

# การพัฒนาตัวแบบทำนายการเป็นโรคติดอินเทอร์เน็ตด้วยเทคนิคต้นไม้เพื่อการตัดสินใจ

## The Development Model for Prediction of Internet Addiction Disorder Using Decision Tree Algorithms

จิราพร สุวรรณประภา และภรณ์ยา ปาลวิสุทธิ

สาขาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม  
paranya@npru.ac.th

### บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อจำแนกกลุ่มนักศึกษาที่มีความเสี่ยงต่อการเป็นโรคติดอินเทอร์เน็ตด้วยเทคนิคต้นไม้เพื่อการตัดสินใจ โดยใช้ข้อมูลจากนักศึกษาในมหาวิทยาลัยราชภัฏนครปฐมเป็นกรณีศึกษา ผลการทดลองพบว่า ตัวแบบทำนายการเป็นโรคติดอินเทอร์เน็ตด้วยเทคนิคต้นไม้เพื่อการตัดสินใจแบบ J48 มีประสิทธิภาพในการจำแนกข้อมูลดีกว่า แบบ JRip และแบบ REP ซึ่งมีค่าความถูกต้องในการจำแนก คือ 75.60 %

**คำสำคัญ:** เหมืองข้อมูล, ต้นไม้เพื่อการตัดสินใจ, โรคติดอินเทอร์เน็ต

### Abstract

*This research aims to development model for prediction of Internet Addiction Disorder using Decision Tree Algorithms. Using data of the students in Nakhon Pathom Rajabhat University. The experimental results showed that percentile of accuracy rate of Decision Tree J48 high more than JRip and REP with accuracy 75.60 %*

**Keywords:** data mining, decision trees, internet addiction disorder

### 1. บทนำ

ปัจจุบันความก้าวหน้าทางเทคโนโลยีการสื่อสารก่อให้เกิดกระแสโลกาภิวัตน์ การใช้อินเทอร์เน็ตถึงเป็นส่วนหนึ่งในการดำเนินชีวิตของคนในยุคนี้ อินเทอร์เน็ตมีอิทธิพลต่ออาชีพของคนในปัจจุบันเป็นอย่างมากไม่ว่าจะเป็นอาชีพไหนก็ต้องใช้อินเทอร์เน็ตในการสื่อสาร แล้วยังส่งผลต่อบุคลากรทุกเพศทุกวัย โดยเฉพาะอาชีพที่ต้องทำงานที่ต้องใช้อินเทอร์เน็ตในการสื่อสาร ติดต่อกับ เช่น นักธุรกิจ นักบิน อาจารย์ นักเล่นหุ้น เป็นต้น การเรียนในสมัยนี้ก็จะแตกต่างกันกับสมัยก่อนเป็นอย่างมาก เพราะการเรียนสมัยนี้จะเน้นการเรียนการสอนที่ต้องผ่านอินเทอร์เน็ต การค้นคว้าหาข้อมูล การส่งงานใน E-mail ก็ต้องผ่านทางอินเทอร์เน็ต การสื่อสารโดยการใช้สื่อออนไลน์ผ่านทาง Facebook , Line , Instagram , Twitter และช่องทางอื่นอีกมากมาย ล้วนแล้วแต่ต้องใช้อินเทอร์เน็ตในการสื่อสารทั้งหมด เนื่องจากการเล่นอินเทอร์เน็ตไม่ได้กำหนดอายุและเพศของผู้เล่น ไม่ว่าจะอายุน้อยเพียงใดก็สามารถเล่นได้โดยไม่มีข้อกีดกัน ซึ่งการเล่นอินเทอร์เน็ตก็จะให้ทั้งที่เป็นคุณและเป็นโทษต่อเยาวชนและบุคลากรในสายอาชีพอื่นๆ

ลักษณะ พลอยเลื่อมแสง (2545) ได้กล่าว การดูสื่อผ่านทางอินเทอร์เน็ตก็ให้ทั้งผลดีและผลเสียเช่นเดียวกัน การดูสื่อผ่านทางอินเทอร์เน็ตนั้นมีข้อแตกต่างที่ดึงดูดมากกว่าสื่ออื่น ๆ ที่อยู่บนโลกอินเทอร์เน็ต คือ ความสามารถในการโต้ตอบ (Interact) กับผู้ที่ติดต่อสื่อสารได้โดยทันที (Real Time) รวมถึงภาพและเสียง ซึ่งทำให้อินเทอร์เน็ตมีความเป็นโลกเสมือนที่มี

ตัวตน ทำให้เกิดสังคมของโลกใหม่ ที่ทำให้รู้สึกว่าคุณภาพความเป็นจริงกับโลกเสมือนคือโลกเดียวกัน อินเทอร์เน็ตมีบทบาทต่อการดำเนินชีวิต การเรียนรู้ ตลอดจนการมีปฏิสัมพันธ์ระหว่างบุคคล

จากการแพร่หลายของการใช้อินเทอร์เน็ตได้พบว่าบางคนเกิดความรู้สึกยึดติดกับการใช้อินเทอร์เน็ต แต่เมื่อใช้ในปริมาณมากเกินไปอาจทำให้เกิดความเสียหายกับชีวิตและไม่สามารถควบคุมได้ เรียกพฤติกรรมการใช้ลักษณะนี้ว่าโรคติดอินเทอร์เน็ต (Internet Addiction Disorder : IAD) ซึ่งในปัจจุบันพบมากในกลุ่มเยาวชน ดังนั้นการวิจัยครั้งนี้จึงเป็นการนำเสนอตัวแบบ (Model) การทำนายความเสี่ยงต่อการเป็นโรคติดอินเทอร์เน็ต ซึ่งเทคนิคการทำนายในงานวิจัยนี้ใช้เทคนิคต้นไม้ตัดสินใจ เมื่อทำการวิจัยเสร็จแล้วสามารถนำตัวแบบการทำนายที่ได้มาใช้อธิบายต่อไป

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 ข้อมูลเกี่ยวกับโรคติดอินเทอร์เน็ต

โรคติดอินเทอร์เน็ต(รวิกานต์ นันทเวช,2550) นั้นจัดเป็นกลุ่มอาการทางจิตใจอย่างหนึ่ง ซึ่งเกิดจากการใช้อินเทอร์เน็ตในการเสพข้อมูลหรือข่าวสารมากเกินไป และหากเปรียบเทียบกับอาการติดดูโทรทัศน์หรือการติดอย่างอื่น ๆ แล้ว อินเทอร์เน็ตจะมีข้อแตกต่างกับสื่ออื่น ๆ สามารถโต้ตอบกับผู้ที่เข้ามาใช้คนอื่น ๆ ได้ทันที (Real Time) ซึ่งทำให้โลกของอินเทอร์เน็ตมีความเสมือนเป็นเหมือนโลกอีกโลกหนึ่งที่ใช้ใช้สามารถ มีตัวตนในโลกนั้นได้ โดยปราศจากกฎเกณฑ์ และไร้ขอบเขตในการเดินทาง และสร้างตัวตนในโลกอินเทอร์เน็ตตามที่ตัวเองต้องการได้ ซึ่งหากผู้ใช้ยึดติดกับสังคมในโลกของอินเทอร์เน็ต จนแยกไม่ออกว่าโลกของความจริงและโลกเสมือน อาจจะนำมาซึ่งสาเหตุของเป็นโรคติดอินเทอร์เน็ตได้

โรคติดอินเทอร์เน็ต (อีวาน โกลเบิร์ก, 1995) ถูกพูดถึงครั้งแรกในปีค.ศ. 1995 ซึ่งเป็นแพทย์ผู้แนะนำให้กลุ่มอาการติดการพนันเข้าไปไว้เกณฑ์วินิจฉัยโรคทางจิตเวชศาสตร์ของสหรัฐอเมริกา ได้วิเคราะห์ว่า บุคคลใดที่มีอาการดังต่อไปนี้ อย่างน้อย 4 ประการ เป็นเวลาไม่น้อยกว่า 1 ปี แสดงว่าเป็นอาการติดอินเทอร์เน็ต

- รู้สึกหมกมุ่นกับอินเทอร์เน็ต แม้ในเวลาที่ไม่ได้ต่อเข้าระบบอินเทอร์เน็ต
- มีความต้องการใช้อินเทอร์เน็ตเป็นเวลานาน
- ไม่สามารถควบคุมการใช้อินเทอร์เน็ตได้
- รู้สึกหงุดหงิดเมื่อใช้อินเทอร์เน็ตน้อยลง
- คิดว่าเมื่อใช้อินเทอร์เน็ตแล้ว ทำให้ตนเองรู้สึกดีขึ้น
- ใช้เป็นอินเทอร์เน็ตในการหลีกเลี่ยงปัญหา
- หลอกคนในครอบครัว หรือเพื่อน เรื่องการใช้อินเทอร์เน็ตของตนเอง
- มีอาการผิดปกติเมื่อเลิกใช้อินเทอร์เน็ต เช่น หดหู่ กระทบกระชวย

วิธีการป้องกันและรักษาจากโรคติดอินเทอร์เน็ต ผู้เชี่ยวชาญบางคนเชื่อว่าภาวะติดอินเทอร์เน็ตสามารถหายไปเองได้ แต่หลายคนก็จำเป็นต้องได้รับการช่วยเหลือ ดร.อีวาน โกลเบิร์ก จิตแพทย์ ผู้ค้นพบโรคติดอินเทอร์เน็ต ได้ให้คำแนะนำในการป้องกันและช่วยเหลือกับผู้ที่คิดว่าตัวเองเป็นหรือกำลัง จะเป็นโรคนี้ว่า ก่อนอื่นต้องรู้ว่าตัวคุณเองใช้อินเทอร์เน็ตมากเกินไปหรือไม่ ในระดับไหน และรูปแบบลักษณะของการใช้ไปในรูปแบบใดเสียก่อน ร่วมกับวิธีการอื่นเช่นการใช้โปรแกรมเพื่อควบคุมเนื้อหาของอินเทอร์เน็ต การให้คำปรึกษา และพฤติกรรมบำบัด

### 2.2 แนวคิดและทฤษฎีเกี่ยวกับเทคนิคเหมืองข้อมูล

การทำเหมืองข้อมูล (Data mining) (ไพฑูรย์ จันทรเรือง, 2550) คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม

เทคนิคเหมืองข้อมูลที่สำคัญเทคนิคหนึ่งคือ Data Classification เป็นกระบวนการสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ โดยการสร้างกฎเพื่อช่วยในการตัดสินใจจากข้อมูลที่มีอยู่เพื่อใช้ทำนายแนวโน้มการเกิดขึ้นของข้อมูลที่ยังไม่เกิดขึ้น โดยงานวิจัยได้นำเทคนิคการจำแนกประเภทข้อมูลมาใช้ในการวิจัย ดังต่อไปนี้

ต้นไม้ช่วยการตัดสินใจ (Decision tree) เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล โดยต้นไม้ช่วยการตัดสินใจจะมีลักษณะคล้ายโครงสร้างต้นไม้ ที่แต่ละโหนดแสดงคุณลักษณะ (attribute) แต่ละกิ่งแสดงเงื่อนไขในการทดสอบ และโหนดปลาย (leaf node) แสดงกลุ่มที่กำหนดไว้

2.2.1 ขั้นตอนการสร้างโมเดลแผนภาพต้นไม้เพื่อการตัดสินใจ อัลกอริทึม J48 หรือ อัลกอริทึมของ C4.5 ซึ่งเป็นอัลกอริทึมอันหนึ่งที่ใช้สร้าง Decision Tree ที่พัฒนาโดย Ross Quinlan โดย C4.5 คือส่วนขยายที่เพิ่มเติมมาจาก อัลกอริทึม ID3 Decision Tree โดยโครงสร้างนี้สามารถถูกใช้สำหรับการคัดแยกข้อมูล(classification) และ ด้วยเหตุผลอันนี้ C4.5 ถูกเรียกใช้บ่อย ๆ สำหรับตัว Statistical Classifier ในส่วนของอัลกอริทึม C4.5 สร้าง Decision Trees มาจากกลุ่มของ training data เหมือนกับ ID3 ที่ใช้หลักการของ Information Entropy ใช้ความถูกต้องของแต่ละแอตทริบิวต์ของข้อมูล เพื่อใช้เป็นการตัดสินใจแบ่งกลุ่มข้อมูลไปยังกลุ่มย่อย ๆ ซึ่ง C4.5 จะพิจารณา ตรวจสอบ Normalized Information Gain (ความแตกต่างใน Entropy) ผลลัพธ์จากการเลือกแอตทริบิวต์สำหรับการแบ่งกลุ่มข้อมูลโดยคุณลักษณะด้วยค่า Normalized Information Gain ที่สูงที่สุดนั้นคือหนึ่งในการสร้างการตัดสินใจ

2.2.2 กฎของริปเปอร์ ถูกสร้างโดย Cohen ในปี 1995 โดย ริปเปอร์ถูกประกอบด้วย 2 เฟส คือ เฟสแรกทำการระบุกฎเริ่มต้น และเฟสที่สองจะระบุค่า Post-Process Rule Optimization โดยข้อมูลที่ ถูกการเรียนรู้ (Training Data) จะถูกแบ่งไปเป็น Growing Set และ Pruning Set ดังนั้น อัลกอริทึมนี้จะสร้างกฎความสัมพันธ์ใน Greedy Fashion ในขณะที่สร้างกฎริปเปอร์นั้น จะหาค่าที่ดีที่สุดสำหรับ Growing Set ใน Rule Space ซึ่ง จะอธิบายได้จาก BNF หลังจากได้ Growing Set ก็จะทำ Pruning ข้อมูลทันทีเมื่อเสร็จแล้วจะได้กลุ่มตัวอย่างที่ เหมือนกันออกมาที่ครอบคลุมกฎของ Training Set จากนั้นก็ จะถูกลบทิ้งซึ่ง Training Data ที่เหลือจะถูกแบ่งใหม่อีกครั้งหนึ่ง หลังจากถูกเรียนรู้ตามกฎแล้ว เพื่อช่วยแก้ปัญหาที่เกิดขึ้นจากการแบ่งแยกกลุ่มที่ผิดพลาด ซึ่งกระบวนการนี้จะกระทำจนกระทั่งผลเป็นที่น่าพอใจ

### 2.3 งานวิจัยที่เกี่ยวข้อง

ผู้อำนวยการศูนย์ศึกษาการติดคอมพิวเตอร์ของโรงพยาบาลแมคลินในมหาวิทยาลัยฮาร์เวิร์ด (รวิกรานต์ นันทเวช, 2550) พบว่าผู้ใช้บริการท่องอินเทอร์เน็ตร้อยละ 5-10 มีปัญหาต้องพึ่งพาอินเทอร์เน็ตจนขาดไม่ได้ และร้อยละ 6 ของคนที่ใช้ยอมรับว่าอินเทอร์เน็ตมีผลกระทบต่อชีวิตของตนเอง ดร. เดวิด กรีนฟิลด์ แห่งศูนย์พฤติกรรมอินเทอร์เน็ต เชื่อว่าบริการบางอย่างในอินเทอร์เน็ตส่งผลกระทบต่อทางจิตวิทยาทำให้เกิดการบิดเบือนของเวลา สร้างความพึงพอใจ กรีนฟิลด์อ้างว่าไม่ว่าจะเป็นเรื่องเพศ การพนัน ซื้อสินค้าออนไลน์ล้วนมีผลกระทบต่อทางอารมณ์ทั้งสิ้น อย่างไรก็ตามเขากล่าวว่าหากคนทั่วไปออนไลน์แล้วเกิดปัญหา ไม่ควรเรียกภาวะนี้ว่าเสพติด แต่ควรเรียกว่าภาวะที่ควบคุมตัวเองไม่ได้ เช่นภาวะควบคุมตนเองไม่ได้ในการพนันหรือซื้อของ (Compulsive Gambling Or Shopping) ซึ่งเราอาจเรียกว่าภาวะควบคุมตนเองไม่ได้ในการใช้อินเทอร์เน็ต

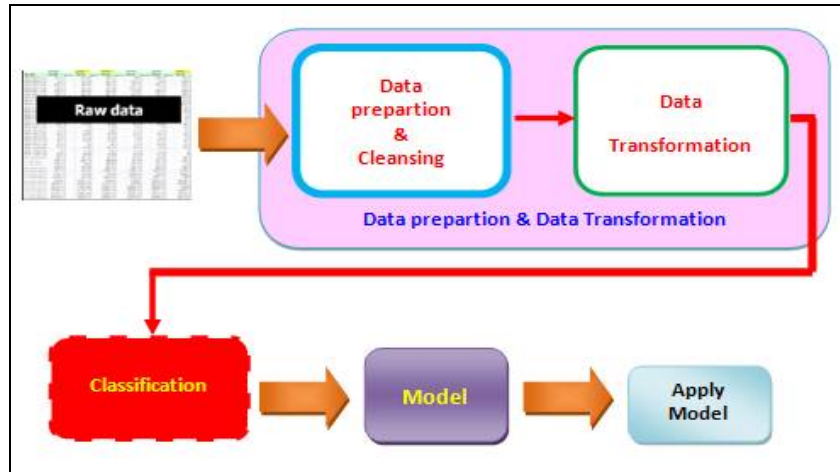
## 3. วิธีการดำเนินการวิจัย

### 3.1 การเตรียมข้อมูลสำหรับทำเหมืองข้อมูล

โรคติดอินเทอร์เน็ต นั้นจัดเป็นกลุ่มอาการทางจิตใจอย่างหนึ่ง ซึ่งเกิดจากการใช้อินเทอร์เน็ตในการเสพข้อมูลหรือข่าวสารมากเกินไป และหากเปรียบเทียบกับการติดดูโทรทัศน์หรือการติดอย่างอื่น ๆ แล้วอินเทอร์เน็ตจะมีข้อแตกต่างกับสื่ออื่นๆ ก็คือ ความสามารถโต้ตอบกับผู้ที่เข้ามาใช้คนอื่นๆ ได้ทันที ซึ่งทำให้โลกของอินเทอร์เน็ตมีความเสมือนเป็นเหมือนโลกอีกโลกหนึ่งที่ผู้ใช้สามารถ มีตัวตนในโลกนั้นได้ โดยปราศจากกฎเกณฑ์ และไร้ขอบเขตในการเดินทาง

เพื่อนำมาเป็นข้อมูลในการวิเคราะห์และออกแบบพัฒนาในขั้นต่อไปผู้พัฒนาจึงได้ทำการวิเคราะห์รูปแบบข้อมูลโดยใช้อัลกอริทึมของต้นไม้เพื่อการตัดสินใจ (Building a Decision Tree) แบบ J48, การสร้างโมเดลกฎริปเปอร์แบ่งแยกข้อมูล (Building Ripper Rule) แบบ JRip และการตัดกิ่งแบบความผิดพลาดลดลง (Reduced-Error Pruning) (REP) มาทำการเปรียบเทียบหาประสิทธิภาพเพื่อความแม่นยำในการทำนายค่าโดยวิเคราะห์ลักษณะของชุดข้อมูล (Datasets)

การออกแบบระบบและการพัฒนาโมเดลจนได้ข้อสรุปในออกแบบและพัฒนาซึ่งเป็นการนำโปรแกรม WEKA เข้ามาจัดสร้างตัวแบบเพื่อใช้ในการหาค่าความเสี่ยงต่อการเกิดโรคติดอินเทอร์เน็ต โดยนำกลุ่มตัวอย่างที่ได้ทำแบบประเมินมาศึกษา กลุ่มตัวอย่างที่ได้ไปศึกษานั้นเป็นกลุ่มเยาวชนและนักศึกษาในมหาวิทยาลัยราชภัฏนครปฐม ได้กรอบแนวคิดของงานวิจัยเพื่อให้ได้โมเดลในการทำนายความเสี่ยงต่อการเกิดโรคติดอินเทอร์เน็ต



ภาพที่ 1: กรอบแนวคิดของงานวิจัย

จากกรอบแนวคิดมีรายละเอียดขั้นตอนดังต่อไปนี้

### ขั้นตอนที่ 1 Data preparation and Cleansing

ทำการจัดเก็บข้อมูล ซึ่งได้จากแบบประเมินอาการติดเชื้อในท่อน้ำเหลือง ในกลุ่มเยาวชนและในมหาวิทยาลัยราชภัฏนครปฐม นำมาวิเคราะห์ให้อยู่ในรูปแบบของ Relation Data Model และทำการเก็บข้อมูลลงใน Microsoft Excel

### ขั้นตอนที่ 2 Data Transformation

โปรแกรมที่ใช้ในงานวิจัยครั้งนี้ ผู้จัดทำได้เลือกใช้โปรแกรม Weka เวอร์ชัน 3.6.13 ซึ่งเป็นซอฟต์แวร์ด้านการทำเหมืองข้อมูลที่ได้รับการยอมรับอย่างแพร่หลายซึ่งต้องจัดเตรียมแฟ้มงาน (File) ให้อยู่ในรูปแบบ ARFF (ARFF Format)

### ขั้นตอนที่ 3 Classification

การจำแนกประเภทข้อมูล เป็นวิธีการหนึ่งในกระบวนการทำเหมืองข้อมูล ซึ่งเป็นเทคนิคในการจำแนกข้อมูล ตามที่กำหนดด้วย ต้นไม้ช่วยการตัดสินใจ เป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากเทคนิคการจำแนกประเภทข้อมูล โดยต้นไม้ช่วยการตัดสินใจจะมีลักษณะคล้ายโครงสร้างต้นไม้ ที่แต่ละโหนดแสดงคุณลักษณะ แต่ละกิ่งแสดงเงื่อนไขในการทดสอบ และโหนดปลาย แสดงกลุ่มที่กำหนดไว้ โดยจะเลือกเทคนิค ดังต่อไปนี้

1. ต้นไม้เพื่อการตัดสินใจ ด้วยเทคนิค J48
2. กฎริบเปอร์แบ่งแยกข้อมูล ด้วยเทคนิค JRip
3. การตัดกิ่งแบบความผิดพลาดลดลงด้วยเทคนิค REP

### 3.2 การเตรียมข้อมูลสำหรับทำเหมืองข้อมูล

ขั้นตอนการเตรียมข้อมูลก่อนที่จะส่งไปเข้าสู่กระบวนการหาอัลกอริธึมและวิเคราะห์ ซึ่งขั้นตอนนี้สำคัญมากเนื่องจากเป็นขั้นตอนการเตรียมข้อมูลก่อนที่จะส่งไปเข้าสู่กระบวนการหาอัลกอริธึมและวิเคราะห์ และนำ Data มาทำการวิเคราะห์ทำการเตรียมข้อมูลโดยมีรายละเอียดข้อมูล คือ จำนวนข้อมูลทั้งหมด 880 ชุด จำนวน แอตทริบิวต์ทั้งหมด 16 แอตทริบิวต์ มี 3 คลาส คือ

- คลาส 1 คือ ติดโรคติดเชื้อในท่อน้ำเหลือง
- คลาส 2 คือ เสี่ยงต่อการติดเชื้อในท่อน้ำเหลือง
- คลาส 3 คือ ไม่ติดเชื้อในท่อน้ำเหลือง

ซึ่งข้อมูลดังตารางที่ 1

ตารางที่ 1 รายละเอียดข้อมูล

ลำดับ	ชื่อแอตทริบิวต์	คำอธิบาย
1	Sex	เพศ 1 = เพศชาย 2 = เพศหญิง
2	Record	ผลการเรียน 1 = ต่ำกว่า 2.00 2= 2.49 – 2.00 3= 2.99 – 2.50 4= 3.00 – 3.49 5= 3.50 ขึ้นไป
3	status	สถานภาพสมรสของบิดามารดา 1 = อยู่ด้วยกัน 2 = แยกกันอยู่ 3 = หย่าร้าง 4= บิดา/มารดาเสียชีวิต
4	Siblings	จำนวนพี่น้องร่วมบิดามารดา 1 = เป็นบุตรคนเดียว 2 = 2 คน 3 = 3 คน 4 = 4 คนขึ้นไป
5	income	รายได้เฉลี่ยต่อเดือน 1 = น้อยกว่า 5,000 2 = 5,000 – 8,000 3 = 8,001 – 10,000 4 = มากกว่า 10,000
6	null_father	อาชีพของบิดา 1 = รับราชการ/รัฐวิสาหกิจ 2 = เกษตรกร 3 = ค้าขาย/ธุรกิจส่วนตัว 4 = รับจ้างทั่วไป 5= อื่นๆ
7	null_mom	อาชีพของมารดา 1= รับราชการ/รัฐวิสาหกิจ 2 = เกษตรกร 3 = ค้าขาย/ธุรกิจส่วนตัว 4 = รับจ้างทั่วไป 5= อื่นๆ
8	null_dwll	อาศัยอยู่กับใคร 1 = บิดาและมารดา 2 = บิดาคนเดียว 3 = มารดาคนเดียว 4 = ญาติ

ตารางที่ 1 รายละเอียดข้อมูล (ต่อ)

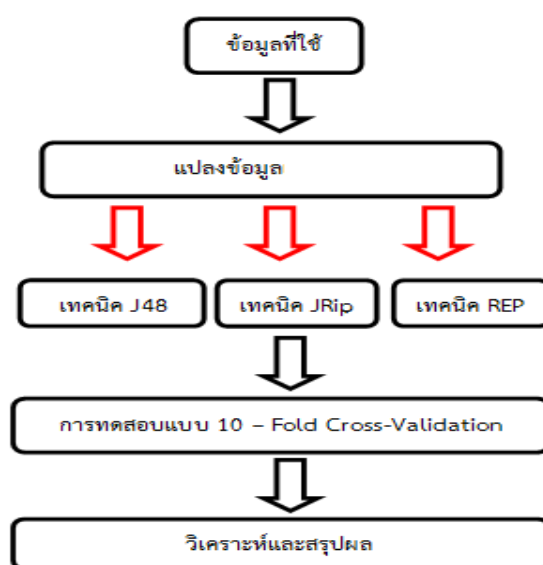
ลำดับ	ชื่อแอตทริบิวต์	คำอธิบาย
9	residence	ที่พักอาศัย 1 = บ้านส่วนตัว 2 = บ้านเช่า 3 = หอพัก/คอนโด
10	null_cafe	ในชุมชนที่อยู่มีร้านอินเทอร์เน็ตมาก น้อยเพียงใด 1 = ไม่มีเลย 2 = 1 ร้าน 3 = 4 - 6 ร้าน 4 = 7 - 9 ร้าน 5 = มากกว่า 10 ร้าน
11	null_period	ระยะเวลาที่เริ่มใช้อินเทอร์เน็ต 1 = ต่ำกว่า 6 เดือน 2 = 6 เดือน - 1 ปี 3 = 1 ปี - 2 ปี 4 = 2 ปีขึ้นไป
12	activities	กิจกรรมยามว่าง 1 = ช่วยงานทางบ้าน 2 = รับจ้างหารายได้พิเศษ 3 = อ่านหนังสือทบทวนบทเรียน 4 = เล่นกีฬาออกกำลังกาย 5 = เล่นเกมอินเทอร์เน็ต 6 = ไปเที่ยวสังสรรค์กับเพื่อน ๆ
13	null_palace	ใช้อินเทอร์เน็ตที่ไหนบ่อยที่สุด 1 = ที่พักอาศัย 2 = สถานศึกษา 3 = โทรศัพท์/อุปกรณ์เคลื่อนที่ 4 = อินเทอร์เน็ตคาเฟ่ 5 = อื่นๆ
14	smoking	การสูบบุหรี่ 1 = ไม่สูบ 2 = บางครั้ง 3 = หลายครั้งต่อสัปดาห์ 4 = ทุกวัน
15	alcohol	ดื่มแอลกอฮอล์ ในระยะเวลา 3 เดือน ที่ผ่านมา 1 = ไม่เคย 2 = 1 - 5 ครั้ง 3 = 6 - 10 ครั้ง 4 = มากกว่า 11 ครั้ง

ตารางที่ 1 รายละเอียดข้อมูล (ต่อ)

ลำดับ	ชื่อแอตทริบิวต์	คำอธิบาย
16	result	1 = IAD 2 = RICK-IAD 3 = NO-RICK

3.3 การวิเคราะห์และการประเมินผล

- ส่วนแรก แผนภาพต้นไม้เพื่อการตัดสินใจ เทคนิค J48
- ส่วนสอง กฎริบเปอร์แบ่งแยกข้อมูล เทคนิค JRip
- ส่วนสาม การตัดกิ่งแบบความผิดพลาดลดลง เทคนิค REP



ภาพที่ 2 ขั้นตอนจำลองแบบจำลอง

4. ผลการทดสอบ

ผลการทดสอบข้อมูลด้วยอัลกอริธึม J48 เป็นการสร้างโมเดลแผนภาพต้นไม้เพื่อการตัดสินใจ ด้วยเทคนิค J48 มาทำการทดสอบจากผลการทดลอง พบว่า ตัวแบบทำนายการสร้างโมเดลแผนภาพต้นไม้เพื่อการตัดสินใจ ด้วยเทคนิค J48 มีความถูกต้องในการจำแนกข้อมูล 75.60 % ค่าความผิดพลาดในการจำแนกข้อมูล 24.65% โดยมีค่าความแม่นยำ 0.40 ค่าความระสีก 0.22 และค่าความถ่วงดุล 0.40

ผลการทดสอบข้อมูลด้วยอัลกอริธึม JRip จะเป็นการสร้างโมเดลกฎริบเปอร์แบ่งแยกข้อมูล ด้วยเทคนิค JRip มาทำการทดสอบ จากผลการทดลอง พบว่า ตัวแบบทำนายการสร้างโมเดลกฎริบเปอร์แบ่งแยกข้อมูลด้วยเทคนิค JRip มีความถูกต้องในการจำแนกข้อมูล 71.14 % ค่าความผิดพลาดในการจำแนกข้อมูล 28.90 % โดยมีค่าความแม่นยำ 0.21 ค่าความระสีก 0.30 และค่าความถ่วงดุล 0.40

ผลการทดสอบข้อมูลด้วยอัลกอริธึม REP จะเป็นการนำการตัดกิ่งแบบความผิดพลาดลดลงด้วยเทคนิค REPมาทำการทดสอบ จากผลการทดลอง พบว่า ตัวแบบทำนายการตัดกิ่งแบบความผิดพลาดลดลง (Reduced-error pruning) ด้วยเทคนิค REPมีความถูกต้องในการจำแนกข้อมูล 70.80 % ค่าความผิดพลาดในการจำแนกข้อมูล 29.20 % โดยมีค่าความแม่นยำ 0.21 ค่าความระสีก 0.24 และค่าความถ่วงดุล 0.40

## ตารางที่ 2 ค่าความถูกต้องของเทคนิคทั้ง 3 อัลกอริธึม

Type	Correctly	Incorrect
เทคนิค J48	75.60	24.65
เทคนิค JRip	71.14	28.90
เทคนิค REP	70.80	29.20

### 5. บทสรุป

จากผลการวิจัยนี้สรุปได้ว่าเทคนิคในการทำเหมืองข้อมูลที่เหมาะสมที่สุดที่ควรใช้กับข้อมูลความเสี่ยงต่อการเป็นโรคติดเชื้ออินเทอร์เน็ต เพื่อให้ได้ตัวแบบที่จะนำมาพัฒนาระบบทำนายนักศึกษาในมหาวิทยาลัยราชภัฏนครปฐมและเยาวชนที่เสี่ยงต่อการเกิดโรคอินเทอร์เน็ตซึ่งข้อมูลที่นำมาใช้ในการวิจัยประกอบด้วยคุณลักษณะทั้งหมด 16 แอตทริบิวต์ มีจำนวนข้อมูลทั้งหมด 880 ชุดข้อมูล โดยเทคนิคของการจำแนกข้อมูลที่เลือกใช้ คือ ต้นไม้ช่วยการตัดสินใจ เทคนิคที่ใช้ คือ J48, REP Tree และ Jrip โดยนำทั้งสามเทคนิคนี้มาเปรียบเทียบค่าประสิทธิภาพในการจำแนกที่เหมาะสมที่สุด ซึ่งจากการทดลองสรุปได้ว่าตัวแบบทำนายด้วยเทคนิค J48 จะเป็นอัลกอริธึมที่เหมาะสมที่ให้ค่าความถูกต้องในการจำแนกข้อมูลสูงกว่าเทคนิค Jrip และ REP จึงเหมาะกับการนำมาพัฒนาระบบทำนายความเสี่ยงต่อการเป็นโรคติดเชื้ออินเทอร์เน็ต

### 6. เอกสารอ้างอิง

- กฤษณะ ไวยมัย, ชิตชนก ส่งศิริ, และธนาวินท์ รักธรรมานนท์, “เทคนิคการทำเหมืองข้อมูล,” .2001
- จิโรจน์ ภาคศิริ และ กาญจนา วิริยะพันธ์. การวิเคราะห์รูปแบบวิเคราะห์การบุกรุกข้อมูลบนเครือข่ายโดยใช้เทคนิคดาต้าไมนิง. วารสารเทคโนโลยีสารสนเทศ ปีที่ : 3 ฉบับที่ : 6 เลขหน้า : 40-46, 2550
- บุญเสริม กิจศิริกุล. อัลกอริธึมการทำเหมืองข้อมูล. รายงานวิจัยฉบับสมบูรณ์ โครงการวิจัยร่วมภาครัฐและเอกชน ปีงบประมาณ 2545. ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมคอมพิวเตอร์จุฬาลงกรณ์ มหาวิทยาลัย, 2546.
- ไพฑูริย์ จันทร์เรือง, “ระบบสนับสนุนการตัดสินใจเลือกสาขาการเรียนของนักศึกษาปริญญาตรีโดยใช้เทคนิคต้นไม้ตัดสินใจ,” ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ, เทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2550.
- รวีกรานต์ นันทเวช. (2550). ปัจจัยที่ส่งผลต่อพฤติกรรมการติดเชื้ออินเทอร์เน็ตของนักเรียนชั้นมัธยมศึกษาตอนปลาย วิทยานิพนธ์ครุศาสตรมหาบัณฑิตจุฬาลงกรณ์มหาวิทยาลัย
- ลักษณา พลอยล้อมแสง. (2545). ผลกระทบของอินเทอร์เน็ตต่อ สุขภาพจิตของผู้ใช้ในเขตอำเภอเมือง จังหวัดเชียงใหม่. วารสารสวนปรุง, 18 (2), 34-46.