

การจำแนกข้อมูลด้วยวิธีแบบร่วมกันตัดสินใจจากพื้นฐานของเทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ร่วมกับการเลือกตัวแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม **Ensemble Data Classification Based on Decision Tree, Artificial Neuron Network and Support Vector Machine Optimized by Genetic Algorithm**

เดช ธรรมศิริ^{1*} และ พยุง มีสัจ²
Dech Thamsiri^{1*} and Phayung Meesad²

บทคัดย่อ

งานวิจัยครั้งนี้ เป็นการนำเสนอวิธีการจำแนกข้อมูลด้วยวิธีร่วมกันตัดสินใจแบบการโหวตเสียงข้างมากจากเทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน โดยเลือกตัวแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรมเพื่อให้ความถูกต้องในการจำแนกข้อมูลที่สูงขึ้น จากการทดสอบความถูกต้องในการจำแนกข้อมูล Australian Credit, German Credit และ Bankruptcy Data พบว่า เทคนิคการร่วมกันตัดสินใจจากหลายโมเดลที่ผ่านการเลือกแทนที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม มีผลลัพธ์ที่ดีกว่าการใช้เทคนิคแบบโมเดลเดี่ยว โดยผลการวิจัยพบว่า สำหรับข้อมูล Australian Credit หลังจากใช้วิธีขั้นตอนวิธีเชิงพันธุกรรมช่วยในการเลือกตัวจำแนกข้อมูลพบว่าการเลือกใช้ตัวจำแนกที่ให้ประสิทธิภาพสูงสุดได้แก่ การใช้ตัวจำแนกข้อมูล 9 โมเดลร่วมกัน ประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 4 โมเดล เทคนิค

ต้นไม้ตัดสินใจ 3 โมเดล และเทคนิคโครงข่ายประสาทเทียม 2 โมเดล ให้ผลลัพธ์ความถูกต้องสูงสุดที่ 89.01% ข้อมูล German Credit หลังจากใช้วิธีขั้นตอนวิธีเชิงพันธุกรรมช่วยในการเลือกตัวจำแนกข้อมูลพบว่าการเลือกใช้ตัวจำแนกที่ให้ประสิทธิภาพสูงสุดได้แก่ การใช้ตัวจำแนกข้อมูล 18 โมเดล ซึ่งประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 7 โมเดลร่วมกัน ประกอบด้วย เทคนิคต้นไม้ตัดสินใจ 6 โมเดล และเทคนิคโครงข่ายประสาทเทียม 5 โมเดล ให้ผลลัพธ์ความถูกต้องสูงสุดที่ 71.50% ข้อมูล Bankruptcy Data หลังจากใช้วิธีขั้นตอนวิธีเชิงพันธุกรรมช่วยในการเลือกตัวจำแนกข้อมูลพบว่าการเลือกใช้ตัวจำแนกที่ให้ประสิทธิภาพสูงสุดได้แก่ การใช้ตัวจำแนกข้อมูล 10 โมเดลร่วมกัน ประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 4 โมเดล เทคนิคต้นไม้ตัดสินใจ 4 โมเดล และเทคนิคโครงข่ายประสาทเทียม 2 โมเดล ให้ผลลัพธ์ความถูกต้องสูงสุดที่ 71.50%

¹ นักศึกษา ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

² ผู้ช่วยศาสตราจารย์ ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

* Corresponding Author, Tel. 08-6775-6065, E-mail : dechit@msn.com

คำสำคัญ: เทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคการร่วมกันตัดสินใจ ขั้นตอนวิธีเชิงพันธุกรรม

Abstract

In this research we propose an ensemble classification technique based on majority vote from decision tree, artificial neural network, and support vector machine models selected by genetic algorithm in order to increase classification accuracy. From classification accuracy test on Australian Credit, German Credit and Bankruptcy Data, we found that the proposed ensemble classification models selected by genetic algorithm yields better performance than that of single model. The results show as follows. For Australian Credit after using genetic algorithm find the best classifier namely using 9 classifier models including 4 support vector models, 3 decision tree modes and 2 artificial neural network models achieved best performance with 89.01% accuracy on classification. For German Credit data after using genetic algorithm find the best classifier namely using 18 classifier models including 7 support vector machine models, 6 decision tree models, and 5 artificial neural network models achieved best performance with 71.50% accuracy on classification. For Bankruptcy Data after using genetic algorithm find the best classifier namely using 10 classifier models composed of 4 support vector machine models, 4 decision tree models, and 2 artificial neural network models achieved best performance with 71.50%.

Keyword: Decision Tree, Artificial Neuron Network, Support Vector Machine, Ensemble, Genetic Algorithm

1. บทนำ

ปัจจุบันนักวิจัยจำนวนมากได้ให้ความสำคัญกับปัญหาที่ต้องการวิธีการจำแนกข้อมูลที่แม่นยำ โดยมีการนำเสนอวิธีการที่หลากหลาย เช่น การนำเอาเทคนิคด้านปัญญาประดิษฐ์ (Artificial Intelligence) มาประยุกต์ใช้ได้แก่ เทคนิคต้นไม้ช่วยตัดสินใจ (Decision Tree Models) [1] เทคนิคโครงข่ายประสาทเทียม (Artificial Neural Networks หรือ ANN) [2] เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine หรือ SVM) [3] เป็นต้น

แต่ถึงแม้ว่าเทคนิคข้างต้นจะให้ผลลัพธ์ในการจำแนกข้อมูลที่แม่นยำแต่การเป็นโมเดลเดี่ยว (Single Model) นั้นทำให้มีการกำหนดกลุ่มของข้อมูลที่ใช้ในการเรียนรู้ รวมทั้งมีการกำหนดค่าพารามิเตอร์ที่ตายตัว บางครั้งเกิดปัญหาความเอนเอียง (Bias) ทำให้ได้ประสิทธิภาพที่ไม่ดีนัก หนทางหนึ่งที่จะสามารถลดค่าความเอนเอียงได้คือการใช้วิธีการร่วมกันตัดสินใจ (Ensemble) ซึ่งสามารถสร้างความหลากหลายและลดค่าความผิดพลาดที่เกิดจากความแปรปรวนได้ แนวความคิดหลักของวิธีการร่วมกันตัดสินใจคือการรวมเอากลุ่มของตัวจำแนกข้อมูล เพื่อแก้ปัญหาเดียวกันและผลลัพธ์ที่ได้จะมีความแม่นยำและความเที่ยงตรงมากกว่าการใช้โมเดลแบบเดี่ยว หากแต่วิธีการรวมกลุ่มถ้าจะให้ประสิทธิภาพที่ดี ได้มีผู้วิจัยดังเช่น Wang และคณะ [4] ให้คำแนะนำว่า “โดยทั่วไปแล้วประสิทธิภาพของการรวมกลุ่มจะขึ้นอยู่กับความหลากหลาย และความแม่นยำของตัวแทนในการจำแนกข้อมูล”

ดังเช่นงานวิจัยของ Yao และคณะ [5] ได้นำเสนอการรวมกลุ่มช่วยในการตัดสินใจโดยใช้เทคนิคต้นไม้ช่วยตัดสินใจ ร่วมกับวิธีการ Boosting โดยได้ทดสอบประสิทธิภาพกับข้อมูลจาก UCI ผลที่ได้แสดงให้เห็นว่าวิธีการดังกล่าวให้ประสิทธิภาพที่ดีมากกว่าโมเดลแบบเดี่ยว Wang และคณะ [6] ได้ใช้เทคนิคการรวมกลุ่มตัดสินใจ ร่วมกับเทคนิคโครงข่ายประสาทเทียม เข้ามาใช้ในการแก้ปัญหาการวิเคราะห์ตัวเลขที่เขียนด้วยลายมือ ซึ่งผู้วิจัยพบว่าปัญหาทั้งในเรื่องของขนาด ความกว้าง ข้อมูลรบกวน และตำแหน่งที่แตกต่างกัน ผู้วิจัยดังกล่าวได้นำเสนอ

การใช้เทคนิคโครงข่ายประสาทเทียมแบบ Hopfield Neural Network เพื่อใช้ในการกำจัดข้อมูลรบกวน และใช้เทคนิค Bagging ทำการรวมกลุ่มตัวจำแนกข้อมูลที่สร้างขึ้นมาจากโครงข่ายประสาทเทียมแบบ Multilayer Feed Forward ผลที่ได้เมื่อเปรียบเทียบกับโมเดลแบบเดี่ยวทั้งในส่วนข้อมูลสอนและข้อมูลทดสอบพบว่า วิธีการที่ได้นำเสนอให้ผลลัพธ์ที่แม่นยำกว่า Chen และคณะ [7] ได้นำเสนอการใช้ตัวจำแนกข้อมูลด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีนที่มีการกำหนดค่าเคอร์เนลฟังก์ชัน รวมทั้งค่าพารามิเตอร์ที่หลากหลาย สำหรับสร้างตัวจำแนกข้อมูลผลที่ได้มีประสิทธิภาพที่ดีกว่าการใช้โมเดลเทคนิคซัพพอร์ตเวกเตอร์แมชชีนแบบเดี่ยว จะเห็นได้ว่าเทคนิคการรวมกลุ่มสามารถช่วยเพิ่มประสิทธิภาพให้กับโมเดล

ในส่วนของ การสร้างตัวแทนการเรียนรู้ที่หลากหลาย ได้มีผู้วิจัยหลายท่านได้นำเสนอไว้ดังนี้ Ying และคณะ [8] ได้ใช้วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีโครงข่ายประสาทเทียม และวิธีต้นไม้ช่วยตัดสินใจร่วมกันจำแนกข้อมูลอีเมลขยะ Pei-Yong และคณะ [9] ได้ใช้ KNN และ Bayes เป็นตัวจำแนกข้อมูลพร้อมด้วยการใช้วิธีการเรียนรู้แบบรวมกลุ่มและใช้วิธีเชิงพันธุกรรมช่วยเลือกคุณลักษณะที่เหมาะสม Zhou และคณะ [10] Chen และคณะ [7] ได้นำเสนอการใช้ตัวจำแนกข้อมูลด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนที่มีการกำหนดค่าเคอร์เนลฟังก์ชัน รวมทั้งค่าพารามิเตอร์ที่หลากหลาย ดังนั้นจะเห็นได้ว่าการสร้างตัวแทนในการเรียนรู้ที่หลากหลาย ทำให้ตัวแทนที่สร้างขึ้นสามารถนำไปแก้ปัญหาที่หลากหลายได้เช่นเดียวกัน อีกทั้งยังเป็นการเพิ่มประสิทธิภาพโดยรวมของโมเดลอีกด้วย

สำหรับการเลือกตัวจำแนกข้อมูลที่เหมาะสมเพื่อนำมาใช้สำหรับการตัดสินใจร่วมกัน ได้มีผู้วิจัยหลายท่านได้นำเสนองานวิจัยดังเช่น Martinez-Muoz และคณะ [11] ได้นำเสนอวิธีการ Pruning Ensemble โดยใช้วิธี Reduce-Error Pruning, Kappa Pruning เพื่อทำการ Pruning ตัวจำแนกข้อมูลที่มีประสิทธิภาพไม่ติดบางส่วนออก Liu และคณะ [12] ได้นำเสนอการใช้วิธี Cluster เพื่อจับกลุ่มตัวจำแนกข้อมูล และทำการเลือกกลุ่มตัวจำแนกข้อมูลที่มีประสิทธิภาพมากที่สุด เพื่อนำมาใช้

ร่วมกันตัดสินใจ Ning และคณะ [13] ได้นำเสนอวิธีการ Virus-Evolutionary Genetic Algorithm สำหรับการเลือกตัวจำแนกข้อมูลที่เหมาะสม เพื่อนำมาใช้ร่วมกับวิธีการเรียนรู้แบบรวมกลุ่ม ในงานวิจัยข้างต้นแสดงให้เห็นถึงความพยายามในการหาวิธีเลือกตัวจำแนกข้อมูลที่เหมาะสม จากตัวจำแนกข้อมูลทั้งหมด และพบว่าวิธีการดังกล่าวเป็นการช่วยเพิ่มประสิทธิภาพ อีกทั้งยังเป็นการลดค่าใช้จ่ายในการคำนวณอีกด้วย

ดังนั้นสำหรับงานวิจัยในครั้ง นี้ ผู้วิจัยได้ให้ความสำคัญด้านการแก้ปัญหาต่างๆ เหล่านี้ เพื่อหาวิธีการที่เหมาะสม สำหรับการเรียนรู้แบบรวมกลุ่ม โดยการสร้างตัวจำแนกข้อมูลจากหลากหลายเทคนิค การเลือกพารามิเตอร์ที่เหมาะสมเพื่อเพิ่มประสิทธิภาพของตัวจำแนกข้อมูลแบบร่วมกันตัดสินใจ พร้อมการเลือกตัวจำแนกข้อมูลที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม สำหรับนำไปสร้างโมเดลใหม่ที่มีประสิทธิภาพด้านการจำแนกข้อมูลได้ถูกต้องมากขึ้น

2. ทฤษฎีที่เกี่ยวข้อง

2.1 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน

เป็นเครื่องมือสำหรับการเรียนรู้แบบใหม่บนพื้นฐานของการเรียนรู้ทฤษฎีทางสถิติได้รับการนำเสนอครั้งแรกโดย Vapnik [14] โดยมีแนวคิดของทฤษฎีดังนี้

กำหนดข้อมูลฝึกสอน $D = \{(x_i, y_i) ; i=1, 2, \dots, n\}$ ในขณะที่ $x_i = (x_{i1}, x_{i2}, \dots, x_{ip}) \in \mathcal{R}^p$ นั้นเป็นข้อมูลนำเข้า และ $y_i \in \{+1, -1\}$ (+1 คือ “ข้อมูล Class 1”, -1 คือ “ข้อมูล Class 0”) เป็นค่าเป้าหมายของ x_i โดยที่ SVM เป็นฟังก์ชันการตัดสินใจที่สามารถแยกแยะข้อมูล x ที่ไม่ทราบได้โดยใช้สมการที่ 1

$$y = \text{sign} \left\{ \sum_{j=1}^{n_s} w_j \phi_j(x) \phi_j(x_j) + b \right\} \quad (1)$$

$$\phi(x) = [\phi_1(x_1), \phi_2(x_2), \dots, \phi_n(x_{n_s})]^T \quad (2)$$

จากตัวอย่างสมการที่ 2 กลุ่มข้อมูล x ซึ่งไม่สามารถแบ่งแยกได้ด้วยสมการเส้นตรงจะถูกแปลงให้

อยู่ในรูปแบบที่สามารถใช้สมการเส้นตรงแบ่งแยกได้ โดยใช้เคอร์เนลฟังก์ชัน $K(\mathbf{x}, \mathbf{x}_j) = \phi(\mathbf{x})\phi(\mathbf{x}_j)$ w_j แทนค่าน้ำหนัก (Weight) ที่เชื่อมโยงจาก Feature Space ในมิติ j ไปสู่ Output Space b เป็นค่าโน้มเอียง (Bias หรือ Threshold) และ $\mathbf{x}_j, j = 1 \dots n_s$ เป็นซัพพอร์ตเวกเตอร์ เมื่อ n_s เป็นจำนวนซัพพอร์ตเวกเตอร์ และทำให้ให้สมการที่ (3) สำหรับการจำแนกข้อมูล

$$f(\mathbf{x}) = \text{sign} \left\{ \sum_{j=1}^{n_s} w_j \phi_j(\mathbf{x}) \phi_j(\mathbf{x}_j) + b \right\} \quad (3)$$

หากแต่บางครั้งไม่สามารถแยกแยะข้อมูลได้ถูกต้องทั้งหมด ทำให้ต้องมีการกำหนดตัวแปรเพื่อยอมรับค่าความผิดพลาด โดยทำการเพิ่มตัวแปร ξ (Slack Variable) โดยเขียนแสดงรายละเอียดได้ดังสมการที่ (4) และ (5)

$$\mathbf{w}^T \mathbf{x} + b \geq +1 - \xi_i \quad ; \quad y_i = +1 \quad (4)$$

$$\mathbf{w}^T \mathbf{x} + b \leq -1 + \xi_i \quad ; \quad y_i = -1 \quad (5)$$

โดยที่ $\xi_i > 0$

ทำให้ได้โครงสร้างของ SVM ซึ่งประกอบด้วย 2 ส่วนหลัก คือการเพิ่มระยะแยกแยะมากที่สุด และการแก้ปัญหาด้วยการลดข้อผิดพลาดให้ต่ำที่สุด ดังแสดงในสมการที่ (6)

$$\text{Minimize}_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (6)$$

โดยที่ $y_i (\mathbf{w}^T \phi(\mathbf{x}) + b) + \xi_i - 1 \geq 0$

$$\xi_i \geq 0, i = 1, 2, \dots, N$$

และมีเคอร์เนลฟังก์ชัน (Kernel Function) ที่พบได้บ่อย 3 แบบด้วยกัน ดังแสดงในสมการที่ 7, 8 และ 9 ได้แก่

Polynomial Kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (1 + \mathbf{x}_i \cdot \mathbf{x}_j)^d \quad (7)$$

Radial Basis Function Kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (8)$$

Sigmoid Kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(k\mathbf{x}_i \cdot \mathbf{x}_j) - \delta \quad (9)$$

2.2 เทคนิคโครงข่ายประสาทเทียม [15]

มีพื้นฐานมาจากการจำลองการทำงานของสมองมนุษย์ ด้วยโปรแกรมคอมพิวเตอร์ จุดมุ่งหมายของโครงข่ายประสาทเทียมคือต้องการให้คอมพิวเตอร์มีความชาญฉลาดในการเรียนรู้เหมือนที่มนุษย์มีการเรียนรู้ สามารถฝึกฝนได้ และสามารถนำความรู้และทักษะ รวมทั้งสามารถนำไปประยุกต์ใช้ได้ดีกับปัญหา Classification, Regression และ Clustering เทคนิคนี้มักถูกเรียกว่า “Black Box” เนื่องจากการทำงานมีความซับซ้อนมากกว่าเทคนิคอื่นๆ ค่อนข้างมาก

การเรียนรู้ของนิวรอนเน็ตเวิร์ก ทำได้โดยการส่งข้อมูลเข้ามายังส่วนที่เรียกว่าเพอร์เซ็ปตรอน (Perceptron) สามารถเทียบได้กับเซลล์สมองของมนุษย์ โดยที่เพอร์เซ็ปตรอนทำการรับข้อมูลที่อยู่ในรูปของแมทริกซ์ซึ่งเป็นตัวเลขเข้ามาคำนวณ สำหรับฟังก์ชันผลรวม (Summation Function) มีการทำงานดังสมการที่ (10)

$$n = \sum_{i=1}^z x_i w_i + b \quad (10)$$

โดยที่

ตัวแปร n คือ ผลรวมที่ได้จากฟังก์ชันผลรวม

ตัวแปร x_i คือ ค่าข้อมูลเข้าตัวที่ i

ตัวแปร w_i คือ ค่าน้ำหนักของนิวรอนตัวที่ i

ตัวแปร z คือ จำนวนนิวรอนชั้นข้อมูลเข้า

ตัวแปร b คือ ค่าความโน้มเอียง

ตัวแปร i มีค่าตั้งแต่ 1 ถึง z

ขั้นตอนการเรียนรู้แบบเพอร์เซ็ปตรอนมีขั้นตอนดังนี้ [16]

1. เริ่มจากการสุ่มค่าน้ำหนัก w_i

2. เทียบเพอร์เซ็ปตรอนกับข้อมูลที่นำมาสอนทำการปรับแก้ค่าน้ำหนัก เมื่อเพอร์เซ็ปตรอนแยกตัวอย่างผิดพลาด

3. ทำการวนซ้ำข้อมูลที่ใช้สำหรับการสอนจนกระทั่งเพอร์เซ็ปตรอนแยกตัวได้อย่างถูกต้องทั้งหมด

4. ในการแก้ไขน้ำหนัก น้ำหนักถูกปรับตาม $W_i \leftarrow W_i + \Delta W_i$

โดย $\Delta W_i = \alpha \times (t - a) \times x_i = \alpha \times error \times input$ เมื่อ t เป็นผลลัพธ์ที่ต้องการ a เป็นผลลัพธ์ที่ได้จากเพอร์เซ็ปตรอน และ α เป็นค่าที่แสดงอัตราการเรียนรู้

2.3 เทคนิคต้นไม้ช่วยตัดสินใจ

เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปของต้นไม้ช่วยตัดสินใจ ซึ่งต้นไม้ช่วยตัดสินใจนั้นทำงานแบบ Supervised Learning คือ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างของข้อมูลที่ได้กำหนดได้ก่อนล่วงหน้า ที่เรียกว่า Training Set ได้อัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วย

วิธีการที่ใช้สร้างต้นไม้ช่วยตัดสินใจมีขั้นตอนดังนี้

1. หา Attribute ที่สำคัญที่สุดมาแบ่งข้อมูลโดย Attribute นี้ถูกนำมาสร้างเป็น Root Node โดยมี Target Attribute เป็นผลลัพธ์ซึ่งเป็น Leaf Node ถูกกำหนดไว้ก่อน

2. นำค่าที่เป็นไปได้ใน Attribute ที่ถูกเลือกมาแตกออกเป็นกลุ่มของตัวเอง

3. แบ่งข้อมูลทั้งหมดตามกลุ่มที่แตกออกจาก Root Node

4. วนกลับไปทำที่ขั้นตอนแรก คือหา Attribute ที่สำคัญที่สุดจากข้อมูลที่เข้ามาเพื่อหาตัวแบ่งต่อไป

2.4 เทคนิคการเรียนรู้แบบรวมกลุ่ม

เป็นเทคนิคในการรวมเอากลุ่มของตัวจำแนกข้อมูลที่หลากหลาย เพื่อแก้ปัญหาเดียวกัน ด้วยวิธีการการนำเอาตัวจำแนกข้อมูลเหล่านั้นมาร่วมกันตัดสินใจ สำหรับงานวิจัยฉบับนี้เลือกใช้วิธีการโหวตเสียงข้างมาก (Majority Vote) ซึ่งวิธีการนี้ง่ายต่อการใช้งาน มีวิธีการคือกำหนดค่า f_k โดยที่ $(k = 1, 2, \dots, K)$ ซึ่ง K คือจำนวนของตัวจำแนกข้อมูลและ $f_k(x) = c, c \in \{+1, -1\}$ โดยที่ฟังก์ชันในการตัดสินใจจะตัดสินใจผลลัพธ์ที่ได้ว่าอยู่ในคลาส +1 หรือ คลาส -1 สุดท้ายนำค่าฟังก์ชัน $f_c(x)$

ที่ได้จากตัวจำแนกข้อมูลมาตัดสินใจร่วมกัน โดยดูที่เสียงข้างมากว่า ตัดสินใจเป็นคลาส +1 หรือคลาส -1 โดยมีสมการดังสมการที่ (11)

$$f_c(x) = \arg \max_c \sum_{k: f_k(x)=c} 1 \quad (11)$$

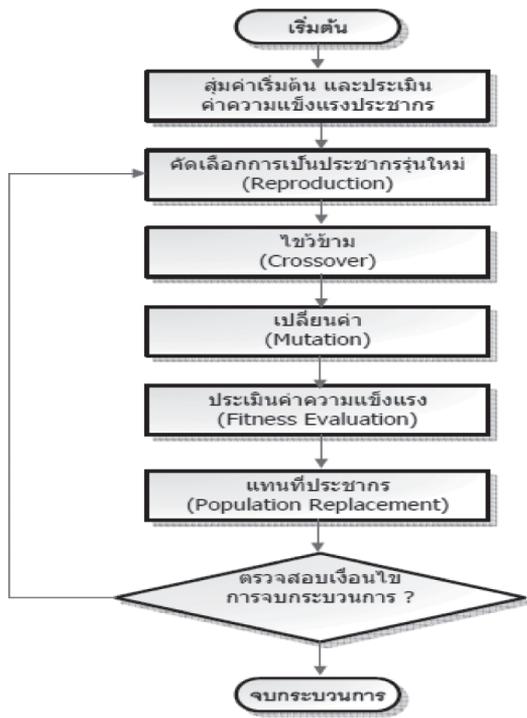
2.5 วิธีเชิงพันธุกรรม (Genetic Algorithm)

เป็นเทคนิคสำหรับค้นหาผลเฉลย (Solutions) หรือคำตอบโดยประมาณของปัญหา โดยอาศัยหลักการจากทฤษฎีวิวัฒนาการจากชีววิทยา และการคัดเลือกตามธรรมชาติ (Natural Selection) นั่นคือสิ่งมีชีวิตที่เหมาะสมที่สุดจึงจะอยู่รอด ด้วยตัวปฏิบัติการทางพันธุกรรม (Genetic Operator) [17] เช่น การถ่ายทอดพันธุกรรม (Inheritance) การคัดเลือกคู่ (Selection) การกลายพันธุ์ (Mutation) และการสลับสายพันธุ์ (Crossover) เป็นต้น

เนื่องจากการหาคำตอบที่เหมาะสม ดังนั้นคำตอบที่ได้จึงไม่ใช่คำตอบที่ดีที่สุดเสมอไป โดยทั่วไปแล้วขั้นตอนวิธีเชิงพันธุกรรมอย่างง่าย [18] ข้างต้นได้ถูกนำมาขยายความเป็นกระบวนการการทำงานของโปรแกรมเพื่อที่จะวิเคราะห์หาคำตอบโดยการกำหนดค่าเริ่มต้นของปัญหาขึ้นมาซึ่งเรียกว่า สายพันธุ์เริ่มต้น จากนั้นผ่านกระบวนการหาคำตอบจากกระบวนการ ทั้งสามข้างต้น จะได้สายพันธุ์ที่สอง จากนั้นตรวจสอบว่าคำตอบที่ได้เป็นคำตอบที่พอใจหรือยัง หากยังไม่พอใจก็ให้วนไปยังสามกระบวนการข้างต้นอีกครั้ง เกิดเป็นสายพันธุ์ที่สาม ที่สี่ และต่อไปเรื่อยๆ จนกว่าจะได้คำตอบที่พอใจ ขั้นตอนกระบวนการทำงานดังรูปที่ 1

3. วิธีการและโมเดลที่นำเสนอ

ในงานวิจัยนี้เริ่มต้นที่การนำข้อมูลมาทำการนอร์มัลไลซ์ (Normalization) โดยปรับค่าข้อมูลให้อยู่ในรูปแบบเดียวกัน จากนั้นได้นำเสนอวิธีการสร้างตัวจำแนกข้อมูลที่หลากหลายโดยใช้วิธีการสุ่มกลุ่มตัวอย่าง เพื่อสร้างความแตกต่างบนกลุ่มข้อมูลสำหรับการฝึกฝน (Training Sets) ซึ่งประกอบด้วย กลุ่มตัวอย่างจำนวน N



รูปที่ 1 ขั้นตอนกระบวนการวิธีเชิงพันธุกรรม

ตามจำนวนของตัวจำแนกข้อมูลที่ได้กำหนดขึ้น สำหรับงานวิจัยนี้ได้ใช้ตัวจำแนกทั้งสิ้น 30 ตัวจำแนกโดยแยกเป็นการสร้างตัวจำแนกข้อมูลด้วยเทคนิคต้นไม้ช่วยตัดสินใจ 10 โมเดล เทคนิคโครงข่ายประสาทเทียม 10 โมเดล เทคนิคซัพพอร์ตเวกเตอร์แมชชีน 10 โมเดล โดยที่ข้อมูลถูกสุ่มและสร้างมาจากกลุ่มข้อมูลสำหรับการสอน โดยที่กลุ่มข้อมูลที่สุ่มสร้างขึ้นมาถูกเรียกว่า Bootstrap [19] สำหรับการสุ่มจำนวนกลุ่มตัวอย่าง จากกลุ่มข้อมูลสำหรับการสอน มีค่าความน่าจะเป็น ดังสมการที่ (12)

$$n_i = 1 - \left(\frac{1 - 1}{m}\right)^m \quad (12)$$

โดยที่

n_i คือ กลุ่มข้อมูลที่ถูกรวมสร้างขึ้นมา

m คือ จำนวนข้อมูลตัวอย่างทั้งหมด ที่อยู่ในกลุ่ม

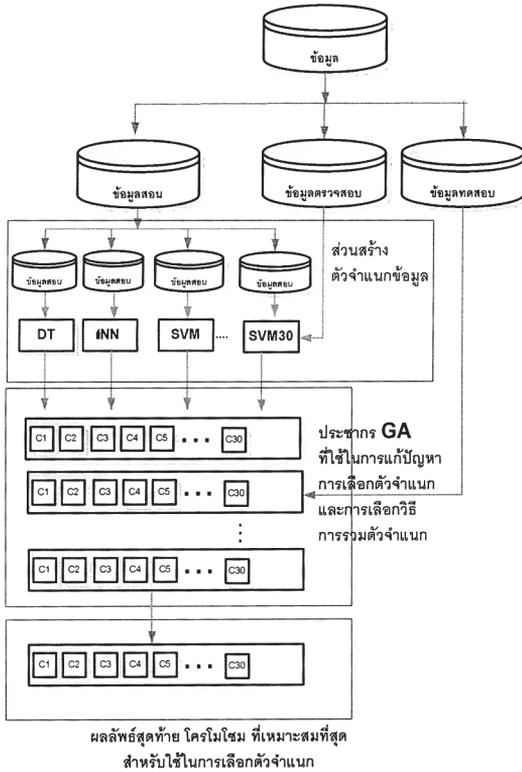
ข้อมูลสำหรับการฝึกฝน

โดยที่กลุ่มข้อมูล n_i จะถูกนำไปสอนให้กับตัวจำแนกข้อมูล ซึ่งได้แก่การใช้เทคนิคต้นไม้ช่วยตัดสินใจ เทคนิคโครงข่ายประสาทเทียมและเทคนิคซัพพอร์ตเวกเตอร์แมชชีน สำหรับเทคนิคโครงข่ายประสาทเทียม ผู้วิจัยได้กำหนดเลือกใช้โครงสร้างที่แตกต่างกัน 6 โครงสร้าง สำหรับการเลือกใช้จะทำการสุ่มเพื่อเลือกโครงสร้างที่จะนำมาสร้างโมเดล ถึงแม้ว่าจะมีการสุ่มเลือกโครงสร้างที่ซ้ำกัน แต่เนื่องจากข้อมูลที่ใช้ในการสอนโมเดลนั้นจะยังคงแตกต่างกันอยู่ ทำให้ได้รูปแบบของโมเดลที่แตกต่างกัน ดังตารางที่ 1

ตารางที่ 1 รายละเอียดโครงสร้างโครงข่ายประสาทเทียม

โครงสร้างที่	จำนวนนิวรอนชั้นซ่อน	รูปแบบฟังก์ชันกระตุ้น	
		ชั้นซ่อน	ชั้นผลลัพธ์
1	30-20-1	'logsig' 'tansig' 'purelin'	'traincgb'
2	3-1	'tansig' 'purelin'	'traingdm'
3	10-50-1	'tansig' 'tansig' 'purelin'	'trainrp'
4	8-10-1	'purelin' 'purelin' 'tansig'	'trainbr'
5	8-20-1	'tansig' 'tansig' 'tansig'	'trainbfg'
6	12-10-1	'tansig' 'purelin' 'logsig'	'trainrp'

นอกจากนั้นทำการกำหนดจำนวนรอบการเรียนรู้ไว้ที่ 500 รอบ กำหนดเป้าหมายการเรียนรู้ไว้ที่ 0.0001 และสำหรับเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ผู้วิจัยได้เลือกใช้เคอร์เนล Radial Basis Function เนื่องจากผู้วิจัยได้ทำการทดลองแล้วให้ผลที่ดีกับข้อมูลที่ผู้วิจัยได้นำมาใช้ สำหรับการกำหนดค่าพารามิเตอร์ให้กับโมเดล ผู้วิจัยเลือกใช้วิธีการค้นหาแบบกริช (Grid Search) เพื่อช่วยในการกำหนดค่าพารามิเตอร์ที่เหมาะสมโดยกำหนดค่า C ซึ่งเป็นค่าถ่วงน้ำหนักสำหรับการให้ความสำคัญระหว่างระยะแยกแยะและค่าความผิดพลาดไว้ที่ [1, 3,



รูปที่ 2 ขั้นตอนการสร้างตัวจำแนกข้อมูล

5, 10, 15, 20, 25, 30] ส่วนค่าแกมมาเป็นพารามิเตอร์สำหรับเคอร์เนลฟังก์ชัน RBF ที่เลือกใช้ได้กำหนดค่าไว้ที่ [0.0001, 0.001, 0.01, 0.1, 0.5, 1] ซึ่งโมเดลที่สร้างขึ้นมาโดยใช้การค้นหาแบบกริชนั้นทำให้ได้โมเดลทั้งสิ้น 48 โมเดลจากการกำหนดค่าพารามิเตอร์ที่แตกต่างกัน หากแต่ในการเลือกใช้งานผู้วิจัยจะเลือกใช้โมเดลที่นำไปเป็นตัวแทนการเรียนรู้เพียงแค่ 1 โมเดล โดยเลือกโมเดลที่มีประสิทธิภาพสูงสุดเท่านั้น โดยมีขั้นตอนการสร้างตัวจำแนกข้อมูลดังรูปที่ 2

3.1 ขั้นตอนการทำงานของโมเดล

วิธีการเข้ารหัสโครโมโซม กำหนดให้ตัวจำแนกทั้งหมด N ตัวในกลุ่มของตัวจำแนกแทนด้วยตัวเลขตั้งแต่ 1 ถึง N โดยที่สายของโครโมโซมจะประกอบไปด้วยตัวเลขแบบไบนารี จำนวน N ตัว โดยตำแหน่งของ



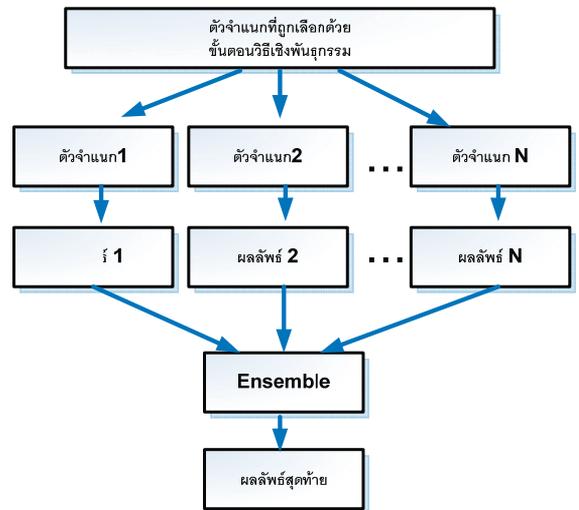
รูปที่ 3 ข้อมูลในโครโมโซม



รูปที่ 4 ตำแหน่งที่ในโครโมโซม



รูปที่ 5 ตำแหน่งตัวจำแนกข้อมูลที่ถูกเลือกมาใช้งาน



รูปที่ 6 โมเดลการเรียนรู้แบบรวมกลุ่ม (Ensemble)

เลขไบนารีจะถูกกำหนดลำดับที่ i เลขไบนารีจะมีค่า 2 ค่า ได้แก่ 0 ซึ่งมีความหมายว่าตัวจำแนกนั้นจะไม่ถูกเลือก และ 1 ซึ่งมีความหมายว่าตัวจำแนกนั้นจะถูกเลือกใช้ เพื่อนำมารวมกันตัดสินใจ ตัวอย่างเช่น โครโมโซม = 1100010001 (เมื่อ $N = 10$) มีความหมายว่าตัวจำแนกที่ 1, 2, 6 และ 10 ถูกเลือกมาใช้งาน ดังรูปที่ 3-5 จากนั้นนำตัวจำแนกข้อมูลที่ถูกเลือกในขั้นตอนก่อนหน้า มาทำการรวมกลุ่มกันเพื่อช่วยในการตัดสินใจ (Ensemble) การตัดสินใจเทคนิคการโหวตเสียงข้างมาก (Majority Vote) ผลลัพธ์ได้มาจากการวัดประสิทธิภาพความแม่นยำในการจำแนกข้อมูล (Accuracy) ดังรูปที่ 6

สำหรับขั้นตอนเชิงพันธุกรรมได้ทำการกำหนดโครโมโซมเริ่มต้นที่ 30 โครโมโซม กำหนดจำนวนรุ่น (Generation) ไว้ที่ 50 รุ่น การไขว้ข้าม (Cross Over) ใช้การไขว้ข้ามแบบสมำเสมอ กำหนดอัตราไว้ที่ 0.6 การเปลี่ยนค่า (Mutation) ใช้วิธีการแบบสุ่ม 1 จุด กำหนดอัตราไว้ที่ .0 การประเมินค่าความแข็งแรง (Fitness Evaluation) ใช้ค่า Mean Absolute Percent Error (MAPE) ซึ่งเป็นผลลัพธ์ที่ได้จากการรวมกันของตัวจำแนกข้อมูลจากนั้นคัดเลือกเป็นประชากรรุ่นใหม่ ใช้วิธีการคัดเลือกแบบวงล้อรูเล็ต (Roulette Wheel Selection) ร่วมกับวิธีการคัดเลือกชั้นยอด (Elitism Selection) กำหนดการคัดเลือกชั้นยอด (Elitism Selection) ไว้ที่ 5 โครโมโซม และการคัดเลือกแบบวงล้อรูเล็ต (Roulette Wheel Selection) ที่ 25 โครโมโซม สุดท้ายเมื่อผ่านกระบวนการวิวิขั้นตอนเชิงพันธุกรรม ประชากรในรุ่นสุดท้ายจะมีค่าตอบที่แสดงให้เห็นว่าจะนำตัวจำแนกข้อมูลตัวใดบ้างมารวมกันแล้วให้ผลที่ดีที่สุด

3.2 ข้อมูลและเครื่องมือที่ใช้ในการวิจัย

การวัดประสิทธิภาพของโมเดล งานวิจัยครั้งนี้ได้ใช้กลุ่มข้อมูล ทั้งสิ้น 2 กลุ่มข้อมูลซึ่งประกอบด้วยข้อมูลจาก UCI ซึ่งเป็นข้อมูลเปิดสำหรับนำไปทดสอบประสิทธิภาพอัลกอริทึมทางด้าน AI ได้แก่ฐานข้อมูล Australian Credit [20], German Credit [21] และข้อมูลอีกส่วนหนึ่งเป็นข้อมูลที่ได้จากการเก็บข้อมูลจริงในงานวิจัยของนักศึกษาระดับปริญญาเอก ได้แก่ข้อมูล Bankruptcy Data [22] รายละเอียดดังตารางที่ 2

ตารางที่ 2 รายละเอียดข้อมูลที่ใช้ในการวิจัย

	ตัวอย่าง	แอตทริบิว	คลาส +1	คลาส -1
Australian Credit	690	14	307	383
German Credit	1000	24	300	700
Bankruptcy Data	240	30	128	112

จากข้อมูลข้างต้นผู้วิจัยได้ทำการแบ่งกลุ่มข้อมูลออกเป็นกลุ่มข้อมูลสำหรับสอน (Training Set) กลุ่ม

ข้อมูลสำหรับตรวจสอบ (Validate Set) และกลุ่มข้อมูลสำหรับทดสอบ (Test Set) โดยสำหรับกลุ่มข้อมูล Australian Credit แบ่งข้อมูลสำหรับสอนไว้ที่ 450 เรคคอร์ด แบ่งข้อมูลสำหรับตรวจสอบไว้ที่ 67 เรคคอร์ด และข้อมูลสำหรับทดสอบไว้ที่ 173 เรคคอร์ด ส่วนกลุ่มข้อมูล German Credit แบ่งข้อมูลสำหรับสอนไว้ที่ 700 เรคคอร์ด แบ่งข้อมูลสำหรับตรวจสอบไว้ที่ 100 เรคคอร์ด และข้อมูลสำหรับทดสอบไว้ที่ 200 เรคคอร์ด ข้อมูล Bankruptcy Data แบ่งข้อมูลสำหรับสอนไว้ที่ 168 เรคคอร์ด แบ่งข้อมูลสำหรับตรวจสอบไว้ที่ 24 เรคคอร์ด และข้อมูลสำหรับทดสอบไว้ที่ 48 เรคคอร์ด การประเมินผลใช้การประเมินผลจากประสิทธิภาพการจำแนกข้อมูล โดยวัดค่าความถูกต้องแม่นยำ (Accuracy) ในการจำแนกข้อมูลในกลุ่มข้อมูลทดสอบ ดังสมการที่ 13

$$Accuracy = \left(\frac{TP + TN}{TP + FN + TN + FP} \right) 100 \quad (13)$$

โดยที่

TP คือค่า True Positive

TN คือค่า True Negative

FP คือค่า False Positive

FN คือค่า False Negative

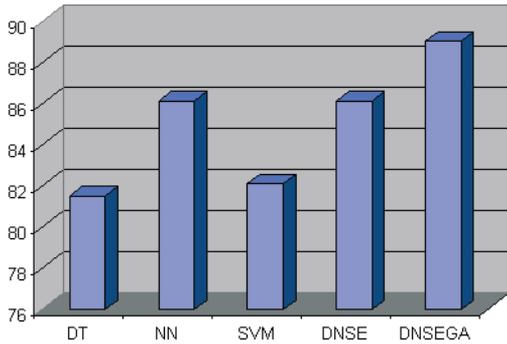
งานวิจัยครั้งนี้ ผู้วิจัยได้ใช้เครื่องคอมพิวเตอร์โน้ตบุ๊ก ซีพียู Intel Celeron M ความเร็ว 1.87 GHz. หน่วยความจำหลัก 1.5 Gb.

3.3 โมเดลที่ใช้เปรียบเทียบเพื่อหาประสิทธิภาพ

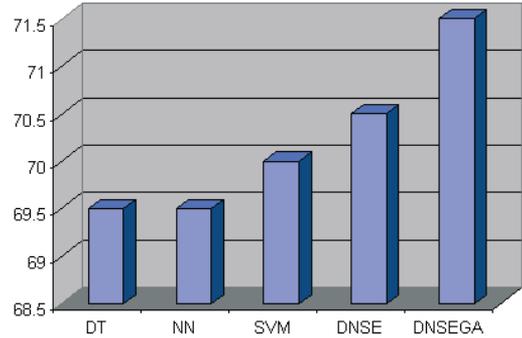
ในงานวิจัยนี้ผู้วิจัยทดลองประสิทธิภาพในการจำแนกข้อมูลสำหรับ 5 โมเดล ดังตารางที่ 3

ตารางที่ 3 รายละเอียดโมเดลที่ทดลอง

โมเดล	รายละเอียดโมเดลที่ทดลอง	ชื่อย่อ
1	Decision Tree	DT
2	Neural Network	NN
3	Support Vector Machine	SVM
4	DT+NN+SVM+Ensemble	DNSE
5	DT+NN+SVM+Ensemble+GA	DNSEGA



รูปที่ 7 กราฟเปรียบเทียบผลที่ได้จากแต่ละโมเดล



รูปที่ 8 กราฟเปรียบเทียบผลที่ได้จากแต่ละโมเดล

4. ผลการวิจัยและการอภิปรายผล

ผลการทดลองที่ได้จากงานวิจัยสำหรับฐานข้อมูล Australian Credit พบว่า โมเดล DNSEGA ซึ่งผลลัพธ์สุดท้ายใช้ตัวจำแนกข้อมูล 9 โมเดล ประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 4 โมเดล เทคนิคต้นไม้ตัดสินใจ 6 โมเดล และเทคนิคโครงข่ายประสาทเทียม 3 โมเดล และเทคนิคโครงข่ายประสาทเทียม 2 โมเดล ให้ผลลัพธ์ความแม่นยำที่ดีที่สุดที่ 89.01% รองลงมาเป็นโมเดล DNSE และ โมเดล NN ให้ผลความแม่นยำ 86.12% โมเดล SVM ให้ผลความแม่นยำที่ 82.08% และสุดท้ายโมเดล DT ให้ผลความแม่นยำที่ 81.05% ตามลำดับ ทั้งนี้จะเห็นว่าโมเดล NN และโมเดล DNSE ให้ความแม่นยำสูงเท่ากัน เนื่องจากจากลักษณะของข้อมูลที่มีลักษณะที่เหมาะสมกับโมเดล NN รวมทั้งการกำหนดเลือกใช้โครงสร้างที่เหมาะสม ได้แก่ การกำหนดจำนวนชั้นซ่อน การเลือกใช้รูปแบบฟังก์ชันกระตุ้น จำนวนนิวรอนชั้นซ่อนของโมเดล NN รวมถึงโมเดล DNSE เป็นลักษณะการรวมเอาหลายโมเดลเข้าด้วยกันเพื่อช่วยตัดสินใจเป็นการช่วยลดความโน้มเอียงทำให้มีความแม่นยำที่สูงขึ้น อีกด้วย ดังตารางที่ 4 และรูปที่ 7

ตารางที่ 4 ผลเปรียบเทียบที่ได้จากแต่ละโมเดล

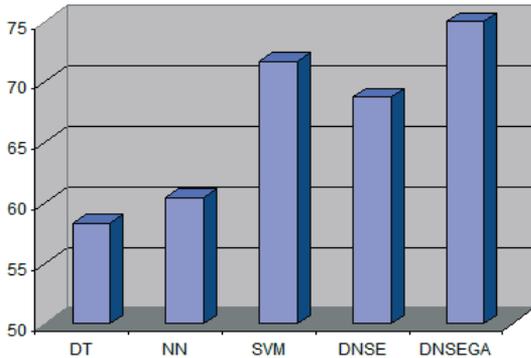
ข้อมูล Australian Credit	
โมเดล	ความแม่นยำ (%)
DT	81.50
NN	86.12
SVM	82.08
DNSE	86.12
DNSEGA	89.01

ผลการทดลองที่ได้จากงานวิจัยสำหรับฐานข้อมูล German Credit พบว่า โมเดล DNSEGA ซึ่งผลลัพธ์สุดท้ายใช้ตัวจำแนกข้อมูล 18 โมเดล ซึ่งประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 7 โมเดล เทคนิคต้นไม้ตัดสินใจ 6 โมเดล และเทคนิคโครงข่ายประสาทเทียม 5 โมเดล ให้ผลลัพธ์ความแม่นยำที่ดีที่สุดที่ 71.50% รองลงมาเป็นโมเดล DNSE ให้ผลความแม่นยำ 70.50% โมเดล SVM ให้ผลความแม่นยำที่ 70.00% โมเดล NN และสุดท้ายโมเดล DT ให้ผลความแม่นยำที่ 69.50% ตามลำดับ ดังตารางที่ 5 และรูปที่ 8

ตารางที่ 5 ผลเปรียบเทียบที่ได้จากแต่ละโมเดล

ข้อมูล German Credit	
โมเดล	ความแม่นยำ (%)
DT	69.50
NN	69.50
SVM	70.00
DNSE	70.50
DNSEGA	71.50

ผลการทดลองที่ได้จากงานวิจัยสำหรับฐานข้อมูล Bankruptcy Data พบว่า โมเดล DNSEGA ซึ่งผลลัพธ์สุดท้ายใช้ตัวจำแนกข้อมูล 10 โมเดล ซึ่งประกอบด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน 4 โมเดล เทคนิคต้นไม้ตัดสินใจ 4 โมเดล และเทคนิคโครงข่ายประสาทเทียม 2 โมเดล ให้ผลลัพธ์ความแม่นยำที่ดีที่สุดที่ 75.00%



รูปที่ 9 กราฟเปรียบเทียบผลที่ได้จากแต่ละโมเดล

รองลงมาเป็นโมเดล DNSE ให้ผลความแม่นยำ 68.75% โมเดล SVM ให้ผลความแม่นยำที่ 71.67% โมเดล NN ให้ผลความแม่นยำที่ 60.41% และโมเดล DT ให้ผลความแม่นยำที่ 58.33% ตามลำดับ ดังตารางที่ 6 และรูปที่ 9

ตารางที่ 6 ผลเปรียบเทียบที่ได้จากแต่ละโมเดล

ข้อมูล Bankruptcy Data	
โมเดล	ความแม่นยำ (%)
DT	58.33
NN	60.41
SVM	71.67
DNSE	68.75
DNSEGA	75.00

จากผลลัพธ์ที่ได้จากการทดลองจะพบว่า แนวทางในการนำตัวจำแนกข้อมูลที่หลากหลายมารวมกันตัดสินใจสามารถช่วยเพิ่มประสิทธิภาพความแม่นยำได้มากกว่าการใช้วิธีการแบบโมเดลเดี่ยว (Single Model) อันเนื่องมาจากการใช้เทคนิคการรวมกลุ่มสามารถลดการเกิดปัญหาความเอนเอียง (Bias) ซึ่งปัญหานี้มักเกิดขึ้นจากการกำหนดค่าของข้อมูลที่ใช้ในการเรียนรู้ รวมทั้งมีการกำหนดค่าพารามิเตอร์ที่ตายตัว อีกทั้งในการทดลองยังพบว่า จำนวนตัวจำแนกที่มากไม่ทำให้โมเดลที่ได้มีประสิทธิภาพเพิ่มขึ้นเสมอไป หากแต่การเลือกใช้ตัวจำแนกที่เหมาะสมจะช่วยทำให้โมเดลที่ได้มีประสิทธิภาพที่ดีที่สุด

5. สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยในครั้งนี้ มีวัตถุประสงค์เพื่อสร้างโมเดลที่มีประสิทธิภาพ สำหรับการจำแนกข้อมูล ผลที่ได้แสดงให้เห็นว่าการนำเอาเทคนิคต้นไม้ช่วยตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และซัพพอร์ตเวกเตอร์แมชชีน ร่วมกันตัดสินใจด้วยเทคนิคการรวมกลุ่ม โดยใช้วิธีการโหวตเสียงข้างมากอีกทั้งการเลือกตัวจำแนกข้อมูลที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม พบว่าเทคนิคการรวมกันตัดสินใจ ให้ผลลัพธ์ที่ดีกว่าการใช้เทคนิคแบบโมเดลเดี่ยว นอกจากนั้นยังพบอีกว่าการเลือกใช้ตัวจำแนกข้อมูลที่เหมาะสมในการรวมกลุ่มด้วยวิธีขั้นตอนเชิงพันธุกรรม เพื่อตัดสินใจจะทำให้โมเดลที่มีประสิทธิภาพยิ่งขึ้น ทั้งนี้เพราะเมื่อนำตัวจำแนกข้อมูลที่หลากหลายมารวมกันด้วยวิธีการร่วมกันตัดสินใจของตัวจำแนกเป็นการลดปัญหาการเกิดความเอนเอียงของข้อมูลและการเลือกใช้ตัวจำแนกที่ดีแต่ละตัวช่วยกันเสริมประสิทธิภาพในการจำแนกข้อมูล ทำให้โมเดลที่ได้มีประสิทธิภาพมากยิ่งขึ้น

ในอนาคตผู้วิจัยมีแนวคิดในการนำเอาวิธีการให้นำหนักสำหรับแต่ละโมเดลเช่นวิธีการเอาดาบัส (Adaboost) อีกทั้งหาวิธีการรวมกลุ่มแบบอื่น ๆ สำหรับประยุกต์ใช้เพื่อให้เกิดความแม่นยำในการจำแนกข้อมูลที่เพิ่มขึ้น

เอกสารอ้างอิง

- [1] K. Zeitouni and N. Chelghoum, "Spatial Decision Tree-Application to Traffic Risk Analysis," in *Proceedings of the ACS/IEEE International Conference on Computer Systems and Applications: IEEE Computer Society*, 2001.
- [2] เดช ธรรมศิริ ณรงค์ โพธิ วาทีน นุ้ยเพียร ภัทรวุฒิ แสงศิริ ภรณ์ยา อำนวยครุฑณ์ และพยุ่ง มีสัง, "การให้คะแนนสินเชื่อโดยวิธีการทำเหมืองข้อมูลด้วยเทคนิคโครงข่ายประสาทเทียม แบบแพร่กระจายย้อนกลับ," NCCIT2008, 2551.
- [3] เดช ธรรมศิริ และพยุ่ง มีสัง, "การจำแนกข้อมูลโดยวิธีการรวมกลุ่มของเทคนิคซัพพอร์ตเวกเตอร์



- แมชชีนโดยการปรับพารามิเตอร์ด้วยขั้นตอนวิธีเชิงพันธุกรรมรวมทั้งการเลือกใช้ลักษณะที่เหมาะสมด้วยวิธีการสัมประสิทธิ์สหสัมพันธ์,” NCCIT2010, 2553.
- [4] S.-j. Wang, A. Mathew, Y. Chen, L.-f. Xi, L. Ma, and J. Lee, “Empirical analysis of support vector machine ensemble classifiers,” *Expert Syst. Appl.*, vol. 36, pp. 6466-6476, 2009.
- [5] Y. Yao, Z. Fu, X. Zhao, and W. Cheng, “Combining Classifier Based on Decision Tree,” *icie*, vol. 2, pp.37-40, 2009 WASE International Conference on Information Engineering, 2009.
- [6] J. Wang, J. Yang, S. Li, Q. Dai, and J. Xie, “Number Image Recognition Based on Neural Network Ensemble,” in *Proceedings of the Third International Conference on Natural Computation-Volume 01*: IEEE Computer Society, 2007.
- [7] S. Chen, W. Wang, and H. van Zuylen, “Construct support vector machine ensemble to detect traffic incident,” *Expert Systems with Applications*, vol. 36, pp. 10976-10986, 2009.
- [8] K.-C. Ying, S.-W. Lin, Z.-J. Lee, and Y.-T. Lin, “An ensemble approach applied to classify spam e-mails,” *Expert Syst. Appl.*, vol. 37, pp. 2197-2201, 2010.
- [9] X. Pei-Yong, D. Xiang-Qian, and J. Bai-Ning, “A GA-based feature selection and ensemble learning for high-dimensional datasets,” in *Machine Learning and Cybernetics, 2009 International Conference on*, 2009, pp. 7-12.
- [10] L. Zhou, K. K. Lai, and L. Yu, “Least squares support vector machines ensemble models for credit scoring,” *Expert Systems with Applications*, vol. 37, pp. 127-133, 2010.
- [11] G. Martinez-Muoz, D. Hernandez-Lobato, and A. Suarez, “An Analysis of Ensemble Pruning Techniques Based on Ordered Aggregation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, pp. 245-259, 2009.
- [12] K.-H. Liu, B. Li, J. Zhang, and J.-X. Du, “Ensemble component selection for improving ICA based microarray data prediction models,” *Pattern Recognition*, vol. 42, pp. 1274-1283, 2009.
- [13] X. B. C. B. Ning, Y.W. Xu, and J. Zhang, “Virus-Evolutionary Genetic Algorithm Based Selective Ensemble Classifier for Pedestrian Detection,” in *GEC'09 Shanghai, China*, 2009.
- [14] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
- [15] พยุง มีสัจ, ระบบฟuzzyและโครงข่ายประสาทเทียม, เอกสารประกอบการสอน คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.
- [16] บุญเสริม กิจศิริกุล, อัลกอริทึมการทำเหมืองข้อมูล, ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2546.
- [17] J. H. Holland, “Outline for a logical theory of Adaptive systems,” *Journal of the Association for Computing Machinery*, vol. 3, pp. 297-314, 1962.
- [18] J. H. Holland, *Adaptation in Natural and Artificial Systems An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The MIT Press, 1992.
- [19] L. Breiman, “Bagging predictors,” *Machine Learning.*, vol. 24, pp. 123-140, 1996.
- [20] Australian Credit Approval Data set. (Feb., 10, 2008), [Online]. Available: <http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/australian/>
- [21] German Credit Data set. (Feb., 10, 2008), [Online]. Available: <http://archive.ics.uci.edu/ml/machine-learning-databases/statlog/german/>
- [22] W. Pietruszkiewicz, “Application of discrete Predicting structures in an earlywarning expert system for financial distress,” Ph.D. Thesis, Szczecin, Technical University, Szczecin., 2004.