

การเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลเพื่อวิเคราะห์ปัจจัยความเสี่ยง  
ที่ส่งผลต่อการเกิดโรคไฮเปอร์ไทรอยด์ด้วยเทคนิคเหมืองข้อมูล

**Comparison of Data Classification Efficiency to Analyze Risk  
Factors that Affect the Occurrence of Hyperthyroid using  
Data Mining Techniques**

ณัฐวดี หงษ์บุญมี<sup>1</sup> และ ประภาสรี ตรีพาณิชกุล<sup>2</sup>

*Nattavadee Hongboonmee<sup>1</sup> and Praphasiri Trepanichkul<sup>2</sup>*

<sup>1,2</sup>ภาควิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร

<sup>1,2</sup>*Department of Computer Science and Information Technology, Faculty of Science, Naresuan University*

Received: March 22, 2019; Revised: May 17, 2019; Accepted: May 20, 2019; Published: June 25, 2019;

**ABSTRACT** – This research aims to compare the efficiency of data classification by 3 types of data mining algorithms, artificial neural networks, naive bayes and decision tree in order to obtain the most efficient model that will be analyzed for factors affecting the risk of hyperthyroid by reducing inputs individually. The data used in the experiment are data from hospitals in Phitsanulok, there are 323 datasets. The data for analysis are 12 factors. The comparison results showed that the classification of data using artificial neural networks gave the highest efficiency with 82.97% accuracy, which is more than a decision trees and naive bayes with efficiency values of 79.87% and 68.11%, respectively. Effect on the risk of hyperthyroid, it was found that the important symptoms were mood swings and fatigue. The personal factors that are important are gender. In addition to finding factors, this research can also use the data classification model developed to predict the risk of hyperthyroid on smart phone. To help support decision-making in the analysis of hyperthyroid disease risk, can be self-screening, and can continue to guide the treatment of doctors and patients.

**KEYWORDS:** Hyperthyroid, Data Mining, Artificial Neural networks, Naive Bayes, Decision Tree.

**บทคัดย่อ** – งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลด้วยอัลกอริทึมเหมืองข้อมูลสามแบบคือโครงข่ายประสาทเทียม การเรียนรู้แบบเบย์และต้นไม้ตัดสินใจ เพื่อให้ได้อัลกอริทึมที่มีประสิทธิภาพสูงสุดที่จะถูกนำมาวิเคราะห์หาปัจจัยที่ส่งผลต่อความเสี่ยงการเกิดโรคไฮเปอร์ไทรอยด์โดยการลดการนำเข้าที่ละปัจจัย ซึ่งข้อมูลที่นำมาใช้ในการทดลองเป็นข้อมูลจากโรงพยาบาลในจังหวัดพิษณุโลกจำนวน 323 ชุดข้อมูล ข้อมูลสำหรับการวิเคราะห์มีจำนวน 12 ปัจจัย ผลการเปรียบเทียบพบว่า การจำแนกข้อมูลโดยใช้โครงข่ายประสาทเทียมให้ค่าประสิทธิภาพสูงสุดโดยมีค่าความถูกต้อง 82.97% ซึ่งมากกว่าต้นไม้ตัดสินใจและการเรียนรู้แบบเบย์ที่มีค่าประสิทธิภาพความถูกต้อง 79.87% และ 68.11% ตามลำดับ ผลการค้นหปัจจัยที่ส่งผลต่อความเสี่ยงโรคไฮเปอร์ไทรอยด์ พบว่าปัจจัยลักษณะอาการที่มีความสำคัญคือ อารมณ์แปรปรวนและเหนื่อยง่าย ส่วนปัจจัยส่วนบุคคลที่มีความสำคัญคือเพศ นอกจากการค้นหปัจจัยแล้วงานวิจัยนี้ยังสามารถนำแบบจำลองการจำแนกข้อมูลที่ได้มาพัฒนาระบบการพยากรณ์ความเสี่ยงโรคไฮเปอร์ไทรอยด์บนสมาร์ตโฟน เพื่อ

**ช่วยสนับสนุนการตัดสินใจในส่วนของการวิเคราะห์ความเสี่ยงโรคไฮเปอร์ไทรอยด์ช่วยคัดกรองด้วยตัวเองเบื้องต้นและสามารถแนะแนวทางการรักษาของแพทย์และผู้ป่วยได้ต่อไป**

**คำสำคัญ:** ไฮเปอร์ไทรอยด์, เหมืองข้อมูล, โครงข่ายประสาทเทียม, การเรียนรู้แบบเบย์, ดัชนีตัดสินใจ

## 1. บทนำ

โรคไทรอยด์เป็นโรคเกี่ยวกับต่อมไร้ท่อชนิดหนึ่งและในปัจจุบันประเทศไทยมีจำนวนผู้ป่วยโรคไทรอยด์เพิ่มมากขึ้น โดยเฉพาะโรคไฮเปอร์ไทรอยด์ซึ่งพบผู้ป่วยในอัตราส่วน 1 คนต่อ 50 คน [1] อาการของโรคไฮเปอร์ไทรอยด์ค่อนข้างคลุมเครือและคล้ายกับอาการเจ็บป่วยอื่นๆ ทั้งนี้ถ้าหากผู้ป่วยมีภาวะของโรคไฮเปอร์ไทรอยด์ที่ไม่รุนแรงจะไม่ค่อยแสดงอาการของโรคอย่างชัดเจนมากนัก อย่างไรก็ตามภาวะโรคไฮเปอร์ไทรอยด์ก็ถือเป็นโรคที่ส่งผลกระทบต่อการใช้ชีวิตประจำวันได้ ถึงแม้ว่าปัจจุบันวิวัฒนาการทางการแพทย์มีความเจริญก้าวหน้ามีการพัฒนาอุปกรณ์ทางการแพทย์ในการตรวจวินิจฉัยโรคที่มีความแม่นยำและรวดเร็ว แต่การประเมินความเสี่ยงในการเกิดโรคไฮเปอร์ไทรอยด์เบื้องต้นจะเป็นประโยชน์ต่อการวางแผนป้องกันรักษาและดูแลสุขภาพเนื่องจากการรักษาโรคไฮเปอร์ไทรอยด์ต้องได้รับการดูแลรักษาที่เหมาะสมกับความรุนแรงและระยะเวลาที่ผู้ป่วยมีอาการ

จากปัญหาดังกล่าว คณะผู้วิจัยจึงได้สร้างระบบพยากรณ์ระดับความเสี่ยงโรคไฮเปอร์ไทรอยด์ขึ้นโดยนำเทคนิคเหมืองข้อมูลมาประยุกต์ใช้วิเคราะห์ความเสี่ยงในการเกิดโรคไฮเปอร์ไทรอยด์ เพื่อให้ผู้ที่ต้องการทราบโอกาสเสี่ยงของตนเองในการเกิดโรคไฮเปอร์ไทรอยด์สามารถวิเคราะห์ความเสี่ยงได้ด้วยตนเองโดยผ่านโมบายแอปพลิเคชันบนสมาร์ตโฟน นอกจากนี้ยังศึกษาถึงปัจจัยที่ส่งผลกระทบต่อความเสี่ยงการเกิดโรคไฮเปอร์ไทรอยด์เพื่อเป็นแนวทางในการป้องกันรักษาโรคไฮเปอร์ไทรอยด์ได้อีกทางหนึ่ง

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 โรคไฮเปอร์ไทรอยด์

ไทรอยด์ (Thyroid gland) [1] เป็นชื่อของต่อมชนิดหนึ่งมีรูปร่างคล้ายผีเสื้ออยู่ที่คอใต้ลูกกระเดือก ทำหน้าที่สร้างฮอร์โมนเพื่อควบคุมการเผาผลาญพลังงานของร่างกาย อุณหภูมิของร่างกาย ความแข็งแรงของกล้ามเนื้อ ระดับไขมันในเลือดยรวมทั้งอารมณ์และความรู้สึก โรคไทรอยด์ที่พบบ่อยมีอยู่ 2 ชนิด คือ 1) โรคไฮโปไทรอยด์ (Hypothyroid) เป็นภาวะที่ต่อมไทรอยด์ผลิตฮอร์โมนออกมาน้อยทำให้ร่างกายเกิดอาการขาดฮอร์โมนไทรอยด์จึงมีการเผาผลาญพลังงานน้อยกว่าปกติ พบอาการเช่น อ้วน เฉื่อยชา เหนื่อยง่าย เป็นต้น 2) โรคไฮเปอร์ไทรอยด์ (Hyperthyroid) หมายถึงภาวะที่ต่อมไทรอยด์สร้างฮอร์โมนออกมามากกว่าปกติ อาการเป็นพิษของไทรอยด์เกิดจากฮอร์โมนที่ถูกสร้างเพิ่มมากขึ้นและหลั่งไปในกระแสเลือดออกฤทธิ์กระตุ้นอวัยวะต่างๆ ให้ทำงานมากขึ้น ซึ่งผู้ป่วยจะมีความเจ็บป่วยรุนแรงไม่เท่ากัน การรักษาจึงทำได้ในระดับที่ควบคุมให้ภาวะของโรคเบาลงเพื่อให้สามารถใช้ชีวิตได้อย่างเป็นปกติมากที่สุด

ในงานวิจัยนี้เลือกศึกษาโรคไฮเปอร์ไทรอยด์ เนื่องจากพบอัตราการเกิดโรคมามากกว่าโรคไฮโปไทรอยด์ [1]

### 2.2 การทำเหมืองข้อมูล

การทำเหมืองข้อมูล (Data mining) คือ [2] การค้นหาความรู้ในฐานข้อมูลเป็นกระบวนการที่กระทำกับข้อมูลที่มีจำนวนมากเพื่อค้นหารูปแบบแนวทางและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยอาศัยหลักสถิติ การเรียนรู้ของเครื่อง (Machine learning) และการจดจำรูปแบบ (Pattern recognition)

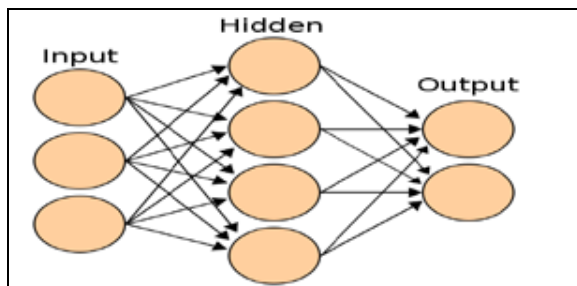
สารสนเทศที่ได้จากการทำเหมืองข้อมูลอาจนำมาสร้างการพยากรณ์หรือสร้างตัวแบบสำหรับการจำแนกหน่วยหรือกลุ่มหรือแสดงความสัมพันธ์ระหว่างหน่วยต่างๆ หรือให้ข้อสรุปของสาระในการสร้างฐานข้อมูล การดำเนินงานมักอยู่ในลักษณะของการสร้างแบบจำลอง (Modeling) [3] ที่อธิบาย

ความเป็นไปหรือสภาพการณ์หนึ่งที่เกิดขึ้นแล้วนำแบบจำลองนี้มาใช้อธิบายสถานการณ์ที่ยังไม่เกิดหรือไม่ทราบคำตอบ

ในปัจจุบันเทคนิคเหมืองข้อมูลมีหลายรูปแบบซึ่งคณะผู้วิจัยได้ศึกษาเทคนิคที่ใช้ทำเหมืองข้อมูลเพื่อการจำแนกข้อมูล ได้แก่ เทคนิคโครงข่ายประสาทเทียม การเรียนรู้แบบเบย์และ ต้นไม้ตัดสินใจ

## 2.3 เทคนิคโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Artificial neural network) หรือ ANN [4] เป็นเทคโนโลยีที่มีพื้นฐานมาจากการจำลองการทำงานของเซลล์สมองของมนุษย์โดยมีโครงสร้างเป็นกลุ่มของ Node ที่เชื่อมโยงถึงกันในแต่ละ Layer คือ Input Layer, Hidden Layer และ Output Layer สามารถแสดงได้ดังรูปที่ 1



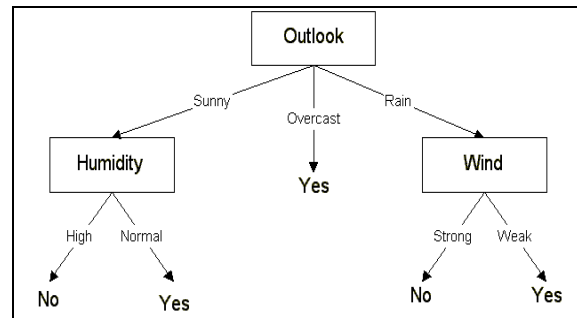
รูปที่ 1. แสดงโครงสร้างโครงข่ายประสาทเทียม

## 2.4 เทคนิคการเรียนรู้แบบเบย์

การเรียนรู้แบบเบย์เป็นเทคนิคที่ใช้ทฤษฎีความน่าจะเป็นตามกฎของเบย์ (Bayes' Theorem) [4] เพื่อหาว่าสมมติฐานใดน่าจะถูกต้องที่สุดโดยใช้ความรู้ก่อนหน้า (Prior knowledge) ได้แก่ ความน่าจะเป็นก่อนหน้าสำหรับสมมติฐานหนึ่งๆ ร่วมกับข้อมูล เช่น ความน่าจะเป็นที่สังเกตได้สำหรับสมมติฐานหนึ่งๆ เพื่อหาสมมติฐานที่ดีที่สุด การเรียนรู้แบบเบย์อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน

## 2.5 เทคนิคต้นไม้ตัดสินใจ

ต้นไม้ตัดสินใจ (Decision tree) [5] เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้ โดยรูปแบบของต้นไม้ตัดสินใจประกอบด้วยโหนดแรกสุดเรียกว่า โหนดราก (Root node) โหนดลูก (Child node) และโหนดใบ (Leaf node) สามารถแสดงได้ดังรูปที่ 2



รูปที่ 2. แสดงโครงสร้างของต้นไม้ตัดสินใจ

## 2.6 การตรวจสอบแบบไขว้

การตรวจสอบแบบไขว้ (K-fold cross validation) เป็นวิธีการทดสอบโดยแบ่งกลุ่มข้อมูลออกเป็น k กลุ่ม [4] ข้อมูลแต่ละกลุ่มจะมีจำนวนเท่ากัน ข้อมูลแต่ละกลุ่มจะถูกละทิ้งให้เป็นข้อมูลฝึกสอนและข้อมูลทดสอบ ตัวอย่างเช่น การแบ่งชุดข้อมูลออกเป็น 5 กลุ่มหรือ  $k = 5$  โดยในรอบที่ 1 จะใช้ชุดข้อมูล 2-5 เป็นชุดข้อมูลฝึกสอนและใช้ชุดข้อมูลชุดที่ 1 เป็นชุดข้อมูลทดสอบ และทำการสลับเปลี่ยนข้อมูลจนครบ 5 รอบ สามารถอธิบายการทำงานดังแสดงในรูปที่ 3

|                       |   |   |   |   |         |   |
|-----------------------|---|---|---|---|---------|---|
| Iteration 1: train on | 2 | 3 | 4 | 5 | test on | 1 |
| Iteration 2: train on | 1 | 3 | 4 | 5 | test on | 2 |
| Iteration 3: train on | 1 | 2 | 4 | 5 | test on | 3 |
| Iteration 4: train on | 1 | 2 | 3 | 5 | test on | 4 |
| Iteration 5: train on | 1 | 2 | 3 | 4 | test on | 5 |

รูปที่ 3. แสดงการแบ่งชุดข้อมูลแบบ k-fold โดยที่  $k=5$

## 2.7. การพัฒนาแอปพลิเคชันบนสมาร์ตโฟน

สมาร์ตโฟนหมายถึงอุปกรณ์อิเล็กทรอนิกส์แบบพกพาเช่น โทรศัพท์มือถือ แท็บเล็ตคอมพิวเตอร์ สำหรับงานวิจัยนี้มุ่งเน้นที่การพัฒนาแบบบนโทรศัพท์มือถือและแท็บเล็ตคอมพิวเตอร์ที่ทำงานบนระบบปฏิบัติการแอนดรอยด์ เนื่องจากปัจจุบันระบบปฏิบัติการแอนดรอยด์มีความนิยมใช้งานแพร่หลายมากกว่าระบบปฏิบัติการบนอุปกรณ์พกพาคืออื่น [6] นอกจากนี้ระบบปฏิบัติการแอนดรอยด์ยังเป็นซอฟต์แวร์ระบบเปิด (Open source) โดยอนุญาตให้นักพัฒนาสามารถดาวน์โหลด Source Code มาปรับแต่งและพัฒนาแอปพลิเคชันบนระบบแอนดรอยด์ด้วยตนเองได้

## 2.8 งานวิจัยที่เกี่ยวข้อง

จากการศึกษาพบว่างานวิจัยจำนวนมากไม่น้อยได้พัฒนาขึ้นเพื่อพยากรณ์การเกิดโรคหลายโรค คณะผู้วิจัยศึกษางานวิจัยที่มีความเกี่ยวข้อง ดังต่อไปนี้

นงเยาว์และพรณี [7] ได้ศึกษาเรื่องการจำแนกผู้ป่วยโรคหัวใจขาดเลือดและผู้ป่วยโรคหัวใจรูปแบบอื่นโดยใช้เทคนิคเหมืองข้อมูล โดยใช้ข้อมูลผลการตรวจของผู้ป่วยจากห้องปฏิบัติการประกอบด้วยคุณลักษณะ 15 ตัวแปร ใช้เทคนิคโครงข่ายประสาทเทียมและเทคนิคเคเนียร์เรสเซเนเบอร์ ผลการทดลองพบว่าโมเดลของโครงข่ายประสาทเทียมมีค่าดีที่สุดโดยได้ค่าความถูกต้องเท่ากับ 87.45%

กรวิภาและคณะ [8] ได้พัฒนาระบบวิเคราะห์ความเสี่ยงของการเกิดโรคอ้วนลงพุงและแนะนำอาหารสำหรับผู้ป่วยโดยประยุกต์ใช้เทคนิคเหมืองข้อมูล ซึ่งเทคนิคที่นำมาใช้คือต้นไม้ตัดสินใจ ข้อมูลที่นำมาใช้ในการวิเคราะห์เป็นข้อมูลที่ได้มาจากการเก็บข้อมูลโดยใช้แบบสอบถามจากบุคลากรในโรงพยาบาล ผลการทดสอบความถูกต้องของโมเดลโดยใช้โปรแกรม WEKA พบว่าโมเดลของการวิเคราะห์ความเสี่ยงการเกิดโรคอ้วนลงพุงมีค่าความถูกต้องเท่ากับ 97.36% และโมเดลของการแนะนำอาหารสำหรับผู้ป่วยมีค่าความถูกต้องเท่ากับ 87.92%

นพรัตน์และคณะ [9] ศึกษาเรื่องระบบสนับสนุนทางการแพทย์เพื่อวิเคราะห์โรคหลอดเลือดแดงโป่งพองโดยใช้เทคนิคเหมืองข้อมูล เลือกใช้เทคนิค Decision Tree, Naïve Bayes และ Neural Network ผลการศึกษาพบว่าแบบจำลองจาก Decision Tree มีประสิทธิภาพค่าความถูกต้องสูงสุดเท่ากับ 98.10%

ณัฐกรและคณะ [10] วิจัยเรื่องระบบพยากรณ์ความรุนแรงและโอกาสเสียชีวิตในเด็กที่ได้รับบาดเจ็บหรืออุบัติเหตุ โดยใช้โครงข่ายประสาทเทียมเพื่อช่วยในการตัดสินใจและช่วยลดความผิดพลาดอันเนื่องมาจากการวินิจฉัยและการรักษาของบุคลากรทางการแพทย์ ข้อมูลที่นำมาใช้ในการเรียนรู้มาจากฐานข้อมูลการเฝ้าระวังการบาดเจ็บโดยเครือข่ายเฝ้าระวังการบาดเจ็บในเด็กแห่งประเทศไทย ผลการทดสอบความแม่นยำของข้อมูลพบว่าโมเดลที่ใช้มีความแม่นยำสูงถึง 99.26%

สุภาพรและคณะ [11] ศึกษาเรื่องการคัดกรองสุขภาพสำหรับผู้ที่เกี่ยวข้องต่อการเป็นโรคข้อเข่าเสื่อมโดยใช้เหมืองข้อมูลเก็บข้อมูลเพื่อใช้ในการวิเคราะห์จำนวน 383 ชุดข้อมูล เลือกใช้ 4 อัลกอริทึมจากเหมืองข้อมูล คือ Decision tree, Naïve byes,

SMO และ Neural network ผลการทดลองพบว่า Naïve byes มีประสิทธิภาพค่าความถูกต้องสูงสุดที่ 92.14%

พวงทองและคณะ [12] วิจัยเรื่องการพัฒนากระบวนการคัดกรองกลุ่มเสี่ยงโรคเบาหวานโดยใช้เทคนิคเหมืองข้อมูลเลือกใช้เทคนิคต้นไม้ตัดสินใจ ผลการทดลองพบว่าประสิทธิภาพด้านความถูกต้องของตัวแบบพยากรณ์การเกิดโรคเบาหวานมีค่าเท่ากับร้อยละ 74.50

ไววิทย์และมหศักดิ์ [13] ศึกษาเรื่องการพัฒนายอดขายปลีกแก่สรณด้วยวิธีการวิเคราะห์อนุกรมเวลาและโครงข่ายประสาทเทียม เนื่องจากเทคนิคโครงข่ายประสาทเทียมนั้นระบบสามารถที่จะเรียนรู้จากปัจจัยต่างๆและสามารถที่จะสกัดลักษณะออกมาได้เอง ดังนั้นงานวิจัยนี้จึงนำเทคนิคโครงข่ายประสาทเทียมมาสร้างโมเดลในการพยากรณ์ ผลการทดลองพบว่าแบบจำลองจากโครงข่ายประสาทเทียมมีความแม่นยำเท่ากับ 89%

การศึกษาศาสนาและงานวิจัยข้างต้นทำให้คณะผู้วิจัยได้แนวคิดที่จะนำเทคนิคเหมืองข้อมูลมาใช้ในการสร้างแบบจำลองเพื่อวิเคราะห์ระดับความเสี่ยงโรคไฮเปอร์ไทรอยด์และศึกษาปัจจัยที่มีผลต่อโรคไฮเปอร์ไทรอยด์ โดยเทคนิคที่เลือกมาทำการทดลองในงานวิจัยครั้งนี้ คือเทคนิคโครงข่ายประสาทเทียม การเรียนรู้แบบเบย์และต้นไม้ตัดสินใจ เนื่องจากทั้งสามเทคนิคนี้เป็นที่นิยมใช้ในการวิเคราะห์พยากรณ์และจำแนกลักษณะข้อมูล

## 3. วิธีการดำเนินงาน

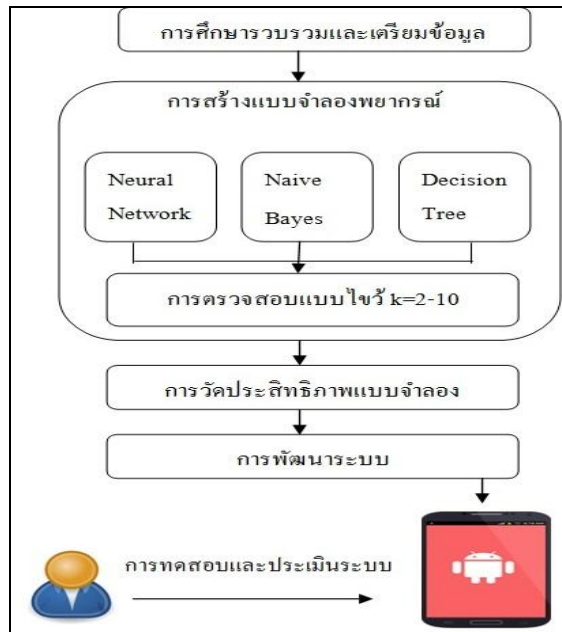
วิธีการดำเนินงานวิจัย มีขั้นตอนการดำเนินงาน 5 ขั้นตอนได้แก่

1) การศึกษารวบรวมและเตรียมข้อมูล 2) การสร้างแบบจำลองพยากรณ์ 3) การวัดประสิทธิภาพแบบจำลอง 4) การพัฒนาระบบและ 5) การทดสอบและประเมินระบบ ดังรูปที่ 4 และรายละเอียดในแต่ละขั้นตอนมีดังนี้

### 3.1 การศึกษารวบรวมและเตรียมข้อมูล

การศึกษาทำความเข้าใจปัญหา พบว่าโรคไฮเปอร์ไทรอยด์เป็นภาวะหนึ่งที่เกิดขึ้นโดยไม่ทราบสาเหตุที่แน่นอนแต่มีปัจจัยเสี่ยงต่างๆ ที่ทำให้เกิดโรคไฮเปอร์ไทรอยด์ ดังนั้นการค้นหาปัจจัยเสี่ยงที่ทำให้เกิดโรคไฮเปอร์ไทรอยด์จะช่วยให้วินิจฉัยโรคได้ถูกต้องและเป็นประโยชน์ต่อการวางแผนป้องกันรักษาเพื่อลดความเสี่ยงของผู้ที่มีโอกาสป่วยเป็นโรคไฮเปอร์ไทรอยด์ได้ การทำความเข้าใจข้อมูลที่นำมาใช้ในการ

วิเคราะห์ โดยการศึกษาวิจัยนี้ใช้ข้อมูลของผู้ป่วยโรคไฮเปอร์ไทรอยด์จากโรงพยาบาลในจังหวัดพิษณุโลก เก็บข้อมูลตั้งแต่เดือนมกราคม-ธันวาคม พ.ศ. 2560 จำนวนทั้งหมด 323 ชุด ข้อมูลได้จากโรงพยาบาลเอกชน 2 แห่ง ได้แก่ โรงพยาบาลรังสีรักษา และโรงพยาบาลพิษณุเวช



รูปที่ 4. แสดงขั้นตอนการดำเนินงานวิจัย

การเตรียมข้อมูล (Data preparation) เป็นการเตรียมข้อมูลให้พร้อมนำไปใช้วิเคราะห์ข้อมูล โดยการรวบรวมข้อมูลและนำไปเตรียมข้อมูลดังนี้ การคัดเลือกข้อมูลโดยเลือกปัจจัยที่เกี่ยวข้องกับการเป็นโรคไฮเปอร์ไทรอยด์ ใช้วิธีการเก็บรวบรวมจากข้อมูลผู้ป่วยและคัดเลือกข้อมูลปัจจัยที่เกี่ยวข้องซึ่งส่งผลต่อความเสี่ยงโรคโรคไฮเปอร์ไทรอยด์ โดยการสอบถามจากผู้เชี่ยวชาญด้านการแพทย์ จึงสามารถสกัดปัจจัยที่เกี่ยวข้องได้จำนวน 12 ปัจจัย ประกอบด้วยปัจจัยส่วนบุคคล ได้แก่ อายุ เพศ ปัจจัยด้านลักษณะอาการ ได้แก่ เหนื่อยง่าย ใจสั่น อารมณ์แปรปรวน น้ำหนักเพิ่มหรือลดลงอย่างผิดปกติ เหนื่อยออกง่าย/หิวน้ำ ร่ายกายอ่อนแรง/อ่อนเพลีย มือสั่น นอนไม่หลับ ตาโปน คอโตกว่าปกติ รวมทั้งหมด 12 ปัจจัยหนึ่งคลาสผลลัพธ์โดยแบ่งระดับความเสี่ยงเป็น 4 ระดับ คือ ความเสี่ยงมาก ความเสี่ยงปานกลาง ความเสี่ยงน้อยและไม่มีความเสี่ยง ข้อมูลปัจจัยนี้ได้ผ่านการตรวจสอบจากอายุรแพทย์เพื่อความถูกต้องตามหลักการของการวิเคราะห์ความเสี่ยงโรคไฮเปอร์ไทรอยด์ โดยการแบ่งระดับความเสี่ยงเป็น 4 ระดับนั้นกำหนดโดย

ผู้เชี่ยวชาญซึ่งมีรายละเอียดดังนี้ 1) ไม่มีความเสี่ยง หมายถึง พบความเสี่ยง 0-3 ข้อ สามารถเก็บข้อมูลได้จำนวน 87 ชุดข้อมูล 2) ความเสี่ยงน้อย หมายถึง พบความเสี่ยง 4-6 ข้อ สามารถเก็บข้อมูลได้จำนวน 89 ชุดข้อมูล 3) ความเสี่ยงปานกลาง หมายถึง พบความเสี่ยง 7-9 ข้อ สามารถเก็บข้อมูลได้จำนวน 74 ชุดข้อมูลและ 4) ความเสี่ยงสูง หมายถึง พบความเสี่ยงมากกว่า 10 ข้อขึ้นไป สามารถเก็บข้อมูลได้จำนวน 73ชุดข้อมูลรวมข้อมูลทั้งหมด 323 ชุดข้อมูล

สำหรับการแปลงข้อมูล (Data transformation) ดำเนินการแปลงข้อมูลให้อยู่ในรูปแบบที่สามารถนำไปวิเคราะห์ได้รายละเอียดในตารางที่ 1

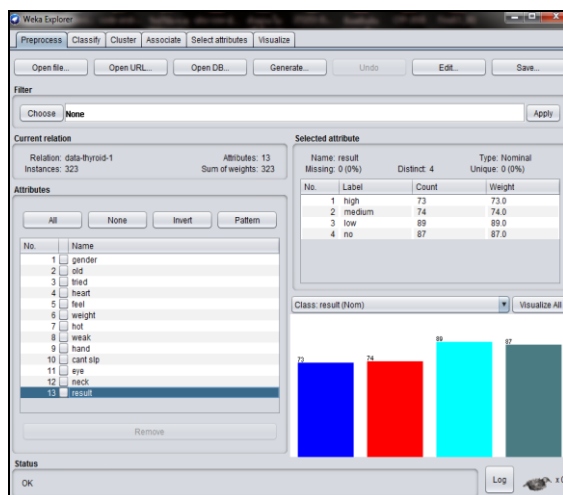
ตารางที่ 1. ปัจจัยที่ใช้ในการสร้างแบบจำลอง

| ลำดับ | ชื่อตัวแปร                       | รายละเอียด                                                      |
|-------|----------------------------------|-----------------------------------------------------------------|
| 1     | เพศ                              | 1=เพศหญิง 2=เพศชาย                                              |
| 2     | อายุ                             | 1=15-25 ปี 2=26-35 ปี<br>3=36-45 ปี 4=46-60 ปี<br>5=61 ปีขึ้นไป |
| 3     | เหนื่อยง่าย                      | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 4     | ใจสั่น                           | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 5     | อารมณ์แปรปรวน                    | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 6     | น้ำหนักเพิ่มหรือลดลงอย่างผิดปกติ | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 7     | เหงื่อออกง่าย/หิวน้ำ             | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 8     | ร่างกายอ่อนแรง/อ่อนเพลีย         | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 9     | มือสั่น                          | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 10    | นอนไม่หลับ                       | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 11    | ตาโปน                            | 1=มีอาการ 2=ไม่มีอาการ                                          |
| 12    | คอโตกว่าปกติ                     | 1=มีอาการ 2=ไม่มีอาการ                                          |

| ลำดับ | ชื่อตัวแปร                                        | รายละเอียด                                                                                |
|-------|---------------------------------------------------|-------------------------------------------------------------------------------------------|
| 13    | ผลการวิเคราะห์<br>ความเสี่ยงโรค<br>ไฮเปอร์ไทรอยด์ | 1=มีความเสี่ยงมาก<br>2=มีความเสี่ยงปาน<br>กลาง<br>3=มีความเสี่ยงน้อย<br>4=ไม่มีความเสี่ยง |

### 3.2 การสร้างแบบจำลอง

การสร้างแบบจำลอง (Model) โดยนำข้อมูลที่เก็บบันทึกมาแปลงให้อยู่ในรูปแบบไฟล์ ARFF และนำไปวิเคราะห์ด้วยขั้นตอนวิธีโครงข่ายประสาทเทียม การเรียนรู้แบบเบย์และต้นไม้ตัดสินใจ เลือกวิธีวัดประสิทธิภาพแบบจำลองด้วยการตรวจสอบไขว้แบบ k=2-10 (K-fold cross validation) ในโปรแกรม WEKA 3.8



รูปที่ 5. แสดงการนำข้อมูลเข้าโปรแกรม WEKA

### 3.3 การวัดประสิทธิภาพแบบจำลอง

การประเมินแบบจำลองพิจารณาจาก 1) ประสิทธิภาพของแบบจำลองประเมินได้จากค่าความถูกต้อง (Accuracy) ในการจำแนกข้อมูล 2) ค่า Confusion matrix ซึ่งเป็นตารางสรุปการจำแนกข้อมูลของแบบจำลองได้ถูกต้องและไม่ถูกต้อง [2] สามารถนำมาคำนวณเพื่อวัดประสิทธิภาพโดยการหาค่าดังนี้ 1) ค่าความแม่นยำ (Precision) คำนวณจากค่าของข้อมูลที่มีการ

จำแนกถูกต้องโดยพิจารณาจากผลรวมของการจำแนกข้อมูลทั้งหมด 2) ค่าความระลึก (Recall) คำนวณจากค่าของข้อมูลที่ผลลัพธ์ถูกต้องโดยพิจารณาจากข้อมูลของผลลัพธ์เดียวกัน 3) ค่าความถ่วงดุล (F-measure) คำนวณได้จากค่าเฉลี่ยระหว่างค่าความแม่นยำและค่าความระลึก รายละเอียดดังสมการที่ 1-4 ต่อไปนี้ [4], [14]

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{F-measure} = \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \times 2 \quad (4)$$

TP = ข้อมูลเป็นจริงและผลการทำนายบอกว่าจริง

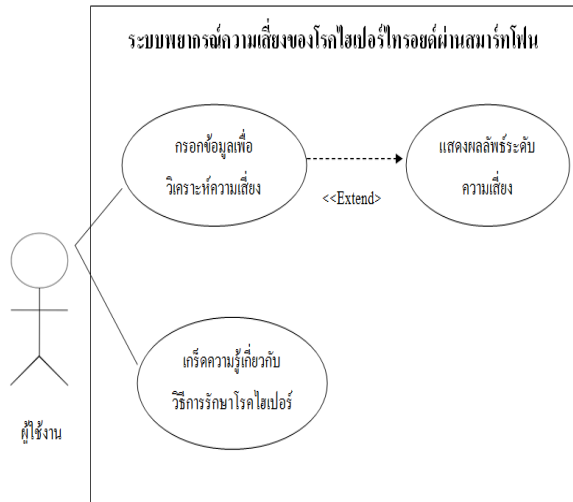
TN = ข้อมูลเป็นข้อมูลไม่จริงและผลการทำนายบอกว่าไม่จริง

FP = ข้อมูลเป็นจริงแต่ผลการทำนายบอกว่าไม่จริง

FN = ข้อมูลเป็นข้อมูลไม่จริงแต่ผลการทำนายบอกว่าจริง

### 3.4 การพัฒนาระบบ

การพัฒนาระบบพยากรณ์ความเสี่ยงของโรคไฮเปอร์ไทรอยด์เริ่มจากการออกแบบระบบเพื่อให้ทราบถึงขั้นตอนการทำงานต่างๆ จึงใช้แผนภาพ UML (Unified Modeling Language) แสดงให้เห็นถึงฟังก์ชันการทำงานของระบบที่สัมพันธ์กับผู้ใช้ โดยรูปที่ 6 แสดงรายละเอียด Use Case Diagram ของระบบ



รูปที่ 6. แผนภาพ Use Case Diagram ของระบบ

แผนภาพ Use Case Diagram แสดงฟังก์ชันหลักของระบบ 2 ส่วน ได้แก่ 1) เมินวิเคราะห์ความเสี่ยงโรคลิพโรยด์ ผู้ใช้ต้องการกรอกข้อมูลเพื่อระบบจะประมวลผลตามแบบจำลองที่สร้างไว้แล้วแสดงผลลัพธ์ระดับความเสี่ยงที่สอดคล้องกับเงื่อนไขที่กำหนดและ 2) เกร็ดความรู้เกี่ยวกับวิธีการรักษาโรคลิพโรยด์

เมื่อออกแบบระบบแล้วขั้นต่อไปจึงดำเนินการพัฒนาระบบในรูปแบบแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์เพื่ออำนวยความสะดวกให้กับผู้ใช้ โดยใช้โปรแกรม Android Studio และภาษา JAVA ร่วมกับภาษา XML เป็นภาษาหลักในการพัฒนา

### 3.5 การทดสอบและประเมินระบบ

การทดสอบและประเมินผลระบบ ได้ทำการสร้างแบบสอบถามประเมินคุณภาพระบบจากผู้ใช้งานทั้งหมด 30 คน

## 4. ผลการทดลองและการอภิปรายผล

ผลการทดลองของงานวิจัยประกอบด้วย 7 ส่วน ได้แก่ 1) ผลการสร้างแบบจำลองด้วยโครงข่ายประสาทเทียม 2) ผลการสร้างแบบจำลองด้วยการเรียนรู้แบบเบย์ 3) ผลการสร้างแบบจำลองด้วยต้นไม้ตัดสินใจ 4) ผลการเปรียบเทียบประสิทธิภาพแบบจำลอง 5) ผลการวิเคราะห์ปัจจัย 6) ผลการพัฒนาระบบและ 7) ผลการประเมินระบบ มีรายละเอียดดังต่อไปนี้

### 4.1 ผลการสร้างแบบจำลองด้วยโครงข่ายประสาทเทียม

การทดสอบโดยใช้แบบจำลองจากเทคนิคโครงข่ายประสาทเทียมแบบ Multilayer Perceptron ทำการกำหนดอัตราการเรียนรู้ (Learning Rate) ที่ 0.3 ทำการเทรนทั้งหมด 500 รอบ แบ่งชุดข้อมูลทดสอบแบบไขว้ กำหนดค่า  $k=2$  ถึง 10 เพื่อเปรียบเทียบค่าความถูกต้องของแบบจำลอง พบว่าแบบจำลองที่ดีที่สุดได้แก่ แบบจำลองที่การแบ่งทดสอบ 4-fold ได้ค่าความถูกต้องสูงที่สุด 82.97% ค่าความแม่นยำ 83.40% ค่าความระลึก 83.00% ค่าความถ่วงดุล 82.80% มีรายละเอียดดังตารางที่ 2

ตารางที่ 2. ค่าประสิทธิภาพจากเทคนิคโครงข่ายประสาทเทียม

| K-fold  | Accuracy (%) | Precision (%) | Recall (%) | F-measure (%) |
|---------|--------------|---------------|------------|---------------|
| 2-fold  | 77.39        | 78.00         | 77.40      | 77.10         |
| 3-fold  | 79.87        | 79.90         | 79.90      | 79.70         |
| 4-fold  | 82.97        | 83.40         | 83.00      | 82.80         |
| 5-fold  | 76.78        | 76.70         | 76.80      | 76.60         |
| 6-fold  | 80.18        | 80.20         | 80.20      | 80.10         |
| 7-fold  | 77.70        | 77.70         | 77.70      | 77.70         |
| 8-fold  | 80.49        | 80.40         | 80.50      | 80.30         |
| 9-fold  | 78.94        | 78.80         | 78.90      | 78.70         |
| 10-fold | 80.49        | 80.40         | 80.50      | 80.30         |

### 4.2 ผลการสร้างแบบจำลองด้วยการเรียนรู้แบบเบย์

การทดสอบแบบจำลองจากเทคนิคการเรียนรู้แบบเบย์โดยใช้การแบ่งชุดข้อมูลด้วยการทดสอบแบบไขว้ กำหนดค่า  $k=2$  ถึง 10 เพื่อเปรียบเทียบค่าความถูกต้องของแบบจำลอง พบว่าแบบจำลองที่ดีที่สุดได้แก่ 5-fold ได้ค่าความถูกต้องสูงที่สุด 68.11% ค่าความแม่นยำ 68.70% ค่าความระลึก 68.10% ค่าความถ่วงดุล 67.70% รายละเอียดในตารางที่ 3

ตารางที่ 3. ค่าประสิทธิภาพจากเทคนิคการเรียนรู้แบบเบย์

| K-fold | Accuracy (%) | Precision (%) | Recall (%) | F-measure (%) |
|--------|--------------|---------------|------------|---------------|
| 2-fold | 65.63        | 65.70         | 65.60      | 65.20         |



| K-fold  | Accuracy (%) | Precision (%) | Recall (%) | F-measure (%) |
|---------|--------------|---------------|------------|---------------|
| 3-fold  | 64.08        | 64.40         | 64.10      | 63.60         |
| 4-fold  | 67.49        | 67.80         | 67.50      | 66.80         |
| 5-fold  | 68.11        | 68.70         | 68.10      | 67.70         |
| 6-fold  | 64.39        | 64.60         | 64.40      | 63.80         |
| 7-fold  | 66.25        | 66.70         | 66.30      | 65.70         |
| 8-fold  | 67.18        | 67.90         | 67.20      | 66.80         |
| 9-fold  | 65.63        | 66.70         | 65.60      | 65.40         |
| 10-fold | 67.18        | 67.60         | 67.20      | 66.80         |

#### 4.3 ผลการสร้างแบบจำลองด้วยต้นไม้ตัดสินใจ

การทดสอบแบบจำลองจากเทคนิคต้นไม้ตัดสินใจ โดยเลือกใช้อัลกอริทึม J48 ใช้การแบ่งชุดข้อมูลด้วยการทดสอบแบบไขว้ กำหนดค่า k=2 ถึง 10 เพื่อเปรียบเทียบค่าความถูกต้องของแบบจำลอง ดังตารางที่ 4. พบว่าแบบจำลองที่ดีที่สุด ได้แก่ การแบ่งทดสอบแบบ 5-fold ได้ค่าความถูกต้องสูงที่สุด 79.87% ค่าความแม่นยำ 80.00% ค่าความระลึก 79.90% ค่าความถ่วงดุล 79.60%

ตารางที่ 4. ค่าประสิทธิภาพจากเทคนิคต้นไม้ตัดสินใจ

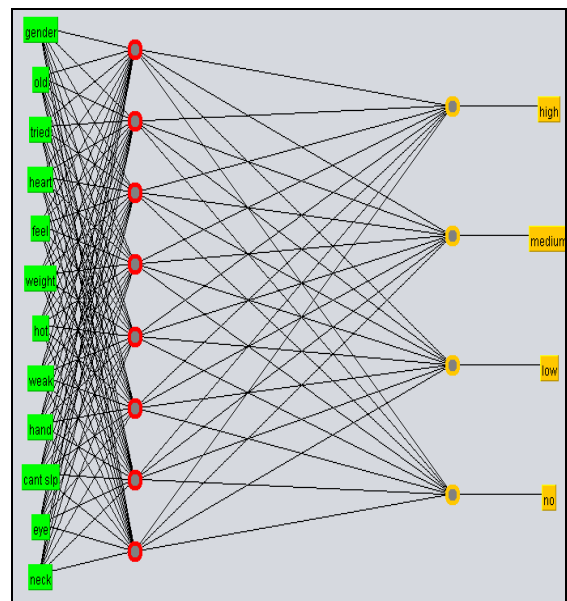
| K-fold  | Accuracy (%) | Precision (%) | Recall (%) | F-measure (%) |
|---------|--------------|---------------|------------|---------------|
| 2-fold  | 68.42        | 68.80         | 68.40      | 68.30         |
| 3-fold  | 77.39        | 77.30         | 77.40      | 77.20         |
| 4-fold  | 75.23        | 75.20         | 75.20      | 75.00         |
| 5-fold  | 79.87        | 80.00         | 79.90      | 79.60         |
| 6-fold  | 78.01        | 77.90         | 78.00      | 77.80         |
| 7-fold  | 78.63        | 78.50         | 78.60      | 78.40         |
| 8-fold  | 78.63        | 78.80         | 78.60      | 78.50         |
| 9-fold  | 76.78        | 76.80         | 76.80      | 76.70         |
| 10-fold | 79.25        | 79.40         | 79.30      | 79.00         |

#### 4.4 ผลการเปรียบเทียบประสิทธิภาพแบบจำลอง

จากการเปรียบเทียบประสิทธิภาพการทำงานของทั้งสามเทคนิค ได้ผลการทดลองดังตารางที่ 5 จะเห็นได้ว่าแบบจำลองจากเทคนิคโครงข่ายประสาทเทียม (ANN) มีค่าประสิทธิภาพดีกว่าแบบจำลองจากเทคนิคต้นไม้ตัดสินใจและการเรียนรู้แบบเบย์ โดยมีค่าความถูกต้อง 82.972% ค่าความแม่นยำ 83.40% ค่าความระลึก 83.00% และค่าความถ่วงดุล 82.80% โดยมีโครงสร้างแบบจำลองดังรูปที่ 7

ตารางที่ 5. ผลการเปรียบเทียบประสิทธิภาพแบบจำลอง

| Model          | Accuracy (%) | Precision (%) | Recall (%) | F-measure (%) |
|----------------|--------------|---------------|------------|---------------|
| Neural Network | 82.97        | 83.40         | 83.00      | 82.80         |
| Naive Bayes    | 68.11        | 68.70         | 68.10      | 67.70         |
| Decision Tree  | 79.87        | 80.00         | 79.90      | 79.60         |



รูปที่ 7. แบบจำลองโครงข่ายประสาทเทียมประสิทธิภาพสูงสุด

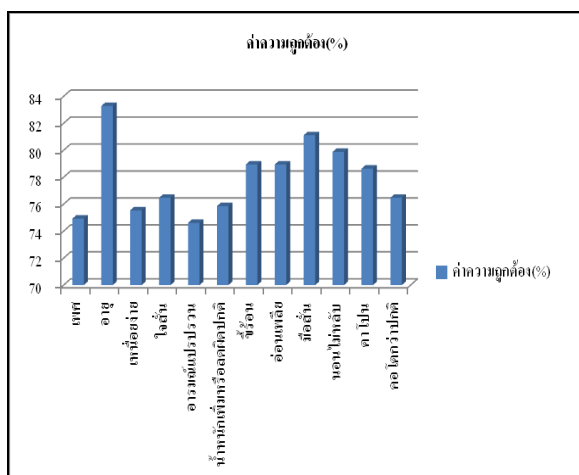


#### 4.5 ผลการวิเคราะห์ปัจจัยที่มีผลต่อการเกิดโรคไฮเปอร์ไทรอยด์

การวิเคราะห์ปัจจัยโดยนำอัลกอริทึมที่สร้างแบบจำลองที่ดีที่สุด ซึ่งก็คือแบบจำลองจากโครงข่ายประสาทเทียมมาวิเคราะห์เพื่อค้นหาปัจจัยที่ส่งผลต่อการเกิดโรคไฮเปอร์ไทรอยด์ ด้วยการตรวจสอบปัจจัยย้อนกลับโดยลดจำนวนนำเข้าที่ละปัจจัยและตรวจสอบค่าประสิทธิภาพความถูกต้องที่ลดลง โดยถ้าค่าความถูกต้องลดลงมากที่สุดนั้นคือปัจจัยที่มีความสำคัญที่สุด จากผลในรูปที่ 8 และตารางที่ 6 แสดงให้เห็นการวิเคราะห์ปัจจัย ซึ่งพิจารณาจากค่าความถูกต้องที่ลดลงมากที่สุด 3 ลำดับ ได้แก่ 1) ปัจจัยด้านอารมณ์แปรปรวน 2) ปัจจัยด้านเพศและ 3) มีอาการเหนื่อยง่ายตามลำดับ

ตารางที่ 6. ผลการวิเคราะห์ปัจจัยสูงสุด 3 อันดับ

| ลำดับ | ปัจจัย        | ความถูกต้อง (%) |
|-------|---------------|-----------------|
| 1     | อารมณ์แปรปรวน | 74.61           |
| 2     | เพศ           | 74.92           |
| 3     | เหนื่อยง่าย   | 75.54           |



รูปที่ 8. ผลการวิเคราะห์ปัจจัยการเกิดโรคไฮเปอร์ไทรอยด์

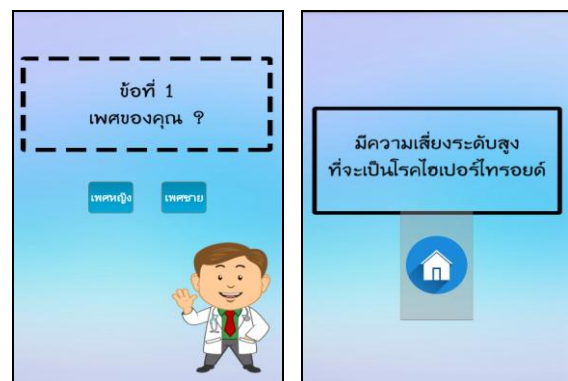
จากผลการทดลองวิเคราะห์ปัจจัยทั้งหมด 12 ปัจจัย ด้วยการตรวจสอบปัจจัยย้อนกลับโดยลดจำนวนนำเข้าที่ละปัจจัย และตรวจสอบค่าประสิทธิภาพความถูกต้องที่ลดลง ดังรูปที่ 8. ผลการทดลองพบว่าปัจจัยด้านอายุ มีค่าความถูกต้องลดลงน้อยที่สุด แสดงว่าอายุเป็นปัจจัยที่ส่งผลกระทบต่อโรคไฮเปอร์ไทรอยด์น้อยที่สุด หรืออาจจะมีผลกระทบต่อโรคไฮเปอร์ไทรอยด์สอดคล้องกับ

ข้อมูลผู้ป่วย [1] ซึ่งพบได้ว่าโรคไฮเปอร์ไทรอยด์สามารถเกิดได้กับประชาชนทุกช่วงอายุ

นอกจากอายุแล้วปัจจัยที่มีค่าความถูกต้องลดลงน้อยรองลงมาได้แก่ ปัจจัยลักษณะอาการมือสั่นและอาการนอนไม่หลับ ซึ่งแสดงว่า อายุ ลักษณะอาการมือสั่น การนอนไม่หลับ เป็นที่ส่งผลกระทบต่อโรคไฮเปอร์ไทรอยด์น้อยที่สุด หรืออาจจะมีผลกระทบต่อโรคไฮเปอร์ไทรอยด์

#### 4.6 ผลการพัฒนาระบบ

ระบบพยากรณ์ความเสี่ยงโรคไฮเปอร์ไทรอยด์พัฒนาระบบในรูปแบบแอปพลิเคชันบนสมาร์ตโฟนระบบปฏิบัติการแอนดรอยด์ การทำงานจะสามารถให้ผู้ใช้งานเข้าไปกรอกข้อมูลบนหน้าจอ โดยมีข้อคำถามทั้งหมด 12 ข้อ จากนั้นระบบจะประมวลผลและแสดงผลการวิเคราะห์ระดับความเสี่ยงแสดงตัวอย่างดังรูปที่ 9.



รูปที่ 9. หน้าจอการวิเคราะห์ข้อมูลและผลลัพธ์การพยากรณ์

การทำงานของระบบประกอบไปด้วยสองส่วนหลักคือ การทำแบบสอบถามเพื่อวิเคราะห์ข้อมูลระดับความเสี่ยงโรคไฮเปอร์ไทรอยด์ ตัวอย่างรูปที่ 9. และการให้ข้อมูลความรู้เกี่ยวกับโรคไฮเปอร์ไทรอยด์ ดังตัวอย่างรูปที่ 10.



รูปที่ 10. หน้าจอวิธีการรักษาโรคไฮเปอร์ไทรอยด์

#### 4.7 ผลการประเมินระบบ

การประเมินผลระบบใช้การประเมินความพึงพอใจด้วยมาตราอันดับ 5 อันดับ โดยมีการทดสอบกับผู้ใช้ทั่วไปด้วยวิธีการสุ่มแบบไม่เจาะจงจำนวน 30 คน ได้ผลดังตารางที่ 7 ซึ่งพบว่าความพึงพอใจของผู้ใช้โดยภาพรวมมีความพึงพอใจอยู่ในเกณฑ์ระดับดี (Mean=4.08, S.D. = 0.54) สามารถสรุปได้ว่าระบบนี้มีประสิทธิภาพสามารถนำไปประยุกต์ใช้งานได้จริง สามารถช่วยวิเคราะห์ระดับความเสี่ยงโรคไฮเปอร์ไทรอยด์ด้วยตนเองและแนะนำแนวทางการรักษาโรคไฮเปอร์ไทรอยด์ได้

ตารางที่ 7. ผลการประเมินระบบ

| รายการประเมิน                             | Mean | S.D. | การแปลผล |
|-------------------------------------------|------|------|----------|
| 1. ด้านการตรงความต้องการของผู้ใช้         | 4.03 | 0.48 | ดี       |
| 2. ด้านการทำงานได้ตามฟังก์ชันงานของระบบ   | 4.20 | 0.65 | ดี       |
| 3. ด้านความสะดวกและง่ายต่อการใช้งาน       | 4.17 | 0.58 | ดี       |
| 4. ด้านการรักษาความปลอดภัยของข้อมูลในระบบ | 3.93 | 0.44 | ดี       |
| ค่าเฉลี่ย                                 | 4.08 | 0.54 | ดี       |

#### 5. บทสรุป

งานวิจัยนี้เป็นการพัฒนาระบบสำหรับพยากรณ์ระดับความเสี่ยงของโรคไฮเปอร์ไทรอยด์ในรูปแบบของโมบายแอปพลิเคชันโดยประยุกต์ใช้เทคนิคเหมืองข้อมูล การดำเนินงานเป็นการศึกษาและวิเคราะห์ปัจจัยที่มีผลต่อการเกิดโรคไฮเปอร์ไทรอยด์เพื่อนำมาสร้างเป็นแบบจำลอง (Model) การพยากรณ์การเกิดโรคไฮเปอร์ไทรอยด์ ข้อมูลที่นำมาใช้เป็นข้อมูลที่เก็บมาจากแบบสอบถามผู้ป่วยโรคไฮเปอร์ไทรอยด์ จากโรงพยาบาลรังสีรักษาและโรงพยาบาลพิษณุเวช จังหวัดพิษณุโลก จำนวน 323 ชุดข้อมูล แบ่งเป็นข้อมูลปัจจัยการเกิดโรค 12 ปัจจัยหนึ่งคลาสผลลัพธ์ ซึ่งปัจจัยการเกิดโรคได้ผ่านการตรวจสอบความถูกต้องจากอายุรแพทย์ วัดประสิทธิภาพแบบจำลองด้วยการตรวจสอบไขว้แบบ  $k = 2-10$  ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง พบว่าแบบจำลอง

จากเทคนิคโครงข่ายประสาทเทียมมีค่าประสิทธิภาพดีที่สุด โดยมีค่าความถูกต้อง 82.97% ค่าความแม่นยำ 83.400% ค่าความระลึก 83.00% และค่าความถ่วงดุล 82.80%

ผลการค้นหาปัจจัยที่ส่งผลต่อโรคไฮเปอร์ไทรอยด์พบว่าปัจจัยลักษณะอาการที่มีความสำคัญคืออารมณ์แปรปรวนและเหนื่อยง่าย ส่วนปัจจัยส่วนบุคคลที่มีความสำคัญคือเพศ เนื่องจากพบผู้ป่วยเพศหญิงมากกว่าเพศชาย ปัจจัยที่มีผลต่อความเสี่ยงไฮเปอร์ไทรอยด์น้อยที่สุด 3 อันดับ ได้แก่ อายุ อาการมือสั่นและอาการนอนไม่หลับ ซึ่งสอดคล้องกับข้อมูลผู้ป่วย ซึ่งพบว่าโรคไฮเปอร์ไทรอยด์สามารถเกิดได้กับประชาชนทุกช่วงอายุ

จากนั้นนำแบบจำลองที่ได้ไปใช้ในการพัฒนาระบบสำหรับพยากรณ์ความเสี่ยงของโรคไฮเปอร์ไทรอยด์ เพื่อนำไปใช้ในการคัดกรองผู้ที่มีภาวะเสี่ยงต่อโรคไฮเปอร์ไทรอยด์ให้สะดวกและรวดเร็วยิ่งขึ้น ระบบนี้พัฒนาด้วยโปรแกรม Android Studio ร่วมกับภาษา JAVA และภาษา XML หลังจากพัฒนาแอปพลิเคชันเรียบร้อยแล้วได้นำระบบไปทดสอบความพึงพอใจจากผู้ใช้งานจำนวน 30 คน ผลการประเมินความพึงพอใจจากผู้ใช้งานพบว่าภาพรวมผู้ใช้มีความพึงพอใจระบบอยู่ในระดับดี (Mean=4.08, S.D. = 0.54) ดังนั้นสามารถนำระบบที่พัฒนาขึ้นนี้ไปใช้งานจริงเพื่อช่วยวิเคราะห์ระดับความเสี่ยงโรคไฮเปอร์ไทรอยด์ด้วยตนเองและแนะนำแนวทางการรักษาโรคไฮเปอร์ไทรอยด์ได้

การพัฒนางานวิจัยในอนาคต ควรใช้จำนวนข้อมูลที่มีมากกว่านี้และเพิ่มข้อมูลปัจจัยให้ครอบคลุมกับการวินิจฉัยทางการแพทย์เพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพมากขึ้นและสามารถนำไปใช้ประโยชน์ได้อย่างกว้างขวาง

#### เอกสารอ้างอิง

- [1] เทพ หิมะทองคำ, ไทรอยด์รู้สัญญาณเตือนแยกแยะและรับมือโรคของต่อมไทรอยด์อย่างเข้าใจ, กรุงเทพฯ: สำนักพิมพ์อมรินทร์สุขภาพ, 2553.
- [2] เอกสิทธิ์ พัทธวงศ์ศักดิ์, การวิเคราะห์ข้อมูลด้วยเทคนิคดาต้าไมน์นิ่งเบื้องต้น, ปทุมธานี: ดาต้าคิวบ์, 2557.
- [3] ภัทริรา สุวรรณโค, นิสาชล จันทน์ศรีและจิตมนต์ อังสกุล, “แบบจำลองการพยากรณ์ความเสี่ยงในการเกิดอุบัติเหตุทางถนนในเทศกาลปีใหมด้วยการทำเหมืองข้อมูล,”

- Journal of Information Science and Technology (JIST)*, Vol. 7, No. 2, pp.10-19, 2017.
- [4] สายชล สันสมบูรณ์ทอง, *การทำเหมืองข้อมูล Data Mining*, กรุงเทพฯ: จามจุรี โปรดักท์, 2558.
- [5] J.R. Quinlan, "Introduction of Decision Trees," *Machine Learning*, Vol. 1, pp.81-106, 1986.
- [6] Panomkorn, "ระบบปฏิบัติการ android," Dec, 2013.  
[Online]. Available:  
<https://panomkorn.wordpress.com/2013/07/26/ระบบปฏิบัติการ-android/>. [Accessed Dec. 15, 2018].
- [7] นงเยาว์ ในอรุณและพรณี สิทธิเดช, "การจำแนกผู้ป่วยโรคหัวใจขาดเลือดและผู้ป่วยโรคหัวใจรูปแบบอื่นโดยใช้เทคนิคเหมืองข้อมูล," บทความวิชาการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 7, 2554, หน้า 551-557.
- [8] กรวิภา ภูมบผา, ระวี ฉวีวงศ์ และวงกต ศรีอุไร, "ระบบวิเคราะห์ความเสี่ยงของการเกิดโรคอ้วนลงพุงและแนะนำอาหารสำหรับผู้ป่วย," บทความการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 8, 2555, หน้า 56-62.
- [9] นพรัตน์ พจน์จิราภรณ์, จิระศักดิ์ นำประดิษฐ์และศักดิ์ชาย ตั้งวรรณวิทย์, "ระบบสนับสนุนทางการแพทย์เพื่อวิเคราะห์โรคหลอดเลือดแดงโป่งพองโดยใช้เทคนิคเหมืองข้อมูลกรณีศึกษาการรักษาหลอดเลือดแดงโป่งพองผ่านสายสวน," บทความวิชาการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 11, 2558, หน้า 346-351.
- [10] ณัฐกร เมืองแก้ว, กาญจนา วิริยะพันธ์, ศักดา วัลลิภากร และพิมพ์ภา เตชะกมลสุข, "ระบบพยากรณ์ความรุนแรงและโอกาสเสียชีวิตในเด็กที่ได้รับบาดเจ็บหรืออุบัติเหตุโดยใช้โครงข่ายประสาทเทียม," บทความวิชาการประชุมวิชาการระดับชาติด้านคอมพิวเตอร์และเทคโนโลยีสารสนเทศ ครั้งที่ 9, 2556, หน้า 27-32.
- [11] สุพาพร บรรดาศักดิ์, เบญญาภา ศรีสว่างและสุภาวดี ทองคำ, "การคัดกรองสุขภาพสำหรับผู้ที่มีความเสี่ยงต่อการเป็นโรคข้อเข่าเสื่อมโดยใช้เหมืองข้อมูล," บทความวิชาการประชุมวิชาการระดับประเทศด้านเทคโนโลยีสารสนเทศ ครั้งที่ 8, 2559, หน้า 168-174.
- [12] พวงทอง วงอนันต์ และธีรพงษ์ สังข์ศรี, "การพัฒนาระบบการคัดกรองกลุ่มเสี่ยงโรคเบาหวานโดยใช้เทคนิคเหมืองข้อมูล," *The 5<sup>th</sup> ASEAN Undergraduate Conference in Computing*, 2017, หน้า 145-150.
- [13] ไวยวิทย์ พานิชอักษร, มหศักดิ์ เกตุน้ำ, "การพยากรณ์ยอดขายปลีกแก่สรณด้วยวิธีการวิเคราะห์อนุกรมเวลาและโครงข่ายประสาทเทียม," *Journal of Information Science and Technology (JIST)*, Vol. 7, No. 1, pp.42-49, 2017.
- [14] บรรจบ คลกฤ, จารี ทองคำและวาทีณี สุขมาก, "การสร้างแบบจำลองเพื่อพยากรณ์การเกิดแผลที่เท้าของผู้ป่วยโรคเบาหวานโดยใช้เทคนิคเหมืองข้อมูล," *Journal of Science and Technology Mahasarakham University*, Vol. 33, No. 6, pp. 703-710, 2014.