้ำมากที่สุด คือ Decision Tree แบบ J48 (Accuracy = 95.048%, Precision = 0.950, Recall = 0.949) และนำโมเดล J48 มาเพิ่มประสิทธิภาพในการจำแนกกลุ่มและหาปัจจัยที่จำเป็นสำหรับการพยากรณ์การระบายน้ำนั้น พบว่าโมเดล J48 ใช้ร่วมกับเทคนิคการลดคุณลักษณะของข้อมูลแบบ CFS (Correlation-based Feature Selection) ทำให้ประสิทธิภาพในการจำแนกกลุ่มการระบายน้ำเพิ่มขึ้นจาก 95.048% เป็น 96.286% นอกจากนี้ทำให้ทราบถึงปัจจัยที่สำคัญซึ่งจำเป็นต้องใช้ในการพยากรณ์การระบายน้ำของประตูระบายน้ำฝักให้ มี 4 ปัจจัย ได้แก่ เดือน ปริมาณน้ำท้ายประตูระบายน้ำ ปริมาณน้ำที่ปล่อยจากเขื่อนเจ้าพระยาและปริมาณน้ำที่ระบายสู่แม่น้ำน้อย

พัฒนาพงษ์ ดลรัตน์ และจारी ทองคำ (2560) ได้ทำวิจัย เรื่อง การเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพงาน วิจัยนี้จึงมีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพ งานวิจัยนี้ได้ใช้ 6 เทคนิคที่มีประสิทธิภาพในการสร้างแบบจำลอง คือ C4.5, Random Forest, Random Tree, Reduced Error Pruning (REP Tree), k-Nearest Neighbors (k-NN) และ Support Vector Machine(SVM) และวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความระลึก (Recall) จากการศึกษาพบว่าแบบจำลอง C4.5 มีประสิทธิภาพในการพยากรณ์ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพมากที่สุดถึง 95.36%

วิธีดำเนินการวิจัย

1. ขั้นตอนการดำเนินการวิจัยประกอบด้วย 3 ขั้นตอน ดังนี้

เกม เป็นลักษณะของกิจกรรมของมนุษย์เพื่อประโยชน์อย่างใดอย่างหนึ่ง เช่น เพื่อความสนุกสนานบันเทิง เพื่อฝึกทักษะ และเพื่อการเรียนรู้ เป็นต้น และในบางครั้งอาจใช้เพื่อประโยชน์ทางการศึกษาได้

1.1 ศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้องกับงานวิจัย

1.2 วิเคราะห์ข้อมูลที่ได้จากการศึกษาในขั้นที่ 1 โดยวิเคราะห์จากข้อมูลที่เลือกมา ดังนี้

1.3 การจำแนกประเภทข้อมูล เป็นขั้นตอนการทดสอบการจำแนกประเภทข้อมูล หลังจากทำการเลือกคุณลักษณะของข้อมูลที่มีความเหมาะสมและสอดคล้องมากที่สุด โดยทำการเปรียบเทียบค่าเฉลี่ยความถูกต้องของการจำแนกประเภทข้อมูลของแต่ละกลุ่มการทดสอบ ด้วยวิธี JRip และ Decision Tree

2. เครื่องมือการวิจัย

2.1 Weka 3.8.4

3. กลุ่มเป้าหมาย/ ประชากรและกลุ่มตัวอย่าง

3.1 กลุ่มตัวอย่าง คือ ข้อมูลที่นำมา 3196 ตัวอย่าง

4. การทดลองใช้โปรแกรม Weka version 3.8.4

การทดลองใช้ โปรแกรม Weka version 3.8.4 ที่พัฒนาขึ้นเพื่อจำแนกข้อมูลการเล่นเกมนักเรียนด้วยเทคนิคเหมืองข้อมูลจำนวน 3196 ตัวอย่าง โดยใช้ 2 เทคนิค คือ JRip และ ต้นไม้ตัดสินใจ (J48) จะแสดงในภาพที่ 1 และ ภาพที่ 2

ตารางที่ 1 ผลการเปรียบเทียบประสิทธิภาพเทคนิคการจำแนกข้อมูลการเล่นเกมส์ด้วยเทคนิค JRIP และต้นไม้ตัดสินใจ

Techniques	Accuracy
Decision Tree (J48)	<u>99.44 %**</u>
JRIP	99.19 %

จากตารางสามารถสรุปได้ว่าเทคนิคที่ให้ประสิทธิภาพในการจำแนกข้อมูลได้ดีที่สุดคือเทคนิคต้นไม้ตัดสินใจหรือ J48 โดยให้ค่าความแม่นยำในการจำแนกข้อมูลที่ 99.44 % และรองลงมา คือ JRIP ซึ่งให้ค่าความแม่นยำในการจำแนกข้อมูลที่ 99.19 % ตามลำดับ

6.สรุปผลและข้อเสนอแนะ

6.1 สรุปผล

ผลการวิจัย ครั้งนี้โดยพัฒนาและเปรียบเทียบตัวแบบการจำแนกทั้ง 2 เทคนิค ได้แก่ JRIP และต้นไม้ตัดสินใจซึ่งผลการประเมินประสิทธิภาพตัวแบบที่มีค่าความแม่นยำมากที่สุด คือ ต้นไม้ตัดสินใจซึ่งได้ค่าที่สูงที่สุดโดยวัดจากค่าความถูกต้องในการจำแนกข้อมูล (Accuracy)

6.2 ข้อเสนอแนะ

การวิจัยครั้งนี้เป็นการวิจัยที่ใช้โปรแกรม Weka 3.8.4 เพื่อนำมาประยุกต์ใช้ในการจำแนกข้อมูล (Performance Classification) โดยที่มีเครื่องมือต่าง ๆ ให้เลือกใช้ อาทิ เช่น IBk,J48,JRipฯ โดยสามารถนำเครื่องมือเหล่านี้มาเปรียบเทียบการจำแนกเพื่อหาค่าความถูกต้องมากที่สุด

8.เอกสารอ้างอิง

- จิราภา เลหาหะวรรณันท์, รชต ลิ้มสุทธิวันภูมิ และบัณฑิต ฐานะโสภณ. (2556). การใช้เทคนิคการทำเหมืองข้อมูลในการจำแนกและคัดเลือกแขนงวิชาสำหรับนักศึกษาคณะเทคโนโลยีสารสนเทศ. คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, กรุงเทพฯ
- นิตินันท์ มาตา, พนิดา หล่อวงศ์ตระกูล และพยุ่ง มีสีจ. (2558). การเปรียบเทียบประสิทธิภาพเทคนิคการลมิติของข้อมูลสำหรับค้นหาปัจจัย และสร้างโมเดลการจำแนกกลุ่มการระบายน้ำของประตुरะบายน้ำ.
- พัฒนพงษ์ ดลรัตน์ และ จาริ ทองคำ. (2560). การเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพงาน. สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม, มหาสารคาม, 44150.