

## การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูลโรคออทิสติกด้วยเทคนิคเหมืองข้อมูล A Comparison of the Efficiency of the Classification of Autism Disease Using Data Mining Technique

จักรกฤษณ์ หงส์เวียงจันทร์<sup>1</sup> นิติมา ลักขณานุรักษ์<sup>2</sup> และไกรุ่ง เสงพะพรหม<sup>1\*</sup>

<sup>1,3</sup>สาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม

<sup>2</sup>สาขาวิชาเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร

\*kairung2011.heng@gmail.com

### บทคัดย่อ

การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบเทคนิคการจำแนกกลุ่มข้อมูลที่เหมาะสมกับการจำแนกข้อมูลโรคออทิสติก ในการวิจัยครั้งนี้ได้ทำการศึกษาคัดเลือกเทคนิคสำหรับใช้ในการวิจัยนี้ทั้งหมด 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และเทคนิคเพื่อนบ้านใกล้เคียงที่สุด เพื่อหาประสิทธิภาพของการจำแนกข้อมูลที่ดีที่สุดในกลุ่มข้อมูลโรคออทิสติก ผลการวิจัยพบว่า เทคนิคที่ให้ค่าความแม่นยำมากที่สุด ซึ่งให้ประสิทธิภาพที่ดีที่สุดคือ เทคนิคโครงข่ายประสาทเทียม โดยให้ค่าความแม่นยำอยู่ที่ 99.65 %

**คำสำคัญ:** การจำแนกข้อมูล ต้นไม้ตัดสินใจ เทคนิคเพื่อนบ้านใกล้เคียงที่สุด โครงข่ายประสาทเทียม

### Abstract

This research presents the results of comparing the efficiency of the classification of autism disease. By comparing the efficiency of data classification of 3 techniques, namely, the Decision Tree technique, K-NN technique and Artificial neural network by using information about the autism disease. The results showed that techniques used in the classification of autism disease with the highest efficiency is Artificial neural network with 99.65 % accuracy which is an acceptable rating level.

**Keywords:** data classification, decision tree, k-NN, artificial neural network

## 1. บทนำ

กลุ่มอาการของโรคออทิสติกจัดเป็นความผิดปกติของพัฒนาการรูปแบบหนึ่ง ซึ่งมีผลต่อการปฏิสัมพันธ์ทางสังคม การสื่อสาร ความสนใจ และพฤติกรรม ในกลุ่มเด็กที่เป็นโรคออทิสติก อาการต่าง ๆ มักจะเริ่มเกิดขึ้นก่อนที่เด็กจะมีอายุครบสามปี แต่ยังมีบางกรณีที่อาจมีการวินิจฉัยว่าเด็กเป็นโรคออทิสติก ภายหลังจากเด็กคนนี้อายุเกินสามปีไปแล้วก็ตาม

อาการและอาการแสดง สำหรับคนที่เป็นโรคออทิสติกมีแนวโน้มที่จะมีปัญหาเกี่ยวกับปฏิสัมพันธ์ทางสังคม และการสื่อสารกับผู้อื่น ในช่วงเป็นเด็กทารกจะพบว่าเด็กบางคนที่มีอาการของโรคออทิสติก ไม่เคยทำเสียงอ้อแอ้ หรือ ฝีกทำเสียง เหมือนกับเด็กทารกคนอื่น เมื่อโตเข้าสู่ช่วงวัยเด็กก็จะมีปัญหาในเรื่องปฏิสัมพันธ์ที่ไม่ใช่คำพูด หรือ เรื่องภาษากายในการตอบโต้กับคนอื่น ๆ เช่น เด็กที่เป็นโรคออทิสติกมักมีปัญหาเรื่องการสบตา การแสดงอาการออกทางสีหน้า ภาษากาย และท่าทางที่ผิดแปลกจากเด็กทั่วไป เด็กเหล่านี้มักไม่สบตา หรือสบตาเป็นระยะสั้น ๆ และมักเพิกเฉยกับคนที่รู้จักหรือไม่รู้จักก็ตาม เด็กที่มีอาการเป็นโรคออทิสติก อาจไม่มีการรับรู้ถึงตัวตนเด็กรอบข้าง และมักไม่สนใจเด็กคนอื่นบ่อยครั้งที่เด็กเหล่านี้มักจะเข้าหาเด็กที่มีอายุมากกว่า หรือ อายุน้อยกว่าตนเอง มากกว่าที่จะปฏิสัมพันธ์กับเด็กวัยเท่า ๆ กัน

ดังนั้นในงานวิจัยนี้จึงเห็นถึงความสำคัญของการจำแนกกลุ่มเด็กที่มีอาการของโรคออทิสติก เพื่อสำหรับเป็นแนวทางในการสนับสนุนการตัดสินใจทางการแพทย์เกี่ยวกับการวินิจฉัยโรคออทิสติก โดยในงานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบเทคนิคการจำแนกกลุ่มข้อมูลที่เหมาะสมกับการจำแนกข้อมูลของโรคออทิสติก โดยทำการศึกษากับ 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้เคียงที่สุด เพื่อหาเทคนิคที่ดีที่สุด สำหรับการจำแนกกลุ่มอาการของโรคออทิสติก และนำเสนอเทคนิคที่ดีที่สุดสำหรับการจำแนกกลุ่มข้อมูลโรคออทิสติก

## 2. วัตถุประสงค์

เพื่อเปรียบเทียบประสิทธิภาพเทคนิคการจำแนกข้อมูลสำหรับโรคออทิสติก ทั้งหมด 3 วิธี ได้แก่ ต้นไม้ตัดสินใจ เครือข่ายประสาทเทียม และ เทคนิคเพื่อนบ้านใกล้เคียงที่สุด

## 3. งานวิจัยที่เกี่ยวข้อง

### 3.1 ทฤษฎีที่เกี่ยวข้อง

#### เทคนิคต้นไม้ช่วยตัดสินใจ (Decision Tree)

เทคนิคต้นไม้ช่วยตัดสินใจ (เดซ ธรรมศิริ, 2554) เป็นเทคนิคที่ได้ผลลัพธ์เป็นแบบจำลองการพยากรณ์ของข้อมูลในรูปแบบของต้นไม้ช่วยตัดสินใจ (Decision Tree) เทคนิคนี้ถือได้ว่าเป็นวิธีการจัดหมวดหมู่หรือแบ่งประเภท (Classification) มีการทำงานเป็นแบบการเรียนรู้โดยมีผู้สอน (Supervised Learning) ซึ่งเป็นเทคนิคของการทำเหมืองข้อมูล (Data Mining) อย่างหนึ่งที่เป็นที่นิยมใช้งานกันเนื่องจากเป็นวิธีที่ง่ายไม่ซับซ้อน

#### โครงข่ายประสาทเทียม (Artificial Neural Networks - ANN)

โครงข่ายประสาทเทียม (วิทยา พรพัชรพงศ์, 2562) เป็นเทคนิคทางเหมืองข้อมูลที่ให้คอมพิวเตอร์จำลองตัวเองให้ทำงานเลียนแบบการทำงานของเซลล์ประสาทในสมองของมนุษย์ ที่สามารถฝึกฝน เรียนรู้ ตลอดถึงการนำความรู้หรือทักษะไปประยุกต์ใช้งานได้จริง โดยมีการทำงานในรูปแบบของการประมวลผลพร้อม ๆ กันเป็นจำนวนมาก ซึ่งเทคนิคนี้เป็นวิทยาการแขนงหนึ่งทางปัญญาประดิษฐ์ (Artificial Intelligence)

ในนิวรอล (Neurons) จะประกอบด้วยข้อมูลเข้า (Inputs) และข้อมูลออก (Outputs) แต่ละข้อมูลเข้าจะมีการจำลองค่าน้ำหนัก (Weight) ในนิวรอลมีค่าเทรชโฮลด์ (Threshold) ซึ่งเป็นตัวกำหนดค่าน้ำหนักรวมของข้อมูลเข้า ว่าต้องมีค่ามากเท่าไรถึงจะสามารถส่งข้อมูลออกไปยังนิวรอลอื่นได้

เทคนิคเพื่อนบ้านใกล้เคียงที่สุด

### เทคนิคเพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbor: KNN)

เทคนิคเพื่อนบ้านที่ใกล้ที่สุด (นพมาศ ปักเข็มและคณะ, 2560) เป็นเทคนิคที่ใช้สำหรับการจัดกลุ่มของข้อมูลโดยคำนวณจากระยะห่างของแต่ละคุณลักษณะในข้อมูล (Data) ซึ่งวิธีนี้จะเหมาะสมกับข้อมูลที่เป็นเชิงตัวเลข การจัดข้อมูลที่อยู่ใกล้กันให้เป็นกลุ่มเดียวกันจะตรวจสอบจากเงื่อนไขของข้อมูลจะตรวจสอบจากจำนวน K ที่กำหนดไว้ แต่เทคนิคนี้จะใช้ระยะเวลาในการประมวลผลที่นาน ถ้าข้อมูลมีปริมาณมากอาจเกิดปัญหาในการคำนวณและใช้ปริมาณทรัพยากรในการประมวลผลสูงมากเนื่องจากจะใช้เวลาสำหรับการประมวลผลเพิ่มขึ้นตามจำนวนข้อมูลที่เพิ่มขึ้นทั้งหมด

### 3.2 งานวิจัยที่เกี่ยวข้อง

สุรวัชร ศรีเปารยะ และสายชล สินสมบุญทอง (2560) ศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศไทย ผลการวิจัยพบว่า วิธีต้นไม้ตัดสินใจมีประสิทธิภาพในการจำแนกของโรคไตเรื้อรังมากที่สุดดังนั้นผู้วิจัยจึงเลือกวิธีต้นไม้ตัดสินใจสำหรับการจำแนกกลุ่มผู้ป่วยโรคไตเรื้อรังโรงพยาบาลพอลโล ประเทศไทย เพราะว่ามีค่าความถูกต้องอยู่ในระดับสูงที่สุดและความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ผลการสรุปดังกล่าวใกล้เคียงกับในช่วงที่ผ่านมาที่มีผู้วิจัยได้ศึกษาวิธีการจำแนกกลุ่มด้วยข้อมูลต่าง ๆ ซึ่งผู้วิจัยส่วนใหญ่ได้สรุปว่าวิธีต้นไม้ตัดสินใจเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการจำแนกกลุ่ม และจากผลการวิจัย ผู้วิจัยจึงสรุปได้ว่าวิธีต้นไม้ตัดสินใจเป็นวิธีที่ดีที่สุด ซึ่งสามารถนำผลลัพธ์ที่ได้จากแบบจำลองของกฎต้นไม้ตัดสินใจเพื่อใช้ในการทำนาย เป็นแนวทางในการสนับสนุนการตัดสินใจทางการแพทย์เกี่ยวกับการวินิจฉัยโรคไตเรื้อรัง เพื่อให้ได้ผลการวินิจฉัยที่รวดเร็ว และแม่นยำมากขึ้น จะทำให้กลุ่มผู้ป่วยโรคไตเรื้อรังลดน้อยลง

สุรัชย์ จันทร์จรัส และคณะ (2556) ศึกษาเรื่อง การพยากรณ์ราคาหลักทรัพย์ด้วยวิธีการนิวโรฟัซซี ผลการวิจัยพบว่า การนำข้อมูลการซื้อขายในอดีตมาพยากรณ์ราคาหลักทรัพย์ของตลาดหลักทรัพย์แห่งประเทศไทยด้วยวิธีการโครงข่ายประสาทเทียมแบบฟัซซี นั้นไม่สามารถที่จะพยากรณ์ได้อย่างถูกต้องและต่อเนื่อง ดังนั้นการทำการสำรวจส่วนเกินจึงไม่สามารถทำได้ในระยะยาว กล่าวได้ว่าตลาดหลักทรัพย์แห่งประเทศไทยเป็นตลาดที่มีประสิทธิภาพในระดับที่หนึ่ง (Weak Form Efficiency) ซึ่งสอดคล้องกับผลการทดสอบประสิทธิภาพตลาดหลักทรัพย์แห่งประเทศไทย ที่สนับสนุนว่า ตลาดหลักทรัพย์แห่งประเทศไทยมีประสิทธิภาพ ในการศึกษาค้นคว้าครั้งนี้ สนใจในเรื่องของจำนวนข้อมูลและระยะเวลาในการพยากรณ์ดัชนีราคาตลาดหลักทรัพย์ จึงไม่มีการทดสอบเปรียบเทียบกับวิธีการพยากรณ์แบบอื่น ว่ามีวิธีการใดให้ความแม่นยำในการพยากรณ์มากกว่ากัน อีกทั้งข้อมูลที่ใช้สำหรับการศึกษา ในครั้งนี้อยู่ในช่วงที่เกิดภาวะวิกฤตการณ์ทางการเมือง ซึ่งส่งผลกระทบต่อราคาหลักทรัพย์ตกลง

สายชล สินสมบุญทอง (2560) ศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดเกมของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร ผลการวิจัยพบว่า การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มพบว่าวิธีต้นไม้ตัดสินใจมีค่าความถูกต้อง ค่าความระลึกลับ ค่าความถ่วงดุลมากที่สุด และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย น้อยที่สุดคือ 92.1739%, 0.976, 0.948 และ 0.0638 ตามลำดับ ส่วนวิธีซัพพอร์ตเวกเตอร์แมชชีน มีค่าความระลึกลับมากที่สุดและค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยน้อยที่สุดคือ 0.976 และ 0.0870 ตามลำดับ และวิธีการถดถอยโลจิสติกแบบ 2 กลุ่ม มีค่าความระลึกลับมากที่สุดคือ 0.976 เนื่องจากวิธีต้นไม้ตัดสินใจมีประสิทธิภาพในการทำนายผลดีที่สุดที่ 4 ใน 5 ค่า ดังนั้นวิธีต้นไม้ตัดสินใจเป็นวิธีที่ดีที่สุด

จิราภา โสมชัย และคณะ (2557) ศึกษาเรื่อง การจำแนกการมีภาวะเสี่ยงสูงต่อโรคความดันโลหิตสูงด้วยเครื่องจักรการเรียนรู้เอ็กซ์ทรีม ผลการวิจัยพบว่า การแก้ปัญหการจำแนกประชากรที่มีภาวะเสี่ยงสูงต่อโรคความดันโลหิตสูงสามารถใช้

เป็นระบบวิเคราะห์ข้อมูลทางด้านสุขภาพเชิงรุกให้กับหน่วยงานที่มหาวิทยาลัยให้ข้อมูลกลุ่มผู้มีความเสี่ยง เพื่อจัดโปรแกรมการปรับพฤติกรรม ดูแลคนไข้แต่ละประเภทกลุ่มเสี่ยงได้อย่างมีประสิทธิภาพ และเฝ้าระวังภาวะความเสี่ยงของโรคความดันโลหิตสูง ที่จะเกิดขึ้นให้เหมาะสมกับความเสี่ยงของแต่ละกลุ่มและตรงเป้าหมาย จากผลการศึกษาพบว่าขั้นตอนวิธีELM มีความสามารถในการจำแนกกลุ่มประชากร ที่มีภาวะเสี่ยงต่อโรคความดันโลหิตสูง ได้ดีกว่าขั้นตอนวิธีดั้งเดิม คือ MLP-BP และ MLP-LM เมื่อเปรียบเทียบกันทั้งในด้านเวลา และความถูกต้องในการ แก้ปัญหา แต่อย่างไรก็ตาม จากการศึกษาครั้งนี้ก็ยังพบประเด็นปัญหาอื่น ๆ เพิ่มเติม เช่น ปัญหาการจำแนกกลุ่มเนื่องจากลักษณะของข้อมูลเป็นข้อมูลแบบ ไม่สมดุล (Imbalanced data sets) ซึ่งควรจะได้มีการศึกษาวิจัยต่อไปในอนาคต อีกทั้งในปัจจุบัน ขั้นตอนวิธี ELM ก็ยังได้มีการพัฒนาปรับปรุงให้มีประสิทธิภาพไปมากกว่าขั้นตอนวิธี ELM ขึ้นพื้นฐาน จากนักวิจัยหลายคน ดังนั้นในการศึกษาการประยุกต์ใช้ขั้นตอนวิธี ELM ใหม่ ๆ ในอนาคต อาจช่วยเพิ่มประสิทธิภาพของการแก้ปัญหาได้เพิ่มมากขึ้น

สุระสิทธิ์ ทรงม้า (2558) ศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ รุ่นที่ 3 ที่ใช้เทคนิคการแบ่งข้อมูลที่แตกต่างกัน ผลการวิจัยพบว่า ประสิทธิภาพของการใช้เทคนิคแบ่งข้อมูลด้วยขนาดความกว้างเท่ากันและจำแนกข้อมูลด้วยต้นไม้ตัดสินใจรุ่นที่ 3 มีประสิทธิภาพสูงที่สุด โดยมีอัตราการจำแนกข้อมูลถูกต้อง เท่ากับ 92.30% มีอัตราการจำแนกข้อมูลผิดพลาด เท่ากับ 4.89% และสามารถจำแนกข้อมูลที่ไม่สามารถจัดเข้ากลุ่มได้ เท่ากับ 2.81% ผลการทดลองยังแสดงให้เห็นว่าเทคนิคการแบ่งข้อมูลชนิดแบบไม่มีผู้สอนแต่ละชนิดส่งผลต่อประสิทธิภาพ การจำแนกข้อมูลของต้นไม้ตัดสินใจ รุ่นที่ 3 และพบว่าการใช้เทคนิคการแบ่งข้อมูลร่วมกับเทคนิคต้นไม้ตัดสินใจ รุ่นที่ 3 สามารถช่วยกรองข้อมูลที่ไม่อยู่ในกฎที่สร้างขึ้นได้ ซึ่งเป็นสิ่งที่แตกต่างกับเทคนิคต้นไม้ตัดสินใจ รุ่นที่ C4.5 ฉะนั้นวิธีที่นำเสนอหากเลือกนำไปใช้ได้เหมาะสมจะเกิดประโยชน์อย่างมาก

#### 4. ขั้นตอนการดำเนินงาน

##### 4.1 ข้อมูลสำหรับใช้ในการทดลอง

ผู้วิจัยได้ทำการรวบรวมข้อมูลเด็กที่เป็นโรคออทิสติกและไม่เป็นโรคออทิสติก โดยนำข้อมูลดังกล่าวมาจาก เว็บไซต์ Autism Screening Adult Data Set จากนั้นนำข้อมูลมาทำความสะอาดโดยตัดข้อมูลที่มีค่าขาดหาย (Missing Values) ออก

##### 4.2 การดำเนินการทดลอง

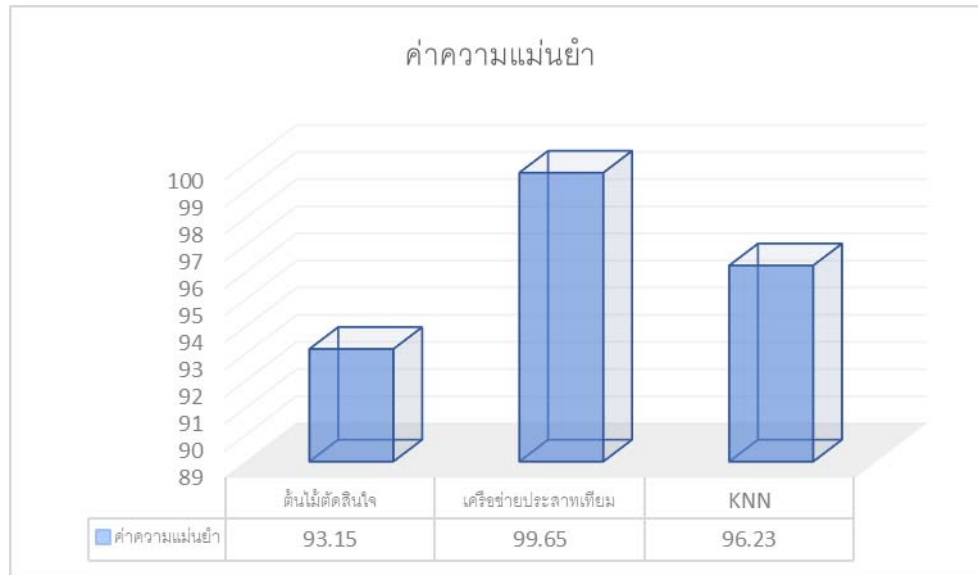
- 1) นำข้อมูลที่ได้จัดเตรียมไว้ในขั้นตอนที่ 4.1 มาทำการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลโรคออทิสติกด้วยเทคนิคการจำแนกข้อมูล จำนวน 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้เคียงที่สุด
- 2) นำผลลัพธ์จากการจำแนกกลุ่มที่ได้จากการทดลองมาหาประสิทธิภาพของการจำแนกกลุ่มโรคออทิสติก
- 3) นำเทคนิคการจำแนกกลุ่มข้อมูลโรคออทิสติกที่มีประสิทธิภาพมากที่สุดไปพัฒนาระบบการจำแนกกลุ่มข้อมูลโรคออทิสติกต่อไป

#### 5. ผลการดำเนินงาน

ผลการทดลองเปรียบเทียบประสิทธิภาพในการการจำแนกกลุ่มข้อมูลโรคออทิสติก จำนวน 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้เคียงที่สุด

ตารางที่ 1 ผลการทดลองเมื่อเปรียบเทียบทั้ง 3 เทคนิค

วิธีวัดประสิทธิภาพ	ค่าความแม่นยำ
ต้นไม้ตัดสินใจ	93.15 %
โครงข่ายประสาทเทียม	99.65 %
เพื่อนบ้านใกล้เคียงที่สุด	96.23 %



ภาพที่ 1 ผลการทดลองเมื่อเปรียบเทียบทั้ง 3 เทคนิค

ผลการทดลองเปรียบเทียบประสิทธิภาพในการจำแนกกลุ่มข้อมูลโรคอหิสติก 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้เคียงที่สุด ผลการวิจัยพบว่า เทคนิคโครงข่ายประสาทเทียม ให้ค่าความแม่นยำมากที่สุด คือ 99.65% รองลงมาเป็น เทคนิคเพื่อนบ้านใกล้เคียงที่สุด ให้ค่าความแม่นยำ คือ 96.23% และสุดท้าย เทคนิคต้นไม้ตัดสินใจ ให้ค่าความแม่นยำ คือ 93.15%

## 6. สรุปผลการดำเนินงานและข้อเสนอแนะ

### 6.1 สรุปผลการดำเนินงาน

งานวิจัยนี้ได้ทำการศึกษาถึงประสิทธิภาพของวิธีการจำแนกกลุ่มข้อมูลโรคอหิสติกด้วยเทคนิคการจำแนกกลุ่มข้อมูลจำนวน 3 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคเพื่อนบ้านใกล้เคียงที่สุด ผลการวิจัยพบว่า เทคนิคทั้ง 3 เทคนิค เทคนิคที่ให้ค่าความแม่นยำมากที่สุด จะมีประสิทธิภาพดีที่สุด ดังนั้นเทคนิคที่ให้ค่าความแม่นยำสูงสุด คือ เทคนิคโครงข่ายประสาทเทียม ให้ค่าความแม่นยำอยู่ที่ 99.65 %

### 6.2 ข้อเสนอแนะ

เนื่องจากเทคนิคการจำแนกกลุ่มข้อมูลมีหลากหลายเทคนิคแต่ 3 เทคนิค ที่เลือกมาเปรียบเทียบนั้นอาจจะไม่ใช่วิธีการที่ดีที่สุด ดังนั้น ควรทดลองใช้วิธีการอื่น ๆ ที่ยังไม่ได้ถูกนำมาทำการเปรียบเทียบทำการจำแนกกลุ่มข้อมูลได้ หากนำทุกวิธีการมาทำการเปรียบเทียบกันหมด อาจจะได้วิธีการจำแนกกลุ่มข้อมูลโรคอหิสติกที่ดีที่สุดจริง ๆ

## 7. เอกสารอ้างอิง

- จิราภา โสมชัย, สิริภัทร เขียวชาญวัฒนา, คำรณ สุนันติ, และนิพนธ์พัทธ์ เมืองโคตร. การจำแนกการมีภาวะเสี่ยงสูงต่อโรคความดันโลหิตสูงด้วยเครื่องจักรการเรียนรู้เอ็กซ์ทรีม. *วารสารวิศวกรรมสาร*. ปีที่ 27. ฉบับที่ 89. (2557): 89-100.
- เดช ธรรมศิริ และพยุง มีสัจ. การจำแนกข้อมูลด้วยวิธีแบบร่วมกันตัดสินใจจากพื้นฐานของเทคนิคต้นไม้ตัดสินใจเทคนิคโครงข่ายประสาทเทียม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ร่วมกับการเลือกตัวแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม. *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*. ปีที่ 21. ฉบับที่ 2. (2554): 293-303.
- นพมาศ ปักเข็ม, ชนิดา จันมณี, และศิวกร อุษย. การจำแนกประเภทภูมิปัญญาท้องถิ่นของไทยแบบอัตโนมัติโดยวิธีการทางเหมืองข้อมูล. *วารสารมหาวิทยาลัยทักษิณ*. ปีที่ 20. ฉบับพิเศษ. (2560): 300-307.
- วิทยา พรพัชรพงศ์. *โครงข่ายประสาทเทียม (Artificial Neural Networks - ANN)*. [ออนไลน์] ค้นเมื่อ 1 พฤษภาคม 2562 จาก <https://www.gotoknow.org/posts/163433>.
- สายชล สินสมบูรณ์ทอง. การเปรียบเทียบประสิทธิภาพในการทำนายผลภาวะการติดยาของเด็กและวัยรุ่นในเขตกรุงเทพมหานคร. รายงานการวิจัย. คณะวิทยาศาสตร์. สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- สุรัชย์ จันทร์จรัส, ชญานิน ชลหาญ, และจิรนนท์ เข็มขันธุ์. การพยากรณ์ราคาหลักทรัพย์ด้วยวิธีการนิวโรฟิวซี. *วารสารวิทยาการจัดการ*. ปีที่ 30. ฉบับที่ 2. (2556): 1-25
- สุรวัชร ศรีเปารยะ, และสายชล สินสมบูรณ์ทอง. การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่งในประเทศอินเดีย. *วารสารวิทยาศาสตร์และเทคโนโลยี*. ปีที่ 25. ฉบับที่ 5. (2560): 839-853.
- สุระสิทธิ์ ทรงม้า. การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจ รุ่นที่ 3 ที่ใช้เทคนิคการแบ่งข้อมูลที่แตกต่างกัน. *SDU Research Journal*. (2558): 29-45.