# การพัฒนาวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค สำหรับการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวาน

# Development of Data Classification using a Hybrid Method of Adaptive Artificial Neural Networks and Particle Swarm Optimization for Identifying Patients at Risk of Diabetes

**อกนิษฐ์ ทองจิตร[1*] พูลพงศ์ สุขสว่าง[1] จตุภัทร เมฆพายัพ[2]**
**Akanit Thongjit[1*], Poonpong Suksawang [1], Jatupat Mekparyup [2]**
*[1] College of Research Methodology and Cognitive Science, Burapha University, Thailand*
*[2] Department of Mathematics, Faculty of Science, Burapha University, Thailand*

**บทคัดย่อ**

งานวิจัยนี้มีวัตถุประสงค์เพื่อ 1) พัฒนาวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค 2) เปรียบเทียบประสิทธิภาพวิธีการจำแนกประเภทข้อมูล 3 วิธี ได้แก่ โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (AANN-PSO) โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (ANN-PSO) และโครงข่ายประสาทเทียมแบบดั้งเดิม (ANN) และ 3) ศึกษาการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวานและกลุ่มปกติ โดยใช้วิธีการจำแนกประเภทข้อมูลที่พัฒนาขึ้นข้อมูลที่ใช้ในการศึกษาเป็นตัวอย่างผู้ป่วยที่มีภาวะเสี่ยงต่อการเป็นโรคเบาหวานในพื้นที่เขตรับผิดชอบของสำนักงานสาธารณสุขจังหวัดนครพนม ปี พ.ศ. 2561 จำนวน 7,000 ระเบียน ผลการวิจัยปรากฏว่า

1. วิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาคที่พัฒนาขึ้น ด้วยฟังก์ชันการแปลงใหม่ $f(s')=\frac{2}{1+e^{-2s'}}-1$ เมื่อ $S'=\frac{S-E(S)}{\sqrt{Var(S)}}$ ทำให้ค่าความชันของฟังก์ชันเป้าหมายลดลง และประสิทธิภาพการจำแนกประเภทข้อมูลมีค่าความแม่นยำเพิ่มขึ้น

2. วิธี AANN-PSO มีประสิทธิภาพการจำแนกประเภทข้อมูลมากกว่าวิธี ANN-PSO และ วิธี ANN ทั้ง 5 สถานการณ์ และมีประสิทธิภาพเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

3. ตัวแปรที่เป็นปัจจัยสำคัญที่ส่งผลต่อการเป็นโรคเบาหวาน ได้แก่ ดัชนีมวลกาย ความดันโลหิตช่วงหัวใจคลายตัว อายุ ความดันโลหิตช่วงหัวใจบีบตัว เส้นรอบเอว และประวัติเบาหวานในญาติสายตรง โดยการจำแนกประเภทข้อมูลกลุ่มเสี่ยงผู้ป่วยเป็นโรคเบาหวาน ด้วยวิธี AANN-PSO มีค่าความแม่นยำ ร้อยละ 92.79 และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.07

**คำสำคัญ:** โครงข่ายประสาทเทียม, การหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค, โรคเบาหวาน

**Corresponding author. E-mail: akanit2524@gmail.com*

**ABSTRACT**

This research aimed to 1) develop a method of data classification using Adaptive Artificial Neural Network and Particle Swarm Optimization (AANN-PSO), 2) compare the performance of the developed data classification method with three types: Adaptive Artificial Neural Network and Particle Swarm Optimization (AANN-PSO), Artificial Neural Network and Particle Swarm Optimization (ANN-PSO) and Artificial Neural Network (ANN) and 3), classify the patients who are at risk of diabetes and normal subjects with the method of Adaptive Artificial Neural Network and Particle Swarm Optimization (AANN-PSO). The data set involved 7,000 patients who were at risk of diabetes, in the area under the responsibility of the Nakhon Phanom Provincial Health Office in the year 2018. The research results were as follows:

1. The data classification using Adaptive Artificial Neural Network and Particle Swarm Optimization (AANN-PSO) with the new conversion function $f(s')=\frac{2}{1+e^{-2s'}}-1$ when $S'=\frac{S-E(S)}{\sqrt{Var(S)}}$ acted to decrease the slope of the target function, while data classification performance increased.

2. Data classification using AANN-PSO resulted in better performance than ANN-PSO and ANN in all five situations. Furthermore, when the sample size was increased, the performance was even better.

3. Factors that affected the risk of diabetes included body mass index, diastolic blood pressure, age, systolic blood pressure, and a family history of diabetes. The classification of patients who are at risk of diabetes by using AANN-PSO had an accuracy of 92.79%, with mean square error of 0.07.

**Keywords**: artificial neural networks, particle swarm optimization, diabetes

## ความนำ

สถานการณ์โรคเบาหวานมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง จากรายงานของสำนักนโยบายและยุทธศาสตร์ สำนักงานปลัดกระทรวงสาธารณสุข พบว่า สถานการณ์โรคเบาหวานในประเทศไทยมีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่องโดยอัตราเสียชีวิตจากโรคเบาหวานต่อประชากรแสนคนของประเทศไทยตั้งแต่ปี พ.ศ. 2556-2558 เท่ากับ 14.93 17.53 และ 17.83 ตามลำดับ (Ministry of Public Health, 2016) จากรายงานดังกล่าวแสดงให้เห็นว่า โรคเบาหวานเป็นหนึ่งในกลุ่มโรคที่เป็นปัญหาสำคัญด้านสาธารณสุข ซึ่งมีปัจจัยนำไปสู่โรคที่มีปัจจัยเสี่ยงร่วม ดังนั้น การคัดกรองหากลุ่มเสี่ยงในการเป็นโรคเบาหวานจึงมีความจำเป็น หากพบผู้ที่มีความเสี่ยงในการเป็นโรคเบาหวานจะสามารถหาแนวทางป้องกัน

หรือชะลอไม่ให้เกิดโรคเบาหวานได้ อีกทั้งยังเป็นข้อมูลเบื้องต้นในการช่วยแพทย์เพื่อวางแผนการรักษาที่เหมาะสมและสามารถควบคุมได้ในระยะแรกตามเป้าหมาย

จากการศึกษางานวิจัยที่ผ่านมาโครงข่ายประสาทเทียม (Artificial Neural Networks) เป็นวิธีการที่นิยมนำมาใช้ในการแก้ปัญหาด้านต่าง ๆ อย่างแพร่หลาย เนื่องจากความสามารถในการเรียนรู้ ปัจจุบันมีการพัฒนาโครงข่ายประสาทเทียมอย่างต่อเนื่อง ซึ่งมีบทบาทอย่างมากในด้านการจำแนกรูปแบบการทำนาย การหาความเหมาะสม และการจัดกลุ่ม ซึ่งมีงานวิจัยจำนวนมากที่นำเสนอวิธีการจัดประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียม ซึ่ง แต่ละงานวิจัยมีสถาปัตยกรรมของโครงข่ายประสาทเทียมที่แตกต่างกันอาทิเช่น ปัญหาด้านธุรกิจ ด้านวิทยาศาสตร์ หรือทางด้านการแพทย์ เป็นต้น ซึ่งการวินิจฉัยโรคหรือจำแนกข้อมูลทางการแพทย์ส่วนใหญ่แล้วนิยมใช้โครงข่ายประสาทเทียม เนื่องจากมีความสามารถในการจำแนกข้อมูล ดังเช่นงานวิจัยของ Botoca, Barden, Botoca, and Alexa (2010) ได้ศึกษาทำนายการป่วยเป็นมะเร็งต่อมลูกหมาก โดยใช้โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (Multi-Layer Perceptron Neural Networks: MLPNN) RBF-MLP และ Recurrent NN (RNN) เปรียบเทียบกับการถดถอยแบบโลจิสติกผลการศึกษาพบว่า ประสิทธิภาพความแม่นยำในการทำนายโดยใช้โครงข่ายประสาทเทียมดีกว่าการถดถอยแบบโลจิสติก ซึ่งสอดคล้องกับ Sumathi and Santhakumaran (2011) ได้ศึกษาโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (MLPNN) สำหรับทำนายการเป็นโรคความดันโลหิตสูง ผลการศึกษาพบว่า MLPNN สามารถทำนายการเป็นโรคความดันโลหิตสูงได้อย่างมีประสิทธิภาพ

จากข้อมูลข้างต้น พบว่าเทคนิคโครงข่ายประสาทเทียมมีความเหมาะสมกับการนำมาทำนายโรคได้อย่างแม่นยำ แต่อย่างไรก็ตาม โครงข่ายประสาทเทียมแบบเดิมนั้นมีข้อด้อย คือ มีการเรียนรู้เพื่อหาค่าคำตอบช้า เนื่องจากมีการปรับค่าน้ำหนักของแต่ละโหนด (Node) ในการนำข้อมูลเข้ามาประมวลผลทั้งหมดทุกรอบจากสาเหตุดังกล่าว ผู้วิจัยจึงเสนอแนวคิดการพัฒนาวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะ (Adaptive Artificial Neural Networks: AANN) มาใช้แทนโครงข่ายประสาทเทียมแบบดั้งเดิม เพื่อให้การทำงานของโครงข่ายประสาทมีความรวดเร็วมากยิ่งขึ้น อีกทั้งยังไม่ลดประสิทธิภาพเดิมของโครงข่ายประสาทเทียม (Sahu & Mishra, 2012)

การหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Particle Swarm Optimization: PSO) เป็นเทคนิคที่นำไปใช้สำหรับหาคำตอบที่เหมาะสมที่สุด ส่งผลให้ตัวแบบสามารถจำแนกประเภทข้อมูลได้แม่นยำมากขึ้นเป็นขั้นตอนการหาค่าเหมาะที่สุดแบบประมาณค่า (Approximation Optimization Algorithms: AOA) และใช้ระยะเวลาในการประมวลผลน้อยแต่ให้คำตอบใกล้เคียงกับคำตอบที่ดีสุด ทำให้วิธีการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (PSO) ได้รับความนิยมมากขึ้น

ดังนั้น การวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีจำแนกประเภทโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค โดยเปรียบเทียบประสิทธิภาพการจำแนกประเภทข้อมูลกับโครงข่ายประสาทเทียมแบบดั้งเดิม และศึกษาการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวานและกลุ่มปกติโดยใช้วิธีการจำแนกประเภทข้อมูลที่พัฒนาขึ้น

**สมมติฐานการวิจัย**

1. วิธีจำแนกประเภทข้อมูลที่พัฒนาขึ้นให้ค่าความแม่นยำมากกว่าวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค และวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิม

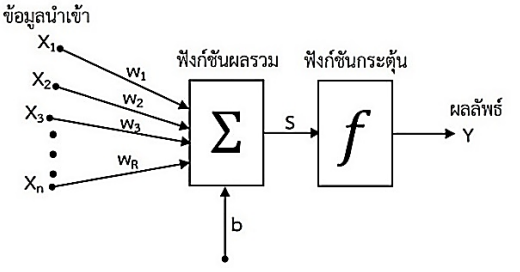
2. วิธีจำแนกประเภทข้อมูลที่พัฒนาขึ้นให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค และวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิม

3. การทำนายผู้ป่วยกลุ่มเสี่ยงในการเป็นโรคเบาหวานโดยใช้วิธีจำแนกประเภทข้อมูลที่พัฒนาขึ้น มีค่าความแม่นยำในการทำนายถูกต้องมากที่สุด โดยมีค่ามากกว่าร้อยละ 80

## วรรณกรรมที่เกี่ยวข้อง

### โครงข่ายประสาทเทียม

ลักษณะทั่วไปของโครงข่ายประสาทเทียม คือ โหนดจำลองมาจากรอยประสานประสาท (Synapse) ระหว่างใยประสาทนำเข้า (Dendrite) ของเซลล์ประสาทตัวหนึ่งและแกนประสาทนำออก (Axon) ของเซลล์ประสาทอีกตัวหนึ่ง โดยมีฟังก์ชันการแปลงเป็นตัวกำหนดสัญญาณส่งออก องค์ประกอบในการทำงานของระบบประสาทในสมองมนุษย์กับโครงข่ายประสาทเทียมมีความคล้ายคลึงกัน ในระบบประสาทของมนุษย์การรับกระแสประสาทจะอาศัยเดนไดรต์ (Dendrites) เป็นตัวนำพา เปรียบเทียบได้กับหน่วยข้อมูลป้อนเข้าในระบบโครงข่ายประสาทเทียม ส่วนการนำกระแสประสาทไปยังเซลล์ประสาทอื่นผ่านทางเอกซอน (Axon) ที่เป็นตัวส่งออกกระแสประสาทเปรียบเทียบได้กับหน่วยผลลัพธ์ในโครงข่ายประสาทเทียม โครงข่ายประสาทเทียมจะผ่านการประมวลผล 2 ขั้นตอน คือ ขั้นตอนการหาผลรวมและขั้นตอนการแปลง ดังภาพที่ 1 โดยโครงข่ายประสาทเทียมจะมีการป้อนข้อมูลเข้าและการกำหนดค่าน้ำหนักซึ่งแบ่งชั้น (Layer) การทำงานออกเป็น 3 ชั้น ดังนี้




**ภาพที่ 1** แบบจำลองการทำงานของโครงข่ายประสาทเทียม

1. ชั้นข้อมูลนำเข้า (Input layer) ทำหน้าที่นำเข้าข้อมูลเข้าสู่โครงข่ายประสาทเทียม โดยข้อมูลนี้ จะนำไปประมวลผลในแต่ละโหนด (Node) ของชั้นถัดไป

2. ชั้นซ่อน (Hidden layer) ทำหน้าที่รับข้อมูลจากชั้นนำเข้าโดยการกำหนดค่าน้ำหนักของข้อมูลจากชั้นนำเข้าก่อนรับข้อมูลเข้าสู่ชั้นซ่อน ชั้นซ่อนจะทำหน้าที่เพิ่มประสิทธิภาพในการจัดกลุ่มข้อมูลก่อนจะส่งต่อข้อมูลไปยังชั้นต่อไป

3. ชั้นข้อมูลส่งออก (Output layer) ทำหน้าที่ส่งออกข้อมูลโดยผ่านการประมวลผลจากฟังก์ชันผลรวม (Summation function: S) เป็นผลรวมของข้อมูลป้อนเข้าที่ถ่วงด้วยค่าน้ำหนัก (W) กับค่าความเอนเอียง (b) และฟังก์ชันการแปลง (Transfer function) จนได้ผลลัพธ์ ดังสมการที่ (1)

$$S = \sum_{i=1}^{n} W_i X_i + b \quad (1)$$

ฟังก์ชันกระตุ้น (Activation function) เป็นส่วนที่ทำหน้าที่แปลงผลรวมของข้อมูลป้อนเข้าให้เป็นผลลัพธ์ในชั้นส่งออก ซึ่งสามารถแบ่งได้ 2 ประเภท ดังนี้

1. ฟังก์ชันการแปลงเชิงเส้น เป็นฟังก์ชันที่สามารถเรียนรู้เพียงความสัมพันธ์เชิงเส้นระหว่างชั้นข้อมูลป้อนเข้าและชั้นข้อมูลส่งออก ดังนั้นจะไม่สามารถหาคำตอบได้สำหรับบางกรณีที่ไม่ใช่เส้นตรง ค่าผลลัพธ์ที่ได้จากการคำนวณจะมีค่าเท่ากับค่าที่ได้จากฟังก์ชันการแปลงเสมอเหมาะสำหรับผลลัพธ์ที่ไม่มีการกำหนดขอบเขตค่าเป้าหมาย เช่น ฟังก์ชันสเต็ป (Step function) ฟังก์ชันเครื่องหมาย (Sign function) ฟังก์ชันการแปลงฮาร์ลิมิต (Hard limit activation function)

2. ฟังก์ชันการแปลงไม่ใช่เชิงเส้น ประกอบด้วยฟังก์ชันการแปลง 2 ชนิด คือ

2.1 ฟังก์ชันซิกมอยด์ (Sigmoid function) ฟังก์ชันการแปลงซิกมอยด์จะบีบช่วงข้อมูลป้อนเข้าที่ไม่จำกัดให้เป็นช่วงของข้อมูลส่งออกที่จำกัด โดยที่ช่วงของข้อมูลส่งออกจะอยู่ในช่วง 0 ถึง 1 ฟังก์ชันการแปลงซิกมอยด์จะแสดงลักษณะของข้อเท็จจริงที่มีความชันเข้าใกล้ศูนย์เมื่อข้อมูลป้อนเข้ามีจำนวนมาก

2.2 ฟังก์ชันไฮเปอร์โบลิกแทนเจนต์ (Hyperbolic tangent function) มีลักษณะเช่นเดียวกับฟังก์ชันการแปลงซิกมอยด์ แต่ต่างกันเพียงช่วงของข้อมูลส่งออกจะอยู่ในช่วง -1 ถึง +1 โดยในงานวิจัยได้เลือกใช้ฟังก์ชันการแปลงนี้ สามารถคำนวณได้ ดังสมการที่ (2)

$$f(s) = \frac{2}{1+e^{-2s}} - 1 \quad (2)$$

โครงข่ายประสาทเทียมสามารถแบ่งได้ 3 แบบ ได้แก่

1. โครงข่ายประสาทแบบป้อนไปข้างหน้าชั้นเดียว (Single-layer feed forward neural networks) ประกอบด้วยชั้นสัญญาณประสาทขาเข้าและชั้นสัญญาณประสาทขาออกเท่านั้น

2. โครงข่ายประสาทแบบป้อนไปข้างหน้าหลายชั้น (Multi-layer feed forward neural networks) มีลักษณะเช่นเดียวกับโครงข่ายประสาทแบบป้อนไปข้างหน้าชั้นเดียว แต่จะมีชั้นซ่อนเพิ่มขึ้น โดยอยู่ระหว่างชั้นข้อมูลนำเข้าและชั้นส่งออก

3. โครงข่ายประสาทแบบแพร่ย้อนกลับ (Back propagation neural networks) มีลักษณะเช่นเดียวกับโครงข่ายประสาทแบบป้อนไปข้างหน้า ซึ่งมีชั้นซ่อนอย่างน้อย 1 ชั้น และมีการวนซ้ำแบบป้อนย้อนกลับอย่างน้อยหนึ่งครั้ง โครงข่ายประสาทเทียมมีขั้นตอนการเรียนรู้ที่แตกต่างกัน สำหรับการวิจัยนี้ใช้โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ โดยมีขั้นตอนการดำเนินการ ดังนี้

3.1 กำหนดชั้นข้อมูลนำเข้า ชั้นซ่อนและชั้นข้อมูลส่งออก อัตราการเรียนรู้ จำนวนรอบ ค่าความคลาดเคลื่อนและค่าน้ำหนักให้โครงข่ายประสาท

3.2 คำนวณค่าผลลัพธ์ของโครงข่าย และปรับค่าน้ำหนักในกรณีที่เพอร์เซ็ปตรอนจำแนกข้อมูลผิดพลาด

3.3 ทำซ้ำชุดข้อมูลการเรียนรู้ (Learning data set) จนกระทั่งเพอร์เซ็ปตรอนจำแนกข้อมูลได้ผลลัพธ์ตามค่าเป้าหมายที่กำหนดไว้

เนื่องจากโครงข่ายประสาทเทียมดั้งเดิม (ANN) ใช้เทคนิคการหาค่าเหมาะสมด้วยวิธี Gradient Descent เพื่อปรับค่าน้ำหนักของโครงข่ายจากความคลาดเคลื่อน (Error) ที่ได้จากการเรียนรู้แต่การเรียนรู้ดังกล่าวทำให้ขั้นตอนการย้อนกลับมักประสบปัญหาการหาค่าที่เหมาะสม (Local optima) ทำให้โครงข่ายประสาทเทียมไม่สามารถจำแนกประเภทข้อมูลได้อย่างมีประสิทธิภาพ (Mohammadi & Mirabedini, 2014)

**การหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค**

การหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (PSO) พัฒนาขึ้นโดย Eberhart and Kennedy (1995) มีแนวคิดการคำนวณเชิงวิวัฒนาการ โดยนำพฤติกรรมเลียนแบบทางสังคมของสัตว์ในการเดินทางหรือหาอาหารของฝูงสัตว์ โดยเฉพาะฝูงนก ฝูงปลา โดยจะเรียกสัตว์แต่ละตัวที่อยู่ในฝูงว่า อนุภาค (Particle) การเคลื่อนที่ประกอบด้วยตำแหน่ง (Position) และความเร็วในการเคลื่อนที่ (Velocity) ซึ่ง Jaroenrat (2015) ได้กำหนดขั้นตอนการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาคประกอบด้วย 7 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 กำหนดอนุภาคเริ่มต้นและค่าความเร็วในการเคลื่อนที่ของแต่ละอนุภาคโดยการสุ่ม

ขั้นตอนที่ 2 คำนวณค่าความเหมาะสม (Fitness value) ของแต่ละอนุภาคด้วยสมการฟังก์ชันเป้าหมาย (Fitness function)

ขั้นตอนที่ 3 เก็บค่าความเหมาะสมที่ดีสุดซึ่งเป็นค่าความเหมาะสมที่ดีสุดในรอบการทำงานปัจจุบัน (Pbest)

ขั้นตอนที่ 4 เก็บค่าน้ำหนักที่ให้ผลลัพธ์ดีที่สุด (Gbest) พิจารณาจากการวนรอบตั้งแต่รอบแรกถึงปัจจุบันซึ่งแตกต่างจาก Pbest ที่พิจารณาการเก็บค่าที่ดีสุดเฉพาะรอบปัจจุบัน

ขั้นตอนที่ 5 ปรับค่าความเร็วในการเคลื่อนที่ของแต่ละอนุภาค ดังสมการที่ (3)

$$V_D^{'} = \left\{w \times V_D\right\} + \left\{n_1 \times \text{random}_1() \right\} \times \left[\text{Pbest}_D - X_D\right] + \left\{n_2 \times \text{random}_2() \times [\text{Gbest}_D - X_D]\right\} \qquad (3)$$

เมื่อ $V_D^{'}$ แทน ค่าความเร็วที่คำนวณได้จากอนุภาครุ่นปัจจุบัน

$V_D$ แทน ค่าความเร็วที่คำนวณได้จากอนุภาครุ่นก่อน

W แทน ค่าการถ่วงน้ำหนัก

$n_1, n_2$ แทน ค่าคงที่อัตราการเร่งในการค้นหา

ขั้นตอนที่ 6 ปรับค่าน้ำหนักของแต่ละอนุภาคโดยใช้ผลลัพธ์จากการคำนวณในสมการ (3) เพื่อปรับปรุงความเร็วที่คำนวณได้จากอนุภาครุ่นปัจจุบัน โดยการเพิ่มข้อมูล ดังสมการที่ (4)

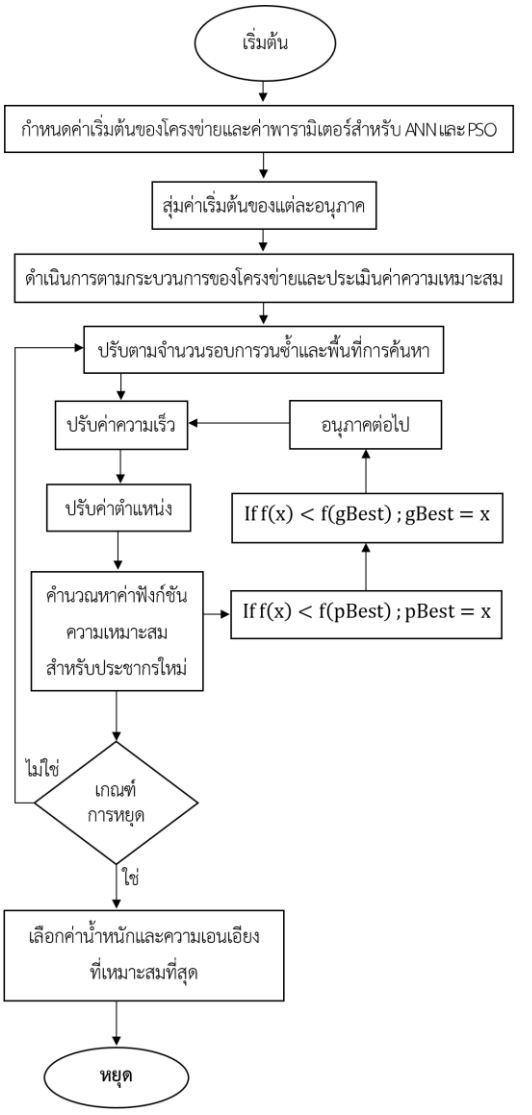
$$X_D^{'} = V_D^{'} + X_D \qquad (4)$$

เมื่อ $X_D^{'}$ แทน ตำแหน่งที่คำนวณได้จากอนุภาครุ่นปัจจุบัน

$X_D$ แทน ตำแหน่งที่คำนวณได้จากอนุภาครุ่นก่อน

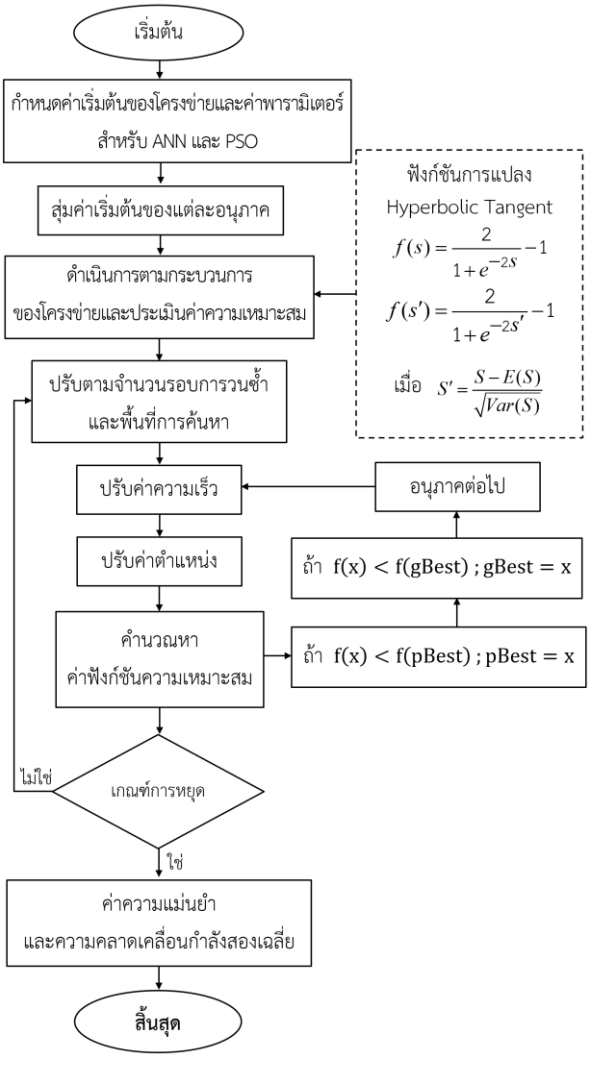
ขั้นตอนที่ 7 ตรวจสอบเงื่อนไข 2 ลักษณะ ได้แก่ เงื่อนไขที่ 1 ผลลัพธ์ที่ได้มีค่าที่สูงหรือดีกว่าคำตอบที่ต้องการและเงื่อนไขที่ 2 จำนวนรอบการทำงานถึงค่าที่กำหนดไว้หากพบว่าการทำงานมีทั้ง 2 เงื่อนไข ให้หยุดการทำงาน แต่ถ้าไม่ครบทั้ง 2 เงื่อนไขให้กลับไปทำซ้ำตั้งแต่ขั้นตอนที่ 2

ซึ่งในการวิจัยครั้งนี้ได้ประยุกต์วิธี PSO มาใช้ปรับค่าน้ำหนักให้กับแต่ละโครงข่ายประสาท เพื่อให้โครงข่ายงานระบบประสาทมีการลู่เข้าหาคำตอบที่เหมาะสม โดยขั้นตอนวิธีการทำงานของโครงข่ายประสาทเทียมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (ANN-PSO) ดังภาพที่ 2




**ภาพที่ 2** ขั้นตอนวิธีโครงข่ายประสาทเทียมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (Geethanjali, Mary Raja Slochanal, & Bhavani, 2008)

อย่างไรก็ตามขั้นตอนวิธี ANN-PSO ยังพบข้อบกพร่องในการเรียนรู้เพื่อหาค่าผลลัพธ์ที่ใช้เวลานาน (Sahu & Mishra, 2012) จากขั้นตอนการคำนวณหาค่าผลลัพธ์ในชั้นซ่อน และจากฟังก์ชันผลรวมด้วยฟังก์ชันการแปลงไฮเปอร์โบลิกแทนเจนต์ พบว่า ผลคูณที่ได้จากการคำนวณด้วยฟังก์ชันผลรวมมีลักษณะการกระจายของข้อมูลแบบไม่ปกติ ส่งผลให้การลู่เข้าหาคำตอบในชั้นข้อมูลส่งออกด้วยฟังก์ชันการแปลงช้า ดังนั้น เพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูล ก่อนจะส่งต่อข้อมูลไปยังชั้นต่อไป ผู้วิจัยจึงมีแนวคิดโดยนำหลักการแปลงข้อมูลให้เป็นค่าปกติมาตรฐาน (Standardization) มาช่วยปรับค่าฟังก์ชันผลรวมของข้อมูลก่อนนำไปใช้ในฟังก์ชันการแปลงให้อยู่ในช่วงแคบลง ส่งผลทำให้การลู่เข้าหาค่าผลลัพธ์เร็วขึ้นและให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยลดลง ดังภาพที่ 3




**ภาพที่ 3** โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (AANN-PSO)

## วิธีดำเนินการวิจัย

บทความวิจัยนี้นำเสนอวิธีการพัฒนาวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค โดยเปรียบเทียบประสิทธิภาพการจำแนกประเภทข้อมูล 3 วิธี ได้แก่ วิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค วิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค และวิธีจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบดั้งเดิม และศึกษาการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวานและกลุ่มปกติโดยใช้วิธีการจำแนกประเภทข้อมูลที่พัฒนาขึ้น

**การเตรียมข้อมูล**

ข้อมูลที่ใช้ในการวิจัยได้จากฐานข้อมูลของสำนักงานสาธารณสุขจังหวัดนครพนม ปี พ.ศ. 2561 ซึ่งเป็นข้อมูลผู้ป่วยที่มาคัดกรองการเป็นโรคเบาหวานที่มีอายุตั้งแต่ 35 ปี ขึ้นไป ในเขตพื้นที่รับผิดชอบ โดยใช้หน่วยตัวอย่างทั้งหมด 7,000 ระเบียน คัดเลือกจากระเบียนข้อมูลที่มีปัจจัยที่ส่งผลต่อการเป็นโรคเบาหวาน จำนวน 10 ตัวแปร ได้แก่ อายุ (AGE) เพศ (GENDER) การสูบบุหรี่ (SMOKE) การดื่มเครื่องดื่มแอลกอฮอล์ (ALCOHOL) ประวัติเบาหวานในญาติสายตรง (DMFAMILY) ประวัติความดันโลหิตสูงในญาติสายตรง (HTFAMILY) ดัชนีมวลกาย (BMI) เส้นรอบเอว (WAIST) ความดันโลหิตช่วงหัวใจบีบตัว (SBP) และความดันโลหิตช่วงหัวใจคลายตัว (DBP) (Nai-aruna & Moungmaia, 2015; Edla & Cheruku, 2017; Karan, Bayraktar, Gümüskaya, & Karlık, 2012)

กำหนดจำนวนโหนดในชั้นข้อมูลป้อนเข้า (N) เท่ากับ 10 โหนด และชั้นซ่อน (M) จำนวน 24 โหนด ชั้นข้อมูลส่งออก (J) จำนวน 2 โหนด และจำนวนรอบเรียนรู้ (R) จำนวน 500 รอบ ค่าพารามิเตอร์ของอัตราการเรียนรู้ เท่ากับ 0.30 สุ่มค่าน้ำหนักที่ได้จากการเรียนรู้โดยใช้โครงข่ายประสาทเทียมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) เท่ากับ $10^{-12}$

**การสร้างตัวแบบจำแนกผู้ป่วยที่มีความเสี่ยงต่อการเป็นโรคเบาหวาน**

การวิเคราะห์ข้อมูลนี้เป็นการวิเคราะห์ประสิทธิภาพของการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวาน ซึ่งใช้วิธีการจำแนกประเภทข้อมูลโดยใช้วิธีที่พัฒนาขึ้น โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค และโครงข่ายประสาทเทียมแบบดั้งเดิมมีขั้นตอนดังนี้

1. แบ่งข้อมูลเป็น 2 ส่วน ได้แก่ ข้อมูลการฝึกเรียนรู้และข้อมูลทดสอบโดยใช้อัตราส่วน 80% และ 20% ตามลำดับ
2. กำหนดจำนวนครั้งในการแบ่งชุดข้อมูลเป็น 100 ครั้ง
3. นำชุดข้อมูลที่ได้มาวิเคราะห์สร้างตัวแบบจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวานโดยใช้วิธีการจำแนกที่พัฒนาขึ้น (AANN-PSO) โครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (ANN-PSO) และโครงข่ายประสาทเทียมแบบดั้งเดิม (ANN)
4. สร้างตัวแบบจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวาน ด้วยชุดข้อมูลฝึกการเรียนรู้
5. ประเมินผลประสิทธิภาพของการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวาน โดยคำนวณค่าความแม่นยำและค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

**การตรวจสอบประสิทธิภาพการจำแนก**

ในงานวิจัยนี้ตรวจสอบประสิทธิภาพของตัวแบบด้วยวิธี K-Fold Cross Validation โดยแบ่งข้อมูลเป็นชุดการฝึกเรียนรู้ (Learning data set) และชุดตรวจสอบ (Validation data set) 10 ส่วนเท่า ๆ กัน โดยใช้ 8 ส่วนเป็นชุด

การเรียนรู้และอีก 2 ส่วนเป็นชุดการตรวจสอบ ซึ่งจะทำสลับกันจนครบทั้งหมด 10 รอบ โดยจะตรวจสอบประสิทธิภาพการจำแนกประเภทข้อมูลด้วยค่าความถูกต้องและค่าคลาดเคลื่อนกำลังสองเฉลี่ย ดังสมการที่ (5) และ (6)

$$\text{Accuracy} = \left( \frac{a+d}{a+b+c+d} \right) \times 100 \tag{5}$$

โดย a คือ จำนวนผู้ป่วยที่เป็นโรคเบาหวานและได้รับการทำนายว่าเป็นโรคเบาหวาน
b คือ จำนวนผู้ป่วยที่เป็นโรคเบาหวานและได้รับการทำนายว่าไม่เป็นโรคเบาหวาน
c คือ จำนวนผู้ป่วยที่ไม่เป็นโรคเบาหวานและได้รับการทำนายว่าเป็นโรคเบาหวาน
d คือ จำนวนผู้ป่วยที่ไม่เป็นโรคเบาหวานและได้รับการทำนายว่าไม่เป็นโรคเบาหวาน

$$\text{MSE} = \sum \frac{\left( Y - Y' \right)^2}{N} \tag{6}$$

โดย $Y$ คือ ค่าที่เกิดขึ้นจริง
$Y'$ คือ ค่าที่เกิดจากการทำนาย

## ผลการวิจัย

### ผลการเตรียมข้อมูล

สำหรับข้อมูลจำแนกประเภท (Categorical data) ทำการแปลงเป็นข้อมูลเชิงตัวเลข (Numerical data) จากนั้นทำการนอร์มอลไลซ์เซชัน (Normalization) ข้อมูลทั้งหมดให้อยู่ในช่วง [-1, 1] เพื่อใช้เป็นข้อมูลนำเข้าในขั้นตอนการเรียนรู้เพื่อสร้างตัวแบบ และจากการวิเคราะห์ข้อมูล สามารถสรุปคุณลักษณะเบื้องต้นของปัจจัยที่ส่งผลต่อกลุ่มมีภาวะเสี่ยงต่อการเป็นโรคเบาหวาน ดังตารางที่ 1

**ตารางที่ 1** คุณลักษณะเบื้องต้นของปัจจัยที่ส่งผลต่อกลุ่มมีภาวะเสี่ยงต่อการเป็นโรคเบาหวาน

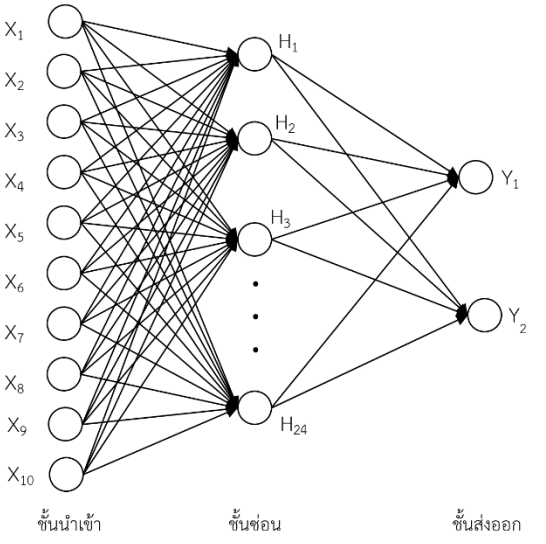
| ตัวแปร | รายละเอียดตัวแปร | ผลการวิเคราะห์ |
|---|---|---|
| 1. AGE | ค่าตัวเลขอายุ | 53.47±0.19 |
| 2. GENDER | 1 = ชาย | 46.10 % |
| | 2 = หญิง | 53.90 % |
| 3. SMOKE | 1 = ไม่สูบ | 71.70 % |
| | 2 = สูบนาน ๆ ครั้ง | 4.70 % |
| | 3 = สูบเป็นครั้งคราว | 0.40 % |
| | 4 = สูบเป็นประจำ | 2.80 % |
| | 9 = ไม่ทราบ | 20.30 % |

**ตารางที่ 1** (ต่อ)

| ตัวแปร | รายละเอียดตัวแปร | ผลการวิเคราะห์ |
|---|---|---|
| 4. ALCOHOL | 1 = ไม่ดื่ม | 69.80 % |
| | 2 = ดื่มนาน ๆ ครั้ง | 7.30 % |
| | 3 = ดื่มเป็นครั้งคราว | 1.50 % |
| | 4 = ดื่มเป็นประจำ | 1.10 % |
| | 9 = ไม่ทราบ | 20.30 % |
| 5. DMFAMILY | 1 = มีประวัติเบาหวานในญาติสายตรง | 18 % |
| | 2 = ไม่มี | 80.40 % |
| | 9 = ไม่ทราบ | 1.60 % |
| 6. HTFAMILY | 1 = มีประวัติความดันโลหิตสูงในญาติสายตรง | 18.10 % |
| | 2 = ไม่มี | 80.40 % |
| | 9 = ไม่ทราบ | 1.50 % |
| 7. BMI | ค่าตัวเลขดัชนีมวลกาย | 23.29±0.17 |
| 8. WAIST | ค่าตัวเลขเส้นรอบเอว | 78.10±0.23 |
| 9. SBP | ค่าตัวเลขความดันโลหิตขณะหัวใจบีบตัว | 115.40±0.16 |
| 10. DBP | ค่าตัวเลขความดันโลหิตขณะหัวใจคลายตัว | 71.53±0.14 |

**ผลการสร้างตัวแบบจำแนกผู้ป่วยที่มีความเสี่ยงต่อการเป็นโรคเบาหวาน**

ตัวแบบจำแนกผู้ป่วยที่มีภาวะเสี่ยงต่อการเป็นโรคเบาหวานโดยใช้โครงข่ายประสาทเทียมมีฟังก์ชันกระตุ้นเชื่อมต่อการทำงานที่โหนดของชั้นนำเข้าชั้นซ่อน และชั้นส่งออกเป็นฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ (Hyperbolic tangent) โดยการปรับน้ำหนักของโครงข่ายระบบประสาทด้วยการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (PSO) แทนการวนปรับค่าแบบเดิมด้วยฟังก์ชันเป้าหมายค่าความคลาดเคลื่อนกำลังสองเฉลี่ย แล้วนำค่าน้ำหนักและค่าเอนเอียงที่เหมาะสมที่สุดที่ได้จากการเรียนรู้ไปใช้สำหรับสร้างตัวแบบการจำแนก ดังภาพที่ 4




**ภาพที่ 4** ตัวแบบโครงข่ายประสาทเทียมจำแนกผู้ป่วยที่มีภาวะเสี่ยงต่อการเป็นโรคเบาหวาน

การเรียนรู้ของโครงข่ายประสาทเทียมในการจำแนกประเภทผู้ป่วยที่มีภาระเสี่ยงต่อการเป็นโรคเบาหวาน สามารถอธิบายได้ ดังนี้

ส่วนที่ 1 จำนวนโหนดของชั้นนำเข้ามี 10 โหนด ซึ่งมีค่าเท่ากับปัจจัยที่ส่งผลต่อการเป็นโรคเบาหวานจำนวน 10 ตัวแปร ได้แก่

$X_1$ แทน อายุ
$X_2$ แทน เพศ
$X_3$ แทน การสูบบุหรี่
$X_4$ แทน การดื่มเครื่องดื่มแอลกอฮอล์
$X_5$ แทน ประวัติเบาหวานในญาติสายตรง
$X_6$ แทน ประวัติความดันโลหิตสูงในญาติสายตรง
$X_7$ แทน ดัชนีมวลกาย
$X_8$ แทน เส้นรอบเอว
$X_9$ แทน ความดันโลหิตช่วงหัวใจบีบตัว
$X_{10}$ แทน ความดันโลหิตช่วงหัวใจคลายตัว

ส่วนที่ 2 จำนวนโหนดของชั้นซ่อนมี 24 โหนด โดยค่าน้ำหนักและค่าความเอนเอียงที่นำมาใช้ในการสร้างตัวแบบได้จากการเรียนรู้ของโครงข่ายประสาทเทียมร่วมกับการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค

ส่วนที่ 3 จำนวนโหนดของชั้นส่งออกมี 2 โหนด แสดงค่าผลการวินิจฉัยของผู้ป่วยมีผลลัพธ์จำนวนสองค่า คือ 1 แทน เป็นโรคเบาหวาน และ 0 แทน ไม่เป็นโรคหวาน

**ผลการตรวจสอบประสิทธิภาพการจำแนกผู้ป่วยที่มีความเสี่ยงต่อการเป็นโรคเบาหวาน**

จากผลการวิจัยสามารถสรุปได้ว่า การจำแนกประเภทด้วยกลุ่มเสี่ยงในการเป็นโรคเบาหวานโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (AANN-PSO) มีประสิทธิภาพในการจำแนกประเภทข้อมูลดีกว่าโครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (ANN-PSO) โดยมีค่าความถูกต้องในการจำแนกผู้ป่วยโรคเบาหวานสูงสุดเท่ากับ 92.79% และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.07 แสดงดังตารางที่ 3

**ตารางที่ 3** ผลการจำแนกผู้ป่วยกลุ่มเสี่ยงในการเป็นโรคเบาหวานโดยใช้ตัวแบบ AANN-PSO ANN-PSO และANN

| วิธีการจำแนกประเภท | Accuracy | MSE | |
|---|---|---|---|
| | | ข้อมูลฝึกเรียนรู้ | ข้อมูลทดสอบ |
| 1. AANN-PSO | 92.79 | 0.1175 | 0.0729 |
| 2. ANN-PSO | 90.57 | 0.1196 | 0.0932 |
| 3. ANN | 88.50 | 0.1563 | 0.1040 |

## สรุปและอภิปรายผลการวิจัย

จากการเปรียบเทียบประสิทธิภาพการจำแนกผู้ป่วยที่มีภาวะเสี่ยงต่อการเป็นโรคเบาหวาน โดยใช้โครงข่ายประสาทเทียมร่วมกับการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาคหรือตัวแบบ BPNN-PSO ดังตารางที่ 3 แสดงให้เห็นว่าตัวแบบการจำแนกประเภทข้อมูลผลการศึกษาการจำแนกประเภทกลุ่มเสี่ยงในการเป็นโรคเบาหวาน โดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (AANN-PSO) ปรากฏว่า ค่าความแม่นยำในการจำแนกกลุ่มมีค่าเท่ากับร้อยละ 92.79 และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย มีค่าเท่ากับ 0.07 ซึ่งมีค่าความแม่นยำมากกว่าร้อยละ 80 เป็นไปตามสมมติฐานของการวิจัยข้อ 3 และมีประสิทธิภาพสูงกว่าโครงข่ายประสาทเทียมแบบดั้งเดิมผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค (ANN-PSO) และโครงข่ายประสาทเทียมแบบดั้งเดิม (ANN) ที่มีค่าความแม่นยำ ร้อยละ 90.57 และ 88.50 ตามลำดับ และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 0.09 และ 0.10 ตามลำดับ และจากผลลัพธ์ที่ได้จากการจำแนกประเภทข้อมูลนี้ แสดงให้เห็นว่าหลักการแปลงค่าข้อมูลให้เป็นค่าปกติมาตรฐานมาประยุกต์ร่วมกับการจำแนกประเภทข้อมูลโดยปรับค่าฟังก์ชันผลรวมข้อมูลให้อยู่ในช่วงเดียวกันและมีค่าพิสัยลดลงทำให้การลู่เข้าหาคำตอบในชั้นข้อมูลส่งออกเร็วขึ้น ส่งผลให้ประสิทธิภาพของตัวแบบการจำแนกประเภทข้อมูลเพิ่มขึ้นและค่าความคลาดเคลื่อนในการทำนายลดลงได้ สอดคล้องกับงานวิจัยของ Jain, Shukla, and Wadhvani (2018), Wang and Tang (2015) และ Smyth and Speed (2003) ที่ใช้หลักการแปลงค่าข้อมูลให้เป็นค่าปกติมาตรฐานมาประยุกต์ร่วมกับการจำแนกประเภทข้อมูล โดยปรับค่าข้อมูลให้อยู่ในช่วงเดียวกันและมีค่าพิสัยลดลงทำให้การลู่เข้าหาคำตอบในชั้นข้อมูลส่งออกเร็วขึ้น ส่งผลให้ประสิทธิภาพของตัวแบบการจำแนกประเภทข้อมูลเพิ่มขึ้นและค่าความคลาดเคลื่อนในการทำนายลดลง

**ข้อเสนอแนะในการนำผลการวิจัยไปใช้**

1. การจำแนกประเภทข้อมูลออกเป็นสองกลุ่ม ควรตรวจสอบคุณลักษณะข้อมูล เพื่อลดความเอนเอียง (Bias) และการจัดกลุ่มที่ผิดประเภท (Misclassified)

2. การกำหนดค่าฟังก์ชันการแปลงควรใช้ฟังก์ชันการแปลงในชั้นซ่อน ส่วนชั้นข้อมูลส่งออกควรใช้ค่าแท้จริงเพื่อประสิทธิภาพการจำแนกประเภทข้อมูล

**ข้อเสนอแนะในการวิจัยต่อไป**

1. การศึกษานี้เป็นการศึกษาข้อมูลตัวอย่างของผู้ป่วยที่มาคัดกรองการเป็นโรคเบาหวาน ในเขตพื้นที่บริการของสำนักงานสาธารณสุขจังหวัดนครพนม ในปี พ.ศ. 2561 ตั้งแต่เดือนมกราคมถึงกันยายน 2561 เท่านั้น เมื่อเวลาเปลี่ยนไปอาจมีผลต่อจำนวนตัวแปรทำนายที่อาจมีปัจจัยอื่นที่ส่งผลต่อการเกิดโรคเบาหวาน จึงควรมีการศึกษาตัวแปรอื่นที่มีความสัมพันธ์กับการเกิดโรคเพิ่มขึ้น และใช้เทคนิควิธีการจำแนกประเภทที่หลากหลาย เพื่อให้ได้คำตอบที่หลากหลายในการหาค่าผลลัพธ์ในการตัดสินใจร่วมกัน เช่น เทคนิควิธีการเรียนรู้ร่วมกัน (Ensemble learning) เป็นต้น

2. ในส่วนของข้อมูลที่นำมาใช้ในการจำแนกประเภท ถ้ามีขนาดใหญ่และตัวแปรทำนายมาก อาจจะใช้วิธีการเลือกคุณลักษณะข้อมูล (Feature selection) เพื่อเลือกเฉพาะตัวแปรที่สำคัญมาใช้ในการเรียนรู้ และสร้างตัวแบบจัดประเภทข้อมูล และถ้าจำนวนกลุ่มต่างกันมากอาจใช้เทคนิคการปรับข้อมูลให้สมดุล (Balance data) เพื่อให้ข้อมูลในแต่ละกลุ่มไม่แตกต่างกันมากจะช่วยเพิ่มประสิทธิภาพค่าความแม่นยำให้กับตัวแบบจัดประเภทข้อมูล

3. ในงานวิจัยนี้ได้นำวิธีการจำแนกประเภทข้อมูลโดยใช้โครงข่ายประสาทเทียมแบบปรับเหมาะผสมผสานการหาค่าเหมาะสมที่สุดแบบกลุ่มอนุภาค มาประยุกต์กับข้อมูลผู้ป่วยที่มาคัดกรองการเป็นโรคเบาหวาน ในเขตพื้นที่บริการของสำนักงานสาธารณสุขจังหวัดนครพนม ในปี พ.ศ. 2561 ซึ่งเป็นข้อมูลที่มีระดับความสัมพันธ์กันไม่มาก ดังนั้นในการศึกษาครั้งต่อไปควรศึกษากับข้อมูลทางด้านอื่นที่ข้อมูลมีคุณลักษณะและการแจกแจงแบบอื่น และมีกลุ่มในการจำแนกมากกว่าสองกลุ่ม เช่น การจำแนกกลุ่มบุคลิกภาพ การจำแนกประเภทเอกสาร เป็นต้น

## กิตติกรรมประกาศ

ขอขอบพระคุณนายแพทย์สำนักงานสาธารณสุขจังหวัดนครพนมที่อนุเคราะห์ข้อมูล และขอขอบคุณบุคลากรของสำนักงานสาธารณสุขจังหวัดนครพนมที่ให้ความสะดวกในการเข้าศึกษาข้อมูล

## เอกสารอ้างอิง

Botoca, C., Bardan, R., Botoca, M., & Alexa, F. (2010). Prostate cancer prognosis evaluation assisted by neural networks. *WSEAS Transactions on Computers*, *9*(2), 164-73.

Eberhart, R., & Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on* (pp. 39-43). IEEE.

Edla, D. R., & Cheruku, R. (2017). Diabetes-finder: a bat optimized classification system for type-2 diabetes. *Procedia Computer Science*, *115*, 235-242.

Geethanjali, M., Mary Raja Slochanal, S., & Bhavani, R. (2008). PSO trained ANN-based differential protection scheme for power transformers. *Neurocomputing*, *71*(4-5), 904-918.

Jain, S., Shukla, S., & Wadhvani, R. (2018). Dynamic selection of normalization techniques using data complexity measures. *Expert Systems with Applications*, *106*(1), 252-262.

Jaroenrat, K. (2015). Particle Swarm Optimization for Open Shortest Path First Network's Traffic Engineering. *Information Technology Journal, 11*(1), 43-52.

Karan, O., Bayraktar, C., Gümüskaya, H., & Karlık, B. (2012). Diagnosing diabetes using neural networks on small mobile devices. *Expert Systems with Applications*, *39*(1), 54-60.

Ministry of Public Health. (2016) *Department of Mental Health*. [Online]. https://www.dmh.go.th/news-dmh/view.asp?id=25634

Mohammadi, N., & Mirabedini, S. J. (2014). Comparison of particle swarm optimization and backpropagation algorithms for training feedforward neural network. *Journal of Mathematics and Computer Science, 12*(1), 113-123.

Nai-aruna, N., & Moungmaia, R. (2015). Comparison of classifiers for the risk of diabetes prediction. *Procedia Computer Science, 69*(1), 132-142.

Sahu, B., & Mishra, D. (2012). A Novel feature selection algorithm using particle swarm optimization for cancer microarray data*. Procedia Engineering, 38*(1*),* 27-31.

Smyth, G. K., & Speed, T. (2003). Normalization of cDNA microarray data. *Methods, 31*(4), 265-273.

Sumathi, B., & Santhakumaran, A. (2011). Pre-diagnosis of hypertension using artificial neural network. *Journal of Computer Science and Technology*, *11*(2), 42-48.

Wang, W. L., & Tang, M. H. (2015). A Normalization process to standardize handwriting data collected from multiple resources for recognition. *Procedia Computer Science, 61*(1), 402-409.