# การหาค่าที่เหมาะสมในการจำแนกข้อมูลด้วยเทคนิคนิวรอลต์เน็ตเวิร์กแบบมัลติเลเยอร์เพอร์เซ็ปตรอน, ซัพพอร์ตเว็กเตอร์แมชชีน และเรเดียลเบสีสฟังก์ชั่น

# Determining appropriate of Data Classification with Multi-Layer Perceptron, Support Vector Machines and Radial Basis Function

*วัชรินทร์ วรินทักษะ[1] กิตติศักดิ์ โชติกิติพัฒน์[2] และทรงพล นคเรศเรืองศักดิ์[3]*

*[1]คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์*

*[2]คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์*

*[3]คณะวิทยาศาสตร์และเทคโนโลยี วิทยาลัยเซาธ์อีสท์บางกอก*

*watcharin.w@rmutr.ac.th, kittisak.cho@rmutr.ac.th, songpon@southeast.ac.th*

## บทคัดย่อ

*วัตถุประสงค์ของงานวิจัยนี้เพื่อทำการเปรียบเทียบการจำแนกข้อมูลของแบบจำลองที่สร้างขึ้นด้วยวิธีการทางเครือข่ายประสาทเทียมแบบมัลติเลเยอร์เพอร์เซ็ปตรอน ซัพพอร์ตเวกเตอร์แมชชีน และเรเดียลเบซีสฟังก์ชั่น โดยสร้างโมเดลจำนวน 30 โมเดล กับชุดข้อมูลมาตรฐานจาก UCI จำนวน 3 ชุดข้อมูล ที่มีคุณลักษณะแตกต่างกัน คือ Vote, Audiology, Ionosphere แล้ววัดประสิทธิภาพด้านความถูกต้องของการจำแนกข้อมูล (Accuracy) ค่าความแม่นยำ (Precision : P) ค่าความระลึก (Recall : R) และการวัดคุณภาพ (F- measure) ผลการวิจัยพบว่าโมเดลที่สร้างจากซัพพอร์ตเว็กเตอร์แมชชีน (SVM) ให้ประสิทธิภาพสูงที่สุดคือ 97.25% รองลงมาคือโมเดลที่สร้างจากวิธีเรเดียลเบซีสฟังก์ชั่น (RBF) คือ 96.80% และสุดท้ายโมเดลที่สร้างจากมัลติเลเยอร์เพอร์เซ็ปตรอน (MLP) คือ 96.09% ตามลำดับ*

**คำสำคัญ:** การจำแนกข้อมูล มัลติเลเยอร์เพอร์เซ็ปตรอน ซัพพอร์ตเวกเตอร์แมชชีน เรเดียลเบซีสฟังก์ชั่น เหมืองข้อมูล

## *Abstract*

*This research is using data mining to compare the identification information of the model created by the Multi-layer Perceptron Neural Network, Support Vector Machines and Redial Basis Function. The model consisted of 30 models of three series data from the UCI that features different is Vote, Audiology and Ionosphere to measure the efficiency and accuracy of data classification, Precision, Recall and quality measurement (F-measure).*

*The results showed that the model based on support vector machine (SVM) performance is 97.25%, followed by the creation of a Redial Basis Function (RBF) is 96.80% and the final model based on Multilayer Perceptron (MLP) is 96.09%, respectively.*

**Keyword:** Classification, Multilayer Perceptron, Support Vector Machines, Redial Basis Function, Data mining

## 1. บทนำ

การทำเหมืองข้อมูล (Data Mining) [1] คือ การนำข้อมูลจากฐานข้อมูลมาวิเคราะห์ และประมวลผล เพื่อค้นหารูปแบบ (Pattern) จากข้อมูลจำนวนมหาศาลโดยอัตโนมัติ โดยการสร้างแบบจำลอง และขั้นตอนวิธีจากวิชาสถิติ ซึ่งเทคนิคในการทำเหมืองข้อมูลนั้นมีหลายวิธี เช่น Decision Tree, K-Nearest Neighbor, Neuron Network, Support Vector Machines เป็นต้น การจะเลือกใช้ควรเลือกให้ตรงกับเป้าหมาย ปัญหาหรือวัตถุประสงค์ของการทำวิจัย โดยทั่วไปจะกระทำงานตามลักษณะของแบบจำลองที่ใช้ในการทำเหมืองข้อมูล ซึ่งสามารถแบ่งได้ 2 ประเภทใหญ่ ๆ คือ แบบจำลองเชิงทำนาย (Predictive Data Mining) และ แบบจำลองเชิงพรรณนา (Descriptive Data Mining)

งานวิจัยนี้ได้นำเทคนิคการจำแนกข้อมูลแบบนิวรอลเน็ตเวิร์กแบบมัลติเลเยอร์เพอร์เซ็ปตรอน ซัพพอร์ตเว็กเตอร์แมชชีน และเรเดียลเบซีสฟังก์ชั่น มาทำการสร้างโมเดลจำนวน 30 โมเดลแล้วทดสอบ 3 ชุดข้อมูลรวมทั้งสิ้น 90 โมเดล พร้อมทั้งทำการประเมินประสิทธิภาพในการจำแนกข้อมูลของแต่ละวิธี แล้วนำผลที่ดีที่สุดของแต่ละวิธีมาเปรียบเทียบเพื่อหาประสิทธิภาพในการจำแนกข้อมูลที่ดีที่สุดว่าเป็นวิธีใด และ

ศึกษาความเหมาะสมระหว่างเทคนิคที่เลือกใช้กับข้อมูลแต่ละชุด รวมทั้งทำการวิเคราะห์หาปัจจัยที่ทำให้ประสิทธิภาพการจำแนกข้อมูลดีขึ้น

## 2. วัตถุประสงค์ของการวิจัย

เพื่อทำการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลด้วยเทคนิคมัลติเลเยอร์เพอร์เซ็ปตอน (Multi-Layer Perceptron: MLP) ซัพพอร์ตเว็กเตอร์แมชชีน (Support Vector Machines: SVM) และเรเดียลเบซิสฟังก์ชั่น (Radial Basis Function: RBF) และศึกษาแต่เทคนิคว่าเหมาะสมกับการจำแนกข้อมูลแบบใด รวมทั้งวิเคราะห์หาปัจจัยที่ทำให้ประสิทธิภาพการจำแนกข้อมูลดีขึ้น

## 3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 3.1 การทำเหมืองข้อมูล (Data Mining)

การทำเหมืองข้อมูล (Data Mining) คือการ ค้นหาความรู้ที่ถูกซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ [2] โดยทำการค้นหารูปแบบความสัมพันธ์ของข้อมูล เพื่อค้นพบ และเกิดความรู้ใหม่ แล้วนำความรู้ที่ค้นพบมาวิเคราะห์ความน่าเชื่อถือด้วยวิธีการต่าง ๆ

ขั้นตอนในการทำเหมืองข้อมูล [3]

1. การคัดเลือกข้อมูล (Data Selection) เป็นการบอกถึงข้อมูลที่จำเป็นต้องใช้ทั้งหมด

2. การเตรียมข้อมูล (Data Preprocessing) เป็นการทำให้เกิดความมั่นใจในคุณภาพของข้อมูลที่นำมาใช้ โดยมีวิธีการดังนี้

2.1 การกรองข้อมูล (Data Cleaning) เป็นการวิเคราะห์ความถูกต้องของข้อมูลด้วยการนำข้อมูลที่ไม่ ถูกต้องออก เช่น ข้อมูลที่ถูกบันทึกผิด ข้อมูลที่ไม่สอดคล้องกัน เป็นต้น

2.2 การแปลงข้อมูล (Data Transformation) เป็นการแปลงข้อมูลที่ถูกเลือกมาให้อยู่ในรูปแบบที่เหมาะสมสำหรับการนำไปวิเคราะห์ข้อมูลตามเทคนิคต่าง ๆ ที่กำหนด

3 การทำเหมืองข้อมูล (Data Mining) เป็นการ นำเทคนิคของการทำเหมืองข้อมูลต่าง ๆ มาทำการ Mining ข้อมูล โดยทั่วไปมีการแบ่งกลุ่มของลักษณะของแบบจำลองที่ใช้ในการทำเหมืองข้อมูลอยู่ 2 ประเภทคือ

3.1 แบบจำลองเชิงทำนาย (Predictive Data Mining) คือ การพยากรณ์ คาดคะเนหรือประมาณค่าที่ชัดเจนของข้อมูลที่จะเกิดขึ้นจากฐานข้อมูลที่ผ่านมาในอดีต

3.2 แบบจำลองเชิงพรรณนา (Descriptive Data Mining) คือ การหาแบบจำลองเพื่ออธิบายลักษณะบางอย่างของข้อมูลที่มีอยู่ แล้วทำการแบ่งกลุ่มให้กับข้อมูลตามลักษณะของข้อมูล

4. การวิเคราะห์และประเมินผลลัพธ์ เป็นขั้นตอนของการแปลความหมาย และการประเมินผลลัพธ์ที่ได้ว่ามีความเหมาะสมหรือตรงกับวัตถุประสงค์ที่ต้องการหรือไม่

### 3.2 การคัดเลือกข้อมูลแบบกรีดดีอัลกอริทึม [4]

กรีดดีอัลกอรึทึม (Greedy Algorithm) เป็นการค้นหาแบบไปข้างหน้า และค้นหาแบบดีที่สุดก่อน (Best First Search) โดยการสร้างกราฟต้นไม้ จากนั้นจะค้นหาแบบชั้นต่อชั้น ระหว่างโหนดจะเชื่อมต่อกันด้วยกิ่งที่มีค่าน้ำหนักของแต่ละกิ่ง คำนวณได้จาก

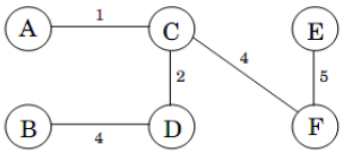
Input: G = (V; E) (1)

เมื่อ V = จำนวนโหนด, E = จำนวนกิ่ง, Output : T = (V; E), เมื่อ $E' \subseteq E$

$$\text{Weight (T)} = \sum w_e \quad (2)$$

เมื่อ $w_e$ = น้ำหนักแต่ละกิ่ง

เทคนิคของกรีดอัลกอริทึม จะพิจารณาเลือกทางที่สามารถเชื่อมต่อกันได้ทุกโหนด แต่ไม่ก่อให้เกิดเป็นกราฟวงกลม และมีค่าน้ำหนักรวมของทุกโหนดน้อยที่สุด



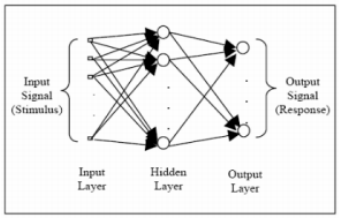

**ภาพที่ 1** การทำงานของกรีดดีอัลกอริธึม

ผลลัพธ์ที่ได้จากเทคนิคกรีดอัลกอริธึมจะได้เฉพาะคุณลักษณะที่สำคัญเท่านั้นและสามารถนำไปใช้ได้ทันที

### 3.3 มัลติเลเยอร์เพอร์เซ็ปตรอน (Multi-Layer Perceptron)

โครงข่ายประสาทเทียมแบบหลายชั้น (Multi-Layer Perceptron) [5] หรือ MLP ซึ่งโหนดของนิวรอลจะถูกจัดเป็นชั้น ๆ ในกรณีที่ใช้สถาปัตยกรรมป้อนไปข้างหน้า เอาท์พุตของ

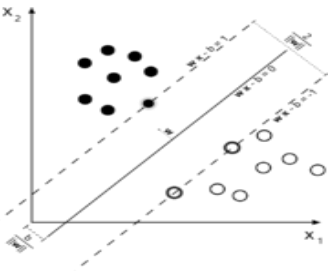
ชั้นหนึ่ง ๆ จะทำหน้าที่เป็นอินพุตของชั้นถัดไป ดังนั้นเอาท์พุตของโครงข่ายจะเป็นชั้นสุดท้ายของโครงข่ายด้วย ส่วนชั้นที่อยู่ระหว่างชั้นอินพุต และชั้นเอาท์พุตจะถูกเรียกว่าชั้นซ่อน (Hidden Layer) ดังแสดงในภาพที่ 3 ในรูปแบบดังกล่าวเป็นโครงข่ายแบบ 3 ชั้น โดยมี $x_1,x_2,x_3,\ldots,x_M$ เป็นอินพุต และ $y_1,y_2,y_3,\ldots.y_N$ เป็นเอาท์พุตระหว่างโครงข่าย และมีชั้นซ่อนเป็นชั้นระหว่างชั้นอินพุตกับชั้นเอาท์พุต




**ภาพที่ 3** แสดงโครงข่ายประสาทเทียมแบบหลายชั้น

**3.4 ซัพพอร์ตเว็กเตอร์แมชชีน** (Support Vector Machine)

ซัพพอร์ทเวกเตอร์แมชชีน [6], [7] หรือ SVM ใช้วิธีการหาระนาบการตัดสินใจในการแบ่งข้อมูลออกเป็นสองส่วน โดยใช้สมการเส้นตรงเพื่อแบ่งเขตข้อมูล 2 กลุ่มออกจากกัน โดยจะพยายามสร้างเส้นแบ่งตรงกึ่งกลางระหว่างกลุ่มให้มี ระยะห่างระหว่างขอบเขตของทั้งสองกลุ่มมากที่สุด แสดงดังภาพที่ 4



**ภาพที่ 4** เส้นแบ่งกลุ่มข้อมูลจากซัพพอร์ตเว็กเตอร์แมชชีน

ซัพพอร์ตเว็กเตอร์แมชชีนจะใช้ฟังก์ชั่นแม็ปสำหรับย้ายข้อมูลจาก Input Space ไปยัง Feature Space และสร้างฟังก์ชั่นวัดความคล้ายที่เรียกว่าเคอร์เนลฟังก์ชั่น (Kernel Function) บน Feature Space เหมาะที่จะใช้สำหรับข้อมูลที่มีมิติของข้อมูลสูง กำหนดให้ $(x_i,y_i),\ldots,(x_n,y_n)$ เป็นตัวอย่างที่ใช้สำหรับการสอน n คือ จำนวนข้อมูลตัวอย่าง m คือ จำนวนมิติข้อมูลเข้า และ y คือผลลัพธ์มีค่า + 1 หรือ -1 ดังสมการ

$$(x_i,y_i),\ldots,(x_n,y_n) \text{ เมื่อ } x \in R^m, y \in \{+1,-1\} \quad (3)$$

สำหรับปัญหาเชิงเส้น มิติข้อมูลขนาดสูงได้ถูกแบ่งเป็น 2 กลุ่ม โดยระนาบตัดสินใจ สามารถคำนวณได้ดังสมการ

$$\mathbf{wx} + b = 0 \quad (4)$$

เมื่อ **w** คือ ค่าน้ำหนัก

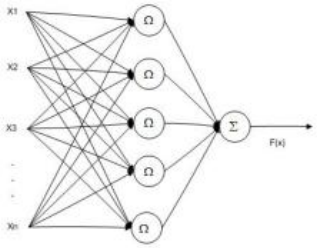
$b$ คือ ค่า bias สมการ

$\mathbf{wx} + b > 0$ ถ้า $y_i = +1$ และ $\mathbf{wx} + b < 0$ ถ้า $y_i = -1$

อย่างไรก็ตามซัพพอร์ตเว็กเตอร์แมชชีนมีเคอร์เนลฟังก์ชั่น (Kernel Function) ที่ผู้วิจัยสามารถประยุกต์ใช้ในการแก้ปัญหาได้หลายวิธี เช่น Normalized Poly Kernel, PUK Kernel และ RBF Kernel ซึ่งได้นำมาใช้ในการทดลองนี้

**3.5 เรเดียลเบซีสฟังก์ชั่น** (Radial Basis Function)

เรเดียลเบซีสฟังก์ชั่น [8] หรือ RBF เป็นโครงข่ายสำหรับการจำแนกข้อมูลแบบป้อนไปข้างหน้าโดยการคำนวณหาจุดศูนย์กลางของข้อมูลและแบ่งข้อมูลออกเป็นกลุ่มเพื่อจัดกลุ่มของข้อมูล โครงข่ายเรเดียลเบซีสฟังก์ชั่นจะประกอบด้วย 3 ชั้น คือ ชั้นข้อมูลเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นข้อมูลออก (Output Layer) แสดงดังภาพที่ 5



**ภาพที่ 5** แสดงโครงข่ายของเรเดียลเบซีสฟังก์ชั่น

จากภาพที่ 5 เมื่อ $x_i$ คือข้อมูลเข้า (Input) ที่ถูกส่งค่าไปยังชั้นซ่อน (Hidden Layer) ซึ่งสามารถคำนวณได้จาก

$$\theta_j(x) = \exp = \frac{x - c_j e}{2\sigma^2 j} \quad \text{เมื่อ } j = 1,2,\ldots,n \quad (5)$$

เมื่อ $\theta$ คือข้อมูลออกที่ j ในชั้นซ่อน, x คือข้อมูลเข้า, $c_j$ คือศูนย์กลางของโหนดที่ j จากนั้นคำนวณค่าข้อมูลออก (Output) ซึ่งสามารถคำนวณได้จาก

$$y = i_c(k+1) = \sum_{i=1}^{n} w_i \theta_i(x) \quad (6)$$

เมื่อ $w_j$ คือค่าน้ำหนักระหว่างชั้นซ่อนและชั้นข้อมูลออก และ y คือผลลัพธ์

จากการศึกษาผู้วิจัยพบว่ามีงานวิจัยที่เกี่ยวกับการเปรียบเทียบการจำแนกข้อมูลด้วยวิธีทาง Data Mining มีดังนี้

เดช และคณะ [9] นำเสนอวิธีการจำแนกข้อมูลด้วยวิธีร่วมกันตัดสินใจแบบการโหวตเสียงข้างมากจากเทคนิคต้นไม้ตัดสินใจด้วยเทคนิค MLP และ SVM โดยเลือกตัวแทนที่

เหมาะสมด้วยขั้นตอนเชิงพันธุกรรม ผลการทดลองพบว่าเทคนิคการร่วมกันตัดสินใจจากหลายโมเดลที่ผ่านการเลือกด้วยขั้นตอนวิธีเชิงพันธุกรรมมีผลลัพธ์ที่ดีกว่าการใช้เทคนิคแบบโมเดลเดี่ยว

ภรัณยา และคณะ [10] นำเสนอการเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกข้อมูลของโรคลมร้อน โดยนำมาวิเคราะห์ทำการเปรียบเทียบประสิทธิภาพ ผลการทดลองพบว่าวิธีการจำแนกข้อมูลด้วย MLP ให้ผลลัพธ์ที่ดีกว่า SVM

ภรัณยา และคณะ [1] นำเสนอการเปรียบเทียบประสิทธิภาพของแบบจำลองในการคัดเลือกและจำแนกข้อมูลโดยใช้ข้อมูล Ozone จากฐานข้อมูล UCI โดยใช้วิธีการคัดเลือกข้อมูล ผลการทดลองพบว่า การใช้ SVM ที่ใช้ kernel ด้วย rbf ในการจำแนกข้อมูลนั้นจะมีประสิทธิภาพที่ดีกว่าการใช้แบบจำลองโครงข่ายประสาทเทียม MLP

## 4. วิธีการดำเนินการวิจัย

การดำเนินการวิจัยครั้งนี้ ผู้วิจัยเลือกใช้โปรแกรม WEKA เวอร์ชั่น 3.7.5 [12] เป็นเครื่องมือในการทดลอง เนื่องจากเป็นซอฟต์แวร์ในการทำเหมืองข้อมูลที่ได้รับการยอมรับอย่างกว้างขวาง โดยกำหนดวิธีวิจัยตามลำดับขั้นตอนดังต่อไปนี้

### 4.1 การเตรียมข้อมูล

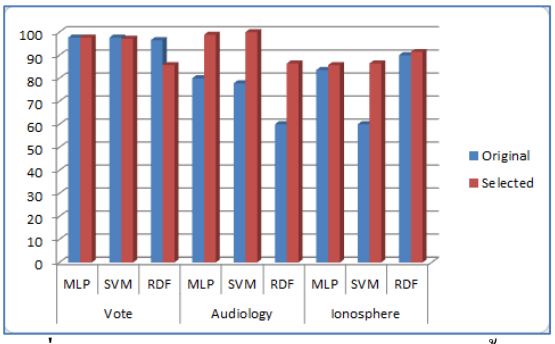
ในการวิจัยครั้งนี้ผู้วิจัยเลือกใช้ชุดข้อมูลมาตรฐานจาก UCI จำนวน 3 ชุด แสดงรายละเอียดของข้อมูลทั้ง 3 ชุดข้อมูลดังตารางที่ 1 รายละเอียดชุดข้อมูลที่ใช้ในงานวิจัย

**ตารางที่ 1** รายละเอียดชุดข้อมูลที่เลือกมาใช้ในงานวิจัย

| ชุดข้อมูล | แอทริบิวต์ | จำนวน | คลาส |
|---|---|---|---|
| **Vote** | 17 | 435 | Republican = 168<br>Democrat = 267 |
| **Audiology** | 70 | 226 | Class f = 224<br>Class t = 2 |
| **Ionosphere** | 35 | 351 | Class b = 126<br>Class g = 225 |

### 4.2 คัดเลือกคุณลักษณะที่สำคัญด้วยกรีดดีอัลกอรึธึม

จากการพิจารณาข้อมูลทั้ง 3 ชุดพบว่ามีจำนวนแอทริบิวต์อยู่เป็นจำนวนมาก จึงทำการคัดเลือกด้วยวิธีกรีดดีอัลกอริทึม ให้คงเหลือเฉพาะคุณลักษณะที่สำคัญ และหลังจากคัดเลือกแล้วคงเหลือแอทริบิวต์ของแต่ละชุดข้อมูลดังนี้ Vote(5), Audiology(16) และ Ionosphere(15) แล้วเปรียบเทียบผลของชุดข้อมูลเดิมกับชุดที่เลือกแอทริบิวต์ด้วยโมเดลดีฟอลต์ของ MLP, SVM และ RDF ได้ผลลัพธ์แสดงดังภาพที่ 6




**ภาพที่ 6** เปรียบเทียบค่าความถูกต้องของชุดข้อมูลทั้ง 3 ชุด

จากภาพที่ 6 พบว่าค่าความถูกต้องของชุดข้อมูล Audiology และ Ionosphere ที่ผ่านการคัดเลือกเฉพาะคุณลักษณะที่สำคัญด้วยกรีดดีอัลกอริทึม ให้ผลลัพธ์ที่สูงกว่า และ Vote ให้ค่าความถูกต้องเท่ากัน ดังนั้นผู้วิจัยจึงเลือกชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะสำคัญแล้ว มาใช้ในการทดลอง

### 4.2 ออกแบบการทดลอง

ออกแบบและสร้างโมเดลสำหรับการทดสอบกับข้อมูลทั้ง 3 ชุดข้อมูลด้วยเทคนิคของการทำเหมืองข้อมูลที่เลือกไว้คือ MLP, SVM และ RDF โดยสร้างเทคนิคละ 10 โมเดล รวม 3 เทคนิค จึงเท่ากับ 30 โมเดล สำหรับทดสอบกับข้อมูลที่เลือกมา 3 ชุด

### 4.3 สร้างโมเดลสำหรับการทดลอง

4.3.1 สร้างโมเดลสำหรับการสอนระบบด้วย MLP จำนวน 10 โมเดลแสดงดังตารางที่ 2

**ตารางที่ 2** โมเดลสำหรับทดสอบด้วย MLP

| Model | Hidden-Node | Learning Late | Momentum |
|---|---|---|---|
| 1 | 3-3 | 0.2 | 0.2 |
| 2 | 4-4 | 0.2 | 0.2 |
| 3 | 5-5 | 0.3 | 0.1 |
| 4 | 3-8-3 | 0.5 | 0.2 |

| | | | |
|---|---|---|---|
| 5 | 3-12-3 | 0.5. | 0.2 |
| 6 | 4-8-4 | 0.5 | 0.2 |
| 7 | 4-8-4 | 0.2 | 0.2 |
| 8 | 4-12-4 | 0.5 | 0.2 |
| 9 | 5-5-5 | 0.5 | 0.2 |
| 10 | 5-8-5 | 0.5 | 0.1 |

4.3.2 สร้างโมเดลสำหรับการสอนระบบด้วย SVM จำนวน 10 โมเดลแสดงดังตารางที่ 3

**ตารางที่ 3** โมเดลสำหรับทดสอบด้วย SVM

| Model | Training Set | Kernel Function | Complexity |
|---|---|---|---|
| 1 | 66% | Normalized Poly Kernel | 1 |
| 2 | 66% | PUK | 1 |
| 3 | 66% | RBF Kernel | 1 |
| 4 | Folds 5 | Normalized Poly Kernel | 1 |
| 5 | Folds 5 | PUK | 1 |
| 6 | Folds 5 | RBF Kernel | 1 |
| 7 | Folds 10 | Normalized Poly Kernel | 2 |
| 8 | Folds 10 | PUK | 2 |
| 9 | Folds 10 | RBF Kernel | 2 |
| 10 | Folds N | RBF Kernel | 2 |

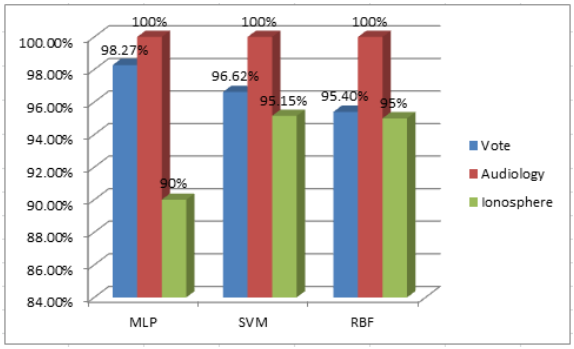
4.3.2 สร้างโมเดลสำหรับการสอนระบบด้วย RDF จำนวน 10 โมเดลแสดงดังตารางที่ 4

**ตารางที่ 4** โมเดลสำหรับทดสอบด้วย RDF

| Model | Training Set | ClusteringSeed | NumberClusters |
|---|---|---|---|
| 1 | 60% | 2 | 1 |
| 2 | 65% | 2 | 1 |
| 3 | 70% | 2 | 1 |
| 4 | 75% | 2 | 1 |
| 5 | 80% | 2 | 1 |
| 6 | 85% | 2 | 1 |
| 7 | 90% | 2 | 1 |
| 8 | Folds 10 | 2 | 1 |
| 9 | Folds 5 | 3 | 2 |
| 10 | Folds N | 5 | 2 |

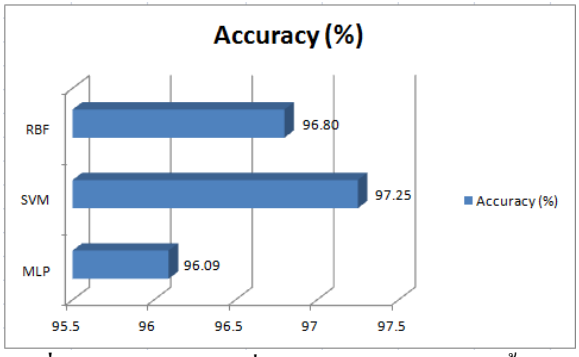
**4.4 ผลการดำเนินงาน**

จากโมเดลในตารางที่ 2, 3, 4 ผู้วิจัยนำไปทดสอบกับข้อมูลทั้ง 3 ชุดที่เลือกไว้จาก UCI คือ Vote, Audiology และ Ionosphere เพื่อวัดประสิทธิภาพในการจำแนกข้อมูล และคัดเลือกเฉพาะโมเดลที่ให้ค่าความถูกต้องสูงที่สุด (Best Model) แล้วนำมาเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลโดยพิจารณาจากค่าความถูกต้อง (Accuracy: A) ของแต่ละวิธีแสดงดังภาพที่ 7




**ภาพที่ 7** เปรียบเทียบค่าถูกต้องของข้อมูลทั้ง 3 ชุดกับ 3 เทคนิค

จากภาพที่ 7 เนื่องจากผู้วิจัยใช้ข้อมูล 3 ชุดทดสอบกับ 3 เทคนิค คือ MLP, SVM และ RDF ทำให้ไม่สามารถนำผลลัพธ์ที่ได้มาเปรียบเทียบประสิทธิภาพของการจำแนกข้อมูลได้อย่างชัดเจน ผู้วิจัยจึงหาเฉลี่ยจากค่าความถูกต้อง (Accuracy) ในการจำแนกข้อมูลของแต่ละเทคนิคมาใช้ในการเปรียบเทียบประสิทธิภาพ ผลลัพธ์ที่ได้แสดงดังภาพที่ 8




**ภาพที่ 8** เปรียบเทียบค่าเฉลี่ยความถูกต้องของข้อมูลทั้ง 3 ชุด

**5. สรุปผลการวิจัย**

งานวิจัยนี้ต้องการนำเสนอผลของการเปรียบเทียบประสิทธิภาพในการจำแนกข้อมูลด้วยเทคนิค MLP, SVM และ

RBF และต้องการทราบว่าเทคนิคใดเหมาะสมกับข้อมูลประเภทใดจากชุดข้อมูล 3 ชุดที่เลือกมา รวมทั้งต้องการทราบว่ามีปัจจัยอะไรบ้างที่มีผลทำให้การจำแนกข้อมูลมีประสิทธิภาพมากยิ่งขึ้น หลังจากที่ได้ทำการทดลองตามวัตถุประสงค์ที่กำหนดไว้ข้างต้น ผู้วิจัยขอสรุปผลการวิจัยจากการทดลองดังนี้

5.1 จากภาพที่ 8 แสดงให้เห็นว่าประสิทธิภาพในการจำแนกข้อมูลด้วยโมเดลของทั้ง 3 วิธีนั้นหากพิจารณาจากค่าเฉลี่ย (Average) พบว่าโมเดลที่สร้างจากวิธี SVM ให้ประสิทธิภาพในภาพรวมสูงที่สุดคือ 97.25% รองลงมาคือโมเดลที่สร้างจากวิธี RBF คือ 96.80% และสุดท้ายโมเดลที่สร้างจาก MLP คือ 96.09% ตามลำดับ

5.2 จากผลการทดลองที่แสดงในภาพที่ 7 พบว่าเทคนิคที่ผู้วิจัยเลือกมาคือ MLP, SVM และ RDF เพื่อทดสอบหาประสิทธิภาพในการจำแนกข้อมูล Vote, Audiology และ Ionosphere ให้ค่าความถูกต้องตั้งแต่ 90% – 100% ซึ่งเป็นค่าที่สูงมาก จึงสรุปได้ว่าทั้ง 3 เทคนิคเหมาะสมกับข้อมูลทั้ง 3 ชุดข้อมูลที่เลือกมา

5.3 จากการทดลองพบว่าปัจจัยที่ทำให้ผลการทดลองมีค่าความถูกต้องสูงขึ้นคือการคัดเลือกคุณลักษณะเฉพาะของแต่ละชุดข้อมูล เพื่อลดจำนวนของแอทริบิวต์ที่ใช้ในการทดสอบลง ทำให้ได้แอทริบิวต์ที่เหมาะสมของแต่ละชุดข้อมูลมาใช้ในการทดลอง

ในการวิจัยครั้งนี้ผู้วิจัยพบว่าการจำแนกข้อมูลโดยการคัดเลือกคุณลักษณะที่สำคัญมีผลอย่างมากในการเพิ่มประสิทธิภาพในการวิเคราะห์ความถูกต้อง และเพิ่มความเร็วในการทำงานของโมเดลให้สูงขึ้น ซึ่งงานวิจัยนี้เลือกใช้วิธีกรีดดีอัลกอริทึม เพราะสามารถนำผลลัพธ์ที่ได้ไปใช้งานได้ทันที แต่วิธีการนี้อาจจะไม่ใช่วิธีที่ดีที่สุด ดังนั้นในงานวิจัยครั้งต่อไปผู้วิจัยจะทดสอบโดยการเปรียบเทียบผลการคัดเลือกคุณสมบัติกับวิธีอื่น ๆ เช่น CfsSubsetEval, Principal Components, Information Gain, Rough Set เป็นต้น แล้วเลือกวิธีที่ให้ผลลัพธ์ดีที่สุดมาใช้ในงานวิจัยครั้งต่อไป

**เอกสารอ้างอิง**


[1] ภรัณยา อำมฤครัตน์, เดช ธรรมศิริ, วาทินี นุ้ยเพียร, ภัทราวุฒิ แสงศิริ, ณรงค์ โพธิ, พยุง มีสัจ. “การเปรียบเทียบประสิทธิภาพการคัดเลือกข้อมูลและจำแนกข้อมูลด้วยวิธีการทางเครือข่ายประสาทเทียม”, *The 5th National Conference on Computing and Information Technology. 2009 :131-138*

[2] Jiawei Han and Micheline Kamber., “Data Mining Concepts and Techniques”, *Morgan Kaufmann Publishers, 2001.*

[3] สุคนธ์ทิพย์ วงศ์พันธ์ และอนงค์ ศรีวิหค, “การจำแนกข้อมูลโดยใช้เทคนิคจีเนติกอัลกอริทึมและสร้างต้นไม้ตัดสินใจเพื่อจำแนกพฤติกรรมการกระทำความผิดของนักเรียนระดับอาชีวะ”, *National Computer Science and Engineering Conference. 2007: 749-756*

[4] จิราภรณ์ สุขใหญ่, สิริภัทร เชี่ยวชาญวัฒนา, คำรณ สุนัติ. “การคัดเลือกคุณลักษณะข้อมูลสำหรับปัญหาการจำแนกประเภทข้อมูลด้วยเทคนิคโครงข่ายประสาทเทียมเอ็กทรีม” in The 14th Nation Computer Science and Engineering Conference: 89 – 94

[5] พยุง มีสัจ. ระบบฟัซซีและโครงข่ายประสาทเทียม, คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ 2555.

[6] C .Cortes and V.Vapnik.1995 "Support-Vector Networks". Machine Learning Volume 20, Number 3 ,September

[7] พรพล ธรรมรงค์รัตน์, ลัดดา ปรีชาวีรกุล, วิภาดา เวทย์ประสิทธิ, “การจำแนกประเภทเว็บเพจโดยใช้ค่าความถี่เอกสารและซัพพอร์ตเวกเตอร์แมชชีน”,*The 12th National Computer Science and Engineering Conference 2008.*

[8] จิราภรณ์ ถมแก้ว และศรัณย์ อินทโกสุม. “การจำแนกข้อมูลโดยการคัดเลือกคุณลักษณะที่สำคัญ”, *วารสารการประชุมวิชาการเสนอผลงาน วิจัยระดับบัณฑิตศึกษาแห่งชาติครั้งที่ 23: 7-12*

[9] เดช ธรรมสิริ และพยุง มีสัจ, “การจำแนกข้อมูลด้วยวิธีแบบร่วมกันตัดสินใจจากพื้นฐานของเทคนิคต้นไม้ตัดสินใจ เทคนิคโครงข่ายประสาทเทียม และเทคนิคซัพพอร์ตเว็กเตอร์แมชชีนร่วมกันคัดเลือกตัวแปรที่เหมาะสมด้วยวิธีเชิงพันธุกรรม”, *วารสารพระจอมเกล้าพระนครเหนือ ปีที่ 21: 293-303*

[10] ภรัณยา อำมฤครัตน์ และพยุง มีสัจ, “การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มข้อมูลของโรคลมร้อนด้วยวิธีการทางเครือข่ายประสาทเทียม”, *The 6th National Conference on Computing and Information Technology 2010: 116-121*

[11] WEKA http://learners.in.th/blog/kanjanablog/110815