บทที่ 1
บทนำ

ในปัจจุบันข้อมูลที่จัดเก็บในฐานข้อมูลได้เพิ่มขึ้นอย่างมาก เช่น ฐานข้อมูลพยากรณ์อากาศที่ต้องมีการเก็บข้อมูลเป็นจำนวนมาก จะมีการเก็บข้อมูลทุกวัน หรือทุก 4 ชั่วโมง ทำให้ข้อมูลอาจมีความไม่สม่ำเสมอและไม่สมบูรณ์ เช่น มีค่าสุดโต่ง มีการบันทึกข้อมูลผิดพลาด ดังนั้นจึงต้องมีการกรองข้อมูลก่อนนำข้อมูลไปประมวลผล เพื่อเพิ่มประสิทธิภาพในการประมวลผลให้มีความถูกต้องมากยิ่งขึ้น

สิ่งรบกวน (Noise) เป็นข้อมูลที่ไม่ต้องการหรือเป็นค่าข้อมูลผิดพลาดที่เกิดขึ้นในข้อมูล รูปแบบของสิ่งรบกวนสามารถพิจารณาได้จากฟังก์ชั่นความหนาแน่นที่น่าจะเป็นไปได้ (Probability Density Function: PDF) เช่น Gaussian Noise จะทำให้ฮิสโตแกรมของข้อมูลมีลักษณะเป็นรูประฆังคว่ำ และ Salt-and-Pepper Noise เป็นสิ่งรบกวนที่มีลักษณะเป็นค่าสุดโต่ง [6] เป็นต้น

การกรองข้อมูล (Data Filtering) เป็นกระบวนการที่ใช้สำหรับขจัดสิ่งรบกวนออกจากข้อมูล เช่น ค่าสุดโต่ง ค่าข้อมูลผิดพลาด เป็นต้น ทำให้ข้อมูลราบเรียบ มีคุณภาพและเหมาะสำหรับนำไปประมวลผล ทำให้การพยากรณ์มีค่าความถูกต้องมากขึ้น ซึ่งเทคนิคการกรองข้อมูลมีหลายเทคนิค เช่น เทคนิคการกรองข้อมูลแบบเฉลี่ยเคลื่อนที่ (Moving Average Filtering) เทคนิคการกรองข้อมูลแบบถดถอยระดับท้องถิ่น (Local Regression Filtering) เทคนิคการกรองข้อมูลแบบสาวิสกี้-โกเลย์ (Savitzky-Golay Filtering) และเทคนิคการกรองข้อมูลแบบหน้าต่างแฮมมิ่ง (Hamming Window Filtering) [7] เป็นต้น

การทำเหมืองข้อมูล (Data Mining) เป็นวิธีการที่นำมาช่วยในการค้นหาความรู้ หรือสกัดความรู้จากฐานข้อมูลที่มีข้อมูลเป็นจำนวนมาก เพื่อเพิ่มประสิทธิภาพการทำงานให้รวดเร็วขึ้น เทคนิคที่ใช้ในการทำเหมืองข้อมูล เช่น เทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network) เป็นเทคนิคที่สามารถคำเนินการกับข้อมูลที่มีปริมาณมากที่อยู่ในรูปของตัวเลขได้อย่างมีประสิทธิภาพ [3, 4] มีการเรียนรู้จากชุดข้อมูลตัวอย่างหรือการเรียนรู้แบบมีผู้สอน (Supervised Learning) สามารถพยากรณ์ผลลัพธ์ที่จะเกิดขึ้นได้ และมีความแม่นยำ (Accuracy) ในการพยากรณ์สูง

ในงานวิจัยนี้ได้ใช้เทคนิคการกรองข้อมูลเพื่อขจัดสิ่งรบกวนในข้อมูลและใช้เทคนิคโครงข่ายประสาทเทียมมาสร้างแบบจำลอง และพัฒนาโปรแกรมทดสอบเทคนิคการกรองข้อมูลแต่ละแบบสำหรับฐานข้อมูลการพยากรณ์อากาศ และฐานข้อมูลด้านการแพทย์ เพื่อให้ได้ผลการพยากรณ์ที่ถูกต้องมากยิ่งขึ้น

1.1 วัตถุประสงค์

1. ออกแบบจำลองการกรองข้อมูลโดยใช้โครงข่ายประสาทเทียมกับเทคนิคการกรองข้อมูลแบบต่างๆ

2. พัฒนาระบบการกรองข้อมูลเพื่อใช้กับโครงข่ายประสาทเทียม

1.2 ขอบเขตการวิจัย

1. งานวิจัยนี้จะพิจารณาการกรองข้อมูลโดยใช้โครงข่ายประสาทเทียมจากข้อมูลทางการแพทย์ และข้อมูลอากาศ

2. งานวิจัยนี้จะทดสอบเทคนิคการกรองข้อมูลแบบต่างๆ สำหรับโครงข่ายประสาทเทียม เช่น เทคนิค Moving Average Filtering, เทคนิค Local Regression Filtering, เทคนิค Savitzky-Golay Filtering และ เทคนิค Hamming Window Filtering

1.3 ระยะเวลาการดำเนินงาน และแผนการดำเนินงาน

มิถุนายน 2549 – พฤษภาคม 2550

กิจกรรม/ขั้นตอนการดำเนินงาน	เดือน											
	2549							2550				
	6	7	8	9	10	11	12	1	2	3	4	5
1. ศึกษาและทำความเข้าใจข้อมูลอากาศและการกรองข้อมูล	■	■										
2. ศึกษาและทำความเข้าใจเอกสารและงานวิจัยที่เกี่ยวข้อง	■	■	■									
3. ศึกษาเทคโนโลยีและเครื่องมือสนับสนุน		■	■	■								
4. วิเคราะห์และออกแบบระบบ				■	■	■	■	■				
5. พัฒนาระบบ						■	■	■	■			
6. ทดสอบและติดตั้งระบบ										■		
7. จัดทำเอกสารประกอบระบบ และเขียนผลงานวิจัย									■	■	■	
8. จัดทำเอกสาร									■	■	■	■

1.4 ขั้นตอนวิธีการวิจัย

1. ศึกษางานวิจัยและเอกสารที่เกี่ยวข้อง

2. เตรียมข้อมูล โดยข้อมูลที่นำมาสกัดความรู้เป็นข้อมูลทางการแพทย์ และข้อมูลพยากรณ์อากาศ ดังนี้

- กลุ่มฐานข้อมูลทางการแพทย์ ที่เป็นมาตรฐาน และ เป็นที่ยอมรับในระดับสากล จาก University of California at Irvine (UCI) ftp://ftp.ics.uci.edu [5]
- กลุ่มฐานข้อมูลทางด้านการแพทย์ ที่เก็บตัวอย่างมาจากโรงพยาบาลในประเทศไทย [1]
- กลุ่มฐานข้อมูลอากาศ จากกรมอุตุนิยมวิทยา ประเทศไทย http://www.tmd.go.th [2]
- กลุ่มฐานข้อมูลพายุ เมืองริโอเดอจาเนโร และ เมืองเซาเปาโล ประเทศบราซิล http://www.wunderground.com [8]

3. การกรองข้อมูล (Data Filtering) เป็นการกำจัดสิ่งรบกวนออกจากข้อมูล ทำการทดลองโดยใช้เทคนิคการกรองข้อมูล 4 แบบคือ Moving Average Filtering, Local Regression Filtering, Savitzky-Golay Filtering และ Hamming Window Filtering เพื่อพิจารณาว่าข้อมูลแต่ละกลุ่มนั้นเหมาะสมกับเทคนิคการกรองข้อมูลแบบใด

4. การสอนโครงข่ายประสาทเทียม (Neural Network Training) เป็นขั้นตอนสอนโครงข่ายประสาทเทียม

5. พัฒนาโปรแกรมกรองข้อมูลโดยใช้โครงข่ายประสาทเทียม โดยใช้โปรแกรมประยุกต์ MATLAB 7.0

6. ทดสอบโปรแกรม สรุป และประเมินผล

7. เขียนผลงานวิจัย และนำเสนอผลงานวิจัย

1.5 สถานที่ทำการวิจัย

ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตหาดใหญ่

1.6 อุปกรณ์ที่ใช้ในงานวิจัย

1. ด้าน Hardware เครื่องคอมพิวเตอร์ 1 เครื่อง หน่วยความจำ 512 MB ฮาร์ดดิสก์ 80 GB

2. ด้าน Software ใช้ระบบปฏิบัติการ Microsoft Windows XP โปรมแกรม MATLAB 7.0

1.7 ผลที่คาดว่าจะได้รับ

1. ได้แบบจำลองการกรองข้อมูลโดยใช้โครงข่ายประสาทเทียมกับเทคนิคการกรองข้อมูลแบบต่างๆ

2. ได้ระบบโปรแกรมการกรองข้อมูลเพื่อใช้กับโครงข่ายประสาทเทียม

1.8 ทฤษฎีที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้องในงานวิจัยชิ้นนี้แบ่งออกเป็น 3 ส่วนคือ การเตรียมข้อมูล (Data Preprocessing) เทคนิคการกรองข้อมูล (Data Filtering) และโครงข่ายประสาทเทียม (Artificial Neural Networks) ซึ่งมีรายละเอียด ดังนี้

1.8.1 การเตรียมข้อมูล (Data Preprocessing)

การจัดเก็บข้อมูลรวบรวมข้อมูลจากหลายแหล่ง ในบางครั้งข้อมูลอาจอยู่ในรูปแบบที่แตกต่างกัน หรือมีความไม่สมบูรณ์ ซึ่งถ้าหากข้อมูลที่ได้มานั้นไม่มีคุณภาพ จะยังผลให้ผลลัพธ์ที่ได้มีคุณภาพต่ำไปด้วย ซึ่งสาเหตุในการเตรียมข้อมูลสามารถแบ่งได้เป็นหัวข้อหลักๆดังนี้ [9] คือ 1) ข้อมูลไม่มีความสมบูรณ์ (Incomplete) คือ มีข้อมูลบางตัวสูญหายไป 2) ข้อมูลมีค่าสุดโต่ง (Outlier) คือ ค่าข้อมูลที่จัดเก็บมาก หรือน้อยเกินกว่าขอบเขตที่กำหนด 3) ข้อมูลที่ได้ไม่มีความสม่ำเสมอ (Inconsistent) คือ ข้อมูลอาจอยู่ในรูปแบบที่แตกต่างกัน และ 4) ข้อมูลไม่อยู่ในรูปแบบที่จะประมวลผลโดยใช้โครงข่ายประสาทเทียม เช่น ข้อมูลอาจอยู่ในรูปแบบสัญลักษณ์ หรือ ตัวอักษร เป็นต้น

1.8.1.1 ทำความสะอาดข้อมูล (Data Cleaning) เป็นขั้นตอนที่แก้ปัญหา Missing Value โดยเราสามารถใช้วิธีจัดการ Missing Value ได้ 3 วิธี [3] คือ 1) ถ้า Attribute ที่เกิด Missing Value มีค่าเป็นตัวเลขจำนวนจริงได้ จะแทน Missing Value ด้วยค่าเฉลี่ยของกลุ่ม (Class Mean) 2) ถ้า Attribute ที่เกิด Missing Value ต้องมีค่าเป็นตัวเลขจำนวนนับจะแทน Missing Value ด้วยค่ามัธยฐานของกลุ่ม (Class Median) และ 3) ลบแถวข้อมูลที่มี Missing Value ทิ้งไป

- การแทนค่าข้อมูลสูญหาย (Missing Value) ด้วยค่าเฉลี่ยของข้อมูลรอบข้างดังสมการที่ (1.1)

$$\textit{ค่าข้อมูลสูญหาย} = \frac{\textit{ค่าก่อนหน้าค่าข้อมูลสูญหาย} + \textit{ค่าหลังค่าข้อมูลสูญหาย}}{2} \qquad (1.1)$$

1.8.1.2 เปลี่ยนรูปข้อมูล (Data Transformation) โดยแปลงข้อมูลที่ไม่พร้อมที่จะนำไปประมวลผล ให้อยู่ในรูปที่พร้อมที่จะประมวลผล [10, 11] ซึ่งการเปลี่ยนรูปแบบข้อมูลสามารถแบ่งออกเป็น 3 วิธี [12] ดังนี้

- การเปลี่ยนรูปข้อมูลนามกำหนด (Nominal Data) คือ ใช้เทคนิคการแบ่งช่วงข้อมูลและแทนค่าข้อมูลเป็นตัวเลข เช่น ข้อมูลสีทั้งหมด 4 สี คือ {สีแดง, สีเขียว, สีฟ้า, สีเหลือง} ถ้าต้องการแปลงข้อมูลให้อยู่ในช่วง [0, 1] จะแทนค่าข้อมูลได้ดังนี้ {0.00, 0.33, 0.67, 1.0}

- การเปลี่ยนข้อมูลรูปตัวเลข (Numerical Data) คือ การแปลงข้อมูลเข้าที่อยู่ในรูปตัวเลข ให้อยู่ในช่วงค่าที่ต้องการ [10] แสดงได้ดังสมการ (1.2) เช่น การแปลงให้อยู่ในช่วง [0, 1] โดยค่าต่ำสุดของช่วงคือ 0 และค่าสูงสุดของช่วงคือ 1 ค่าข้อมูลเข้า คือ {100, 200, 300, 400} เมื่อผ่านการแปลงรูปข้อมูลจะได้ผลลัพธ์ คือ {0.0, 0.33, 0.66, 1.0}

$$\textit{ค่าข้อมูลใหม่} = \frac{(\textit{ค่าข้อมูลเดิม}-\textit{ค่าต่ำสุดของช่วง})}{(\textit{ค่าสูงสุดช่วง}-\textit{ค่าต่ำสุดช่วง})} \tag{1.2}$$

- การเปลี่ยนรูปข้อมูลโดยการเพิ่มโหนดข้อมูลเข้า (Use of Additional Input Node) คือ มีการเพิ่มโหนด เช่น ข้อมูลสี 4 สีสามารถกำหนดโหนดข้อมูลเข้า 2 โหนด คือ สีแดง = [0, 0] สีเขียว = [0, 1] สีฟ้า = [1, 0] และ สีเหลือง = [1, 1] เป็นต้น

1.8.2 การกรองข้อมูล (Data Filtering)

เป็นการกำจัดสิ่งรบกวนออกจากข้อมูล ทำให้ข้อมูลคุณภาพ เพื่อให้สามารถทำนายผลลัพธ์ได้ดี ในงานวิจัยนี้ได้ทดลองเทคนิคการกรองข้อมูล 4 เทคนิค ดังนี้ คือ Moving Average Filtering, Local Regression Filtering, Savitzky-Golay Filtering และ Hamming Window Filtering ซึ่งมีรายละเอียดดังนี้ [7]

1.8.2.1 Moving Average Filtering เป็นการกรองข้อมูลอย่างง่าย โดยใช้หลักการหาค่าเฉลี่ยของทุกจุดในหน้าต่างที่ได้กำหนดไว้ แสดงได้ตามสมการที่ (1.3)

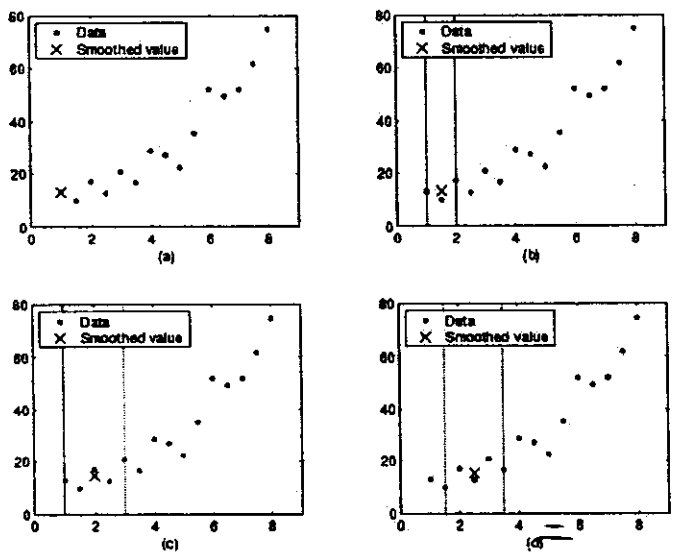
$$(y_k)_s = \frac{\sum_{i=-n}^{i=n} y_{k+i}}{2n+1} \tag{1.3}$$

กำหนดให้ ตัวแปร $(y_k)_s$ คือ ข้อมูลตำแหน่งที่ k ที่ผ่านการกรองแล้ว

ตัวแปร y_k คือ ข้อมูลตำแหน่งที่ k

ตัวแปร n คือ จำนวนของจุดข้อมูลในแต่ละข้างด้านซ้ายหรือด้านขวาไปจนถึงขอบหน้าต่าง

ภาพประกอบที่ 1.1 เป็นตัวอย่างการกรองข้อมูลแบบ Moving Average Filtering โดยกำหนดขนาดของหน้าต่างไว้ 5 จุด ภาพประกอบที่ 1.1 (a) แสดงว่าในจุดแรกของข้อมูลจะไม่มีการกรอง เพราะไม่มีหน้าต่างกำหนดไว้ ภาพประกอบที่ 1.1 (b) แสดงว่าข้อมูลจุดที่สองจะถูกปรับค่าโดยใช้ค่าเฉลี่ยของข้อมูล 3 จุด ภาพประกอบที่ 1.1 (c) และ (d) แสดงว่าข้อมูลจุดที่ 3 และจุดที่ 4 ถูกปรับค่าโดยใช้ค่าเฉลี่ยของข้อมูล 5 จุด เป็นต้น



ภาพประกอบที่ 1.1 เทคนิคการกรองแบบ Moving Average Filtering

1.8.2.2 Local Regression Filtering เป็นเทคนิคการกรองข้อมูลที่ใช้หลักการของการถดถอยแบบกำลังสองน้อยที่สุดมาใช้ในการคำนวณ (Linear Least Square Regression) ตัวอย่างในการคำนวณสมการถดถอยเชิงเส้นโดยใช้พหุนามเชิงเส้น (Linear Polynomial) ดังนี้ ถ้าให้ x เป็นตัวแปรต้น และ y เป็นตัวแปรตาม จากนั้นนำข้อมูลมาสร้างแผนภาพการกระจาย หากพบว่า y และ x มีความสัมพันธ์เป็นเส้นตรง จะได้ รูปแบบการถดถอยเชิงเส้นอย่างง่ายตามสมการที่ (1.4) [13]

$$y = a_0 + a_1 x \tag{1.4}$$

กำหนดให้ ตัวแปร y คือ ตัวแปรตาม

ตัวแปร x คือ ตัวแปรต้น

ตัวแปร a_0 คือ ค่าสัมประสิทธิ์ (จุดตัดแกนตั้ง)

ตัวแปร a_1 คือ ค่าสัมประสิทธิ์ (ค่าความชัน)

จากนั้นต้องทำการแก้สมการเพื่อหาค่า a_0 และ a_1 การคำนวณค่าสัมประสิทธิ์จากวิธีกำลังสองน้อยที่สุดนั้นจะต้องทำให้ผลรวมกำลังสองของเศษเหลือ (r) มีค่าน้อยที่สุดซึ่งคำนวณหาค่า r ตามสมการที่ (1.5)

$$r_i = y_i - \hat{y}_i \tag{1.5}$$

กำหนดให้ ตัวแปร r_i คือ เศษเหลือของข้อมูลตำแหน่งที่ i

ตัวแปร y_i คือ ตัวแปรตามตำแหน่งที่ i

ตัวแปร $\hat{y}$ คือ ค่าตัวแปรตามที่คำนวณจากสมการถดถอย

คำนวณผลรวมกำลังสองของเศษเหลือจากสมการที่ (1.6)

$$S = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{1.6}$$

กำหนดให้ ตัวแปร n คือ จำนวนข้อมูลทั้งหมด

ตัวแปร S คือ ผลรวมกำลังสองของเศษเหลือ

จากสมการถดถอยเชิงเส้น $\hat{y} = a_0 + a_1 x$ จะได้

$$S = \sum_{i=1}^{n} (y_i - (a_0 + a_1 x_i))^2 \tag{1.7}$$

เนื่องจากต้องการสัมประสิทธิ์ที่ทำให้คาดคะเนค่าตัวแปรตามได้ผิดพลาดน้อยที่สุด (ผลรวมกำลังสองของเศษเหลือน้อยที่สุด) ดังนั้น จึงกำหนดผลลัพธ์ของสมการเป็น 0

$$\frac{\partial S}{\partial a_1} = -2\sum_{i=1}^{n} x_i (y_i - (a_1 x_i + a_0)) = 0 \tag{1.8}$$

$$\frac{\partial S}{\partial a_0} = -2\sum_{i=1}^{n} (y_i - (a_1 x_i + a_0)) = 0 \tag{1.9}$$

แก้สมการจะได้สมการดังนี้

$$\sum x_i (y_i - (a_1 x_i + a_0)) = 0 \tag{1.10}$$

$$\sum(y_i - (a_1 x_i + a_0)) = 0 \tag{1.11}$$

เมื่อผลรวม รวมตั้งแต่ $i = 1$ ถึง n จะได้สมการปกติ ดังนี้

$$a_1 \sum x_i^2 + a_0 \sum x_i = \sum x_i y_i \tag{1.12}$$

$$a_1 \sum x_i + n a_0 = \sum y_i \tag{1.13}$$

แก้สมการจะได้ค่าสัมประสิทธิ์ a_0 และ a_1 ดังนี้

$$a_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \tag{1.14}$$

$$a_0 = \frac{1}{n}(\sum y_i - a_1 \sum x_i) \tag{1.15}$$

นอกจากนั้น Local Regression Filtering มีการกำหนดน้ำหนักให้ข้อมูลทุกจุด ในหน้าต่างที่กำหนดไว้ โดยใช้ฟังก์ชั่นน้ำหนักถดถอย (Regression Weight Function) แสดงได้ตามสมการที่ (1.16) โดยจุดที่มีน้ำหนักมากจะมีอิทธิพลต่อผลการกรองข้อมูลมาก

$$w_i = \left(1 - \left|\frac{x - x_i}{d(x)}\right|^3\right)^3 \tag{1.16}$$

กำหนดให้ ตัวแปร x คือ ตัวแปรต้น

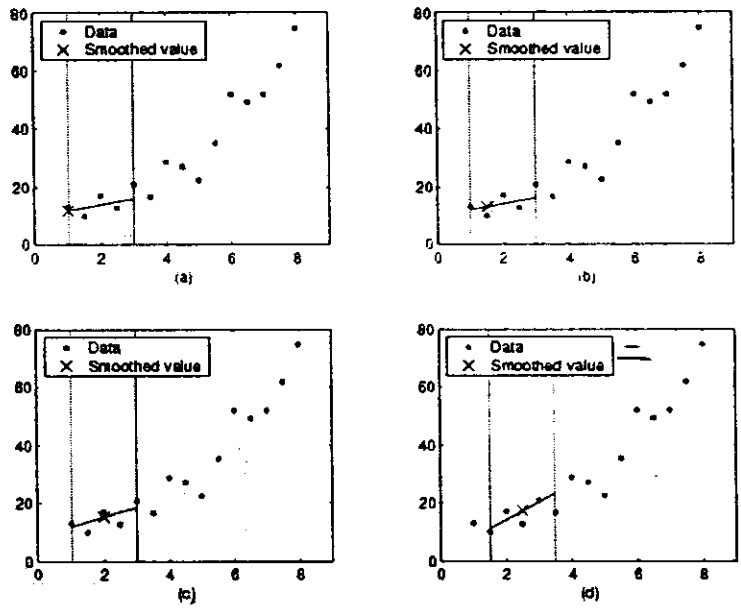
ตัวแปร x_i คือ ตัวแปรต้นที่อยู่ใกล้ x ตำแหน่งที่ i

ตัวแปร $d(x)$ คือ ระยะห่างของ x ไปยัง x_i ที่ไกลที่สุด

เมื่อได้ค่าน้ำหนักของแต่ละจุดข้อมูลแล้วจะนำค่าน้ำหนักไปรวมกับสมการที่ (1.6) เป็นสมการใหม่ที่เป็นสมการถดถอยแบบถ่วงน้ำหนัก แสดงตามสมการ (1.17) และจากนั้นก็จะแก้สมการและหาค่าสัมประสิทธิ์ต่อไป

$$S = \sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2 \tag{1.17}$$

เทคนิคการกรองข้อมูล Local Regression Filtering สามารถแยกเป็นสองแบบ คือ lowess และ loess ความแตกต่างของทั้งสองแบบ คือ รูปแบบของสมการถดถอย โดยที่ lowess ใช้พหุนามเชิงเส้น (Linear Polynomial) และ loess ใช้พหุนามกำลังสอง (Quadratic Polynomial) [7] ตัวอย่างการกรองข้อมูลแบบ lowess แสดงได้ดังภาพประกอบที่ 1.2 ซึ่งการกรองข้อมูลแบบ lowess นั้น ขนาดของหน้าต่างจะไม่เปลี่ยนแปลง ภาพประกอบที่ 1.2 (a) และ (b) แสดงการใช้ฟังก์ชั่นน้ำหนักแบบไม่สมมาตร ภาพประกอบที่ 1.2 (c) และ (d) ใช้ฟังก์ชั่นน้ำหนักแบบสมมาตร และนอกจากนั้นสามารถประยุกต์ใช้ Robust Function เข้ากับ lowess และ loess เพื่อไม่ให้ Outlier มีผลกับการกรองข้อมูล เมื่อประยุกต์เข้ากับ lowess กลายเป็น rlowess และประยุกต์เข้ากับ loess กลายเป็น rloess



ภาพประกอบที่ 1.2 เทคนิคการกรองแบบ Local Regression Filtering (lowess)

1.8.2.3 Savitzky-Golay Filtering เป็นเทคนิคการกรองข้อมูลที่นิยมใช้กันในข้อมูลที่มีลักษณะเป็นความถี่ โดยใช้หลักการของสมการถดถอยแบบกำลังสองน้อยที่สุดมาใช้ในการคำนวณ (Linear Least Square Regression) เช่นเดียวกับเทคนิคการกรองข้อมูลแบบ Local Regression Filtering แต่แตกต่างตรงที่เทคนิคการกรองข้อมูลแบบ Savitzky-Golay Filtering จะไม่มีการกำหนดค่าน้ำหนักให้กับจุดข้อมูลในหน้าต่าง [7,14] ลักษณะเด่นของเทคนิคการกรองข้อมูลแบบนี้ คือ จะเก็บรายละเอียดของข้อมูลเอาไว้ไม่กรองรายละเอียดข้อมูลออกไปจนไม่เหลือรายละเอียดของข้อมูล เทคนิคการกรองแบบนี้ได้ถูกนำไปประยุกต์ใช้ในหลายสาขา เช่น การลดสิ่งรบกวนในภาพอัลตร้าซาวด์ [15] การลดสิ่งรบกวนในภาพถ่ายเรดาร์ (Synthetic Aperture Radar SAR) [16] เป็นต้น

ในการกรองข้อมูลโดยใช้เทคนิคนี้จะต้องระบุพารามิเตอร์ 2 ตัวคือ ขนาดของหน้าต่าง และระดับขั้นพหุนาม (Degree of Polynomial Orders) โดยที่ขนาดของหน้าต่างจะต้องมากว่าระดับขั้นของพหุนามเสมอ

1.8.2.4 Hamming Window Filtering เป็นเทคนิคการกรองข้อมูลโดยใช้เทคนิคของการดำเนินการกับสัญญาณ (Signal Processing Technique) ใช้การกรองความถี่แบบอิมพัลส์จำกัด (Finite Impulse Response) [17] แสดงได้ดังสมการที่ (1.18)

$$\hat{h}(k) = h(k) \bullet w(k) \tag{1.18}$$

กำหนดให้ ตัวแปร $\hat{h}(k)$ คือ ข้อมูลใหม่ที่ผ่านการกรองแล้ว

ตัวแปร $h(k)$ คือ ข้อมูลเดิม

ตัวแปร $w(k)$ คือ ฟังก์ชันหน้าต่างซึ่งเป็นฟังก์ชันที่ใช้ในการลดการกระเพื่อมและปรับปรุงช่วงการตอบสนองการเปลี่ยนแปลง (Transition)

สามารถคำนวณหาค่า $w(k)$ ได้จากสมการที่ (1.19)

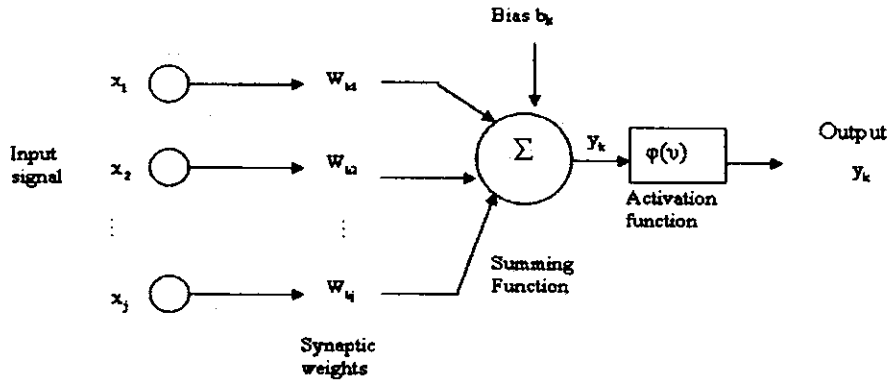
$$w(k)=\begin{cases} 0.54 - 0.5\cos\dfrac{2\pi k}{K-1} & ; 0 \le k \le K-1 \\ 0 & ; \text{อื่นๆ} \end{cases} \tag{1.19}$$

กำหนดให้ ตัวแปร k คือ ขนาดของจุดข้อมูลที่ทำการประมวลผล

ตัวแปร K คือ ขนาดของจุดข้อมูลของฟังก์ชันหน้าต่างที่ผู้ใช้กำหนด

1.8.3 โครงข่ายประสาทเทียม (Neural Networks)

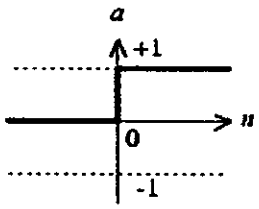
เป็นรูปแบบการประมวลผลที่เลียนแบบการทำงานของสมองมนุษย์ มีการทำงานแบบขนาน หลักการทำงานประกอบด้วยหน่วยประมวลผลย่อย (นิวรอน) หลายหน่วยทำงานเชื่อมต่อกัน แสดงได้ดังภาพประกอบที่ 1.3 ซึ่งแต่ละหน่วยสามารถปรับค่าพารามิเตอร์ประจำหน่วยได้จากกระบวนการเรียนรู้โดยใช้ตัวอย่าง แบบจำลองที่ผ่านการฝึกฝนแล้ว จะสามารถนำไปใช้แก้ปัญหาจริงได้ [18] ข้อดีของโครงข่ายประสาทเทียมคือสามารถทำนายค่าที่มีความแม่นยำสูง และทนทานต่อการทำงานล้มเหลว (Fault Tolerance) [19]



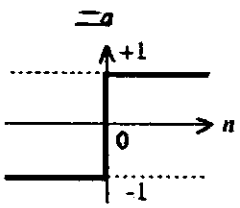
ภาพประกอบที่ 1.3 หน่วยประมวลผลย่อย (นิวรอน)

1.8.3.1 ฟังก์ชั่นในการในการคำนวณของนิวรอน ประกอบด้วย

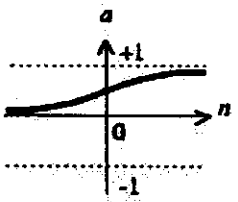
1. ฟังก์ชั่นในการหาผลรวมของค่าน้ำหนักข้อมูลเข้า (Summation Function)
2. ฟังก์ชั่นกระตุ้น (Activation Function) มีเพื่อจำกัดการขยายตัวของนิรอน เช่น ฟังก์ชั่นสเตป ฟังก์ชั่นซายน์ ฟังก์ชั่นซิกมอยด์ และฟังก์ชั่นแทนซิกมอยด์ แสดงดังภาพประกอบที่ 1.4 – 1.7 ตามลำดับ



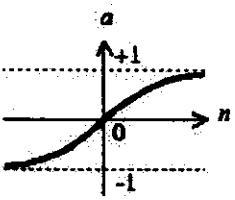
ภาพประกอบที่ 1.4 ฟังก์ชั่นสเตป



ภาพประกอบที่ 1.5 ฟังก์ชั่นซายน์

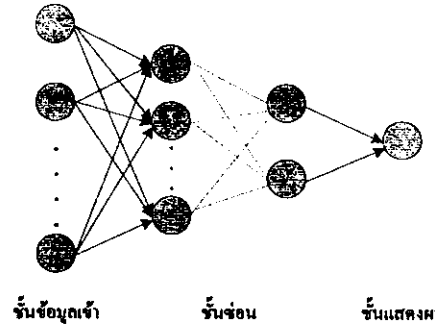


ภาพประกอบที่ 1.6 ฟังก์ชั่นซิกมอยด์



ภาพประกอบที่ 1.7 ฟังก์ชั่นแทนซิกมอยด์

1.8.3.2 โครงสร้างของโครงข่ายประสาทเทียม แบ่งเป็น 3 ระดับ คือ ชั้นข้อมูลเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นแสดงผล (Output Layer) โดยในแต่ละชั้นจะมีกี่นิวรอนก็ได้ และในชั้นซ่อนจะมีกี่ระดับก็ได้ ดังแสดงในภาพประกอบที่ 1.8



ภาพประกอบที่ 1.8 โครงสร้างของโครงข่ายประสาทเทียม

1.8.3.3 ประเภทของโครงข่ายประสาทเทียม มี 3 ประเภท คือ

1. โครงข่ายประสาทเทียมไปข้างหน้าชั้นเดียว (Single-Layer Feedforward Networks) เป็นโครงข่ายประสาทเทียมที่มีทิศทางไปข้างหน้าและมีเฉพาะชั้นข้อมูลเข้าและชั้นแสดงผล

2. โครงข่ายประสาทเทียมแบบไปข้างหน้าหลายชั้น (Multilayer Feedforward Networks) เป็นโครงข่ายประสาทเทียมที่มีทิศทางไปข้างหน้า มีชั้นข้อมูลเข้า ชั้นซ่อน และชั้นแสดงผล

3. โครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Recurrent Network) เป็นโครงข่ายประสาทเทียมที่สามารถมีลูปในการวนกลับจากชั้นแสดงผล ไปเป็นชั้นข้อมูลเข้าได้อย่างน้อย 1 รอบ

1.8.3.4 การสอนโครงข่ายประสาทเทียม เป็นขั้นตอนสอนโครงข่ายประสาทเทียม ซึ่งประกอบด้วย 3 ขั้นตอนย่อย ดังนี้

1. Random Training Set and Testing Set เป็นขั้นตอนในการสุ่มข้อมูลเพื่อสร้างชุดสอน (Training set) และ ชุดทดสอบ (Testing set)

2. Create Neural Network เป็นขั้นตอนในการสร้างโครงสร้างของโครงข่ายประสาทเทียม

3. Train Neural Network เป็นขั้นตอนในการสอนโครงข่ายประสาทเทียม