การพยากรณ์ปริมาณการใช้ยาในโรงพยาบาล โดยใช้โครงข่ายประสาทเทียม

พนิดา ยืนยงสวัสดิ์* และพยุง มีสัจ**

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโปรแกรมการพยากรณ์ปริมาณการใช้ยาโดยใช้โครงข่ายประสาทเทียม ใช้ข้อมูลปริมาณการใช้ยาที่ถูกเก็บไว้ในลักษณะของอนุกรมเวลา (Time Series) สำหรับทำการสอน (Training) โครงข่ายประสาทเทียมแบบฟีดฟอร์เวิร์ดหลายชั้น การเรียนรู้แบบมีการควบคุม (Supervised Learning) สำหรับเครื่องมือพัฒนาได้ใช้โปรแกรมแมตแล็บ 6.5 สร้างโมเดลโครงข่ายประสาทเทียมสำหรับการพยากรณ์ และนำโมเดลที่ได้ไปใช้สำหรับการพยากรณ์ปริมาณการใช้ยาโดยใช้โปรแกรมวิชวลเบสิค 6.0 ผลการทดสอบระบบการพยากรณ์แบบรายวันซึ่งใช้โมเดลแบบ 3-5-1 และมีการซ้อนทับของหน้าต่างแบบ 2 จุดให้ค่าความผิดพลาดที่ 0.011 ส่วนการพยากรณ์แบบรายสัปดาห์ซึ่งใช้โมเดลแบบ 3-5-1 และมีการซ้อนทับของหน้าต่างแบบ 2 จุดซึ่งให้ค่าความผิดพลาด 0.052 จากการเปรียบเทียบการพยากรณ์แบบรายวันมีความถูกต้องมากกว่าการพยากรณ์แบบรายสัปดาห์

คำสำคัญ : การพยากรณ์ โครงข่ายประสาทเทียม ปริมาณการใช้ยา ดาต้าไมนิง

1. บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันองค์กร หรือหน่วยงานต่าง ๆ ได้นำเทคโนโลยีสารสนเทศมาใช้งานกันอย่างกว้างขวางทุกวงการ ทำให้มีการเก็บข้อมูลของระบบงานต่างๆ ไว้ในคลังข้อมูลอยู่เป็นจำนวนมากจึงควรนำข้อมูลที่มีอยู่มากมายนั้นมาใช้ให้เกิดประโยชน์อย่างสูงสุดและจำเป็นอย่างยิ่งที่จะต้องควบคุมกระบวนการแยกหรือสืบค้นเพื่อเจาะลึกข้อมูลทางความรู้จากฐานข้อมูล ซึ่งมักใช้เครื่องมืออย่างดาต้าไมนิง (Data mining) [1], [2] ที่สามารถดึงข้อมูลที่ซ่อนเร้นอยู่และยังไม่เคยถูกค้นพบ มาสกัดเป็นสารสนเทศ หรือการคาดเดาสารสนเทศที่ถูกจัดเก็บในคลังข้อมูล [3] เพื่อสนับสนุนผู้บริหารในการวิเคราะห์รูปแบบ (Pattern) และแนวโน้ม (Trend) ในการประกอบการตัดสินใจ และปรับเปลี่ยนกลยุทธ์ การทำดาต้าไมนิงมีเทคนิคต่าง ๆ หลายเทคนิค ซึ่งเทคนิคของโครงข่ายประสาทเทียม (Neural Networks) ถูกนำไปประยุกต์ใช้ในงานดาต้าไมนิงได้หลายรูปแบบ เช่น การพยากรณ์อากาศ การพยากรณ์หุ้นในตลาดหลักทรัพย์ เป็นต้น การใช้โครงข่ายประสาทเทียมจะสามารถสร้างโมเดลเพื่อที่จะทำการทำนายหรือพยากรณ์ได้อย่างมีประสิทธิภาพโดยการหารูปแบบและแนวโน้มของข้อมูลของสิ่งที่เราสนใจ ซึ่งโรงพยาบาลเป็นองค์กรหนึ่งที่ประกอบด้วยระบบงานต่าง ๆ ที่มีการเก็บข้อมูลจำนวนมหาศาลยากต่อการทำความเข้าใจและการนำมาใช้ อีกทั้งยังไม่ได้มีการดึงข้อมูลมาสกัดให้เป็นสารสนเทศเพื่อที่จะใช้ประโยชน์ได้ ปัญหาที่มีของโรงพยาบาลคือต้องการสำรองยาให้เป็นไปตามความเป็นจริงโดยไม่ส่งผลกระทบต่องบประมาณในการจัดซื้อยา และมีความเพียงพอกับความต้องการใช้ยา ซึ่งเป็นสิ่งสำคัญในการวางแผนและตัดสินใจสำหรับการสำรองยาในอนาคตด้วย ด้วยเหตุผลนี้จึงจำเป็นต้องทราบค่าพยากรณ์ปริมาณการใช้ยาในอนาคตให้ใกล้เคียงกับความเป็นจริง

1.2 วัตถุประสงค์

1.2.1 เพื่อสร้างโปรแกรมการพยากรณ์ปริมาณการใช้ยาในอนาคต จากโมเดลของโครงข่ายประสาทเทียม

1.2.2 เพื่อประเมินผลโมเดลข้อมูลในการพยากรณ์ปริมาณการใช้ยาในอนาคต

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ดาต้าไมนิง (Data mining)

ดาต้าไมนิง คือ กระบวนการที่จะสกัดความรู้ที่มีประโยชน์จากข้อมูลดิบที่ซ่อนเร้นอยู่ ที่เราไม่ทราบมาก่อน ทำให้เกิดศักยภาพในการใช้ข้อมูลในฐานข้อมูล [4]

2.1.1 ขั้นตอนในการทำดาต้าไมนิง โดยทั่วไปกระบวน

* *นักศึกษา ระดับปริญญาโท คณะเทคโนโลยีสารสนเทศ สจพ.*

** *อาจารย์ ระดับ 7 ภาควิชาครุศาสตร์ไฟฟ้า คณะครุศาสตร์อุตสาหกรรม สจพ. และ รองคณบดีฝ่ายบริหาร คณะเทคโนโลยีสารสนเทศ สจพ.*

การในการทำดาต้าไมนิง ประกอบด้วย 5 ขั้นตอนใหญ่ๆ คือ

1) การกำหนดวัตถุประสงค์ (Business Objective Determination)

2) การเตรียมข้อมูล (Data Preparation) เป็นการทำปรับข้อมูลให้อยู่ในรูปแบบที่เหมาะสม ซึ่งประกอบด้วย 3 ขั้นตอนดังนี้

2.1) การคัดเลือกข้อมูล (Data Selection) ซึ่งเป็นการกำหนดลักษณะ และคัดเลือกข้อมูลที่ต้องการและนำข้อมูลที่ไม่ต้องการออก โดยพิจารณาจากวัตถุประสงค์ที่กำหนดไว้ และทำการแปลงข้อมูลที่ใช้จากหลายแหล่งให้อยู่ในรูปแบบเดียวกัน ในการคัดเลือกข้อมูลที่จะใช้นั้นควรต้องเข้าใจความหมาย ประเภทของข้อมูล และค่าที่สามารถเป็นไปได้สำหรับข้อมูล ซึ่งลักษณะของข้อมูลนั้นสามารถแบ่งได้เป็น 2 ลักษณะคือ

ก) ข้อมูลเชิงคุณภาพ (Categorial) ประกอบด้วย 2 แบบคือ แบบเชิงนาม (Nominal) หมายถึง ค่าของข้อมูลที่ไม่สามารถทำการจัดลำดับได้ เช่น สถานภาพสมรส (โสด แต่งงาน หม้าย) และแบบลำดับ (Ordinal) หมายถึงค่าของข้อมูลที่สามารถนำมาจัดลำดับได้ และลำดับของข้อมูลมีความสำคัญ เช่น อัตราการใช้บัตรเครดิตของลูกค้า (ดี ปานกลาง ไม่ดี)

ข) ข้อมูลเชิงปริมาณ (Quantitative) ประกอบด้วย 2 แบบ คือแบบต่อเนื่อง (Continuous) หมายถึงตัวแปรที่เป็นตัวเลขและค่าของข้อมูลมีความต่อเนื่อง มักแสดงเป็นตัวเลขที่เป็นจำนวนจริง เช่น ข้อมูลรายได้ และแบบไม่ต่อเนื่อง (Discrete) หมายถึง ตัวแปรที่เป็นตัวเลขและค่าของข้อมูลเป็นค่าที่ไม่ต่อเนื่องกัน มักแสดงเป็นตัวเลขที่เป็นจำนวนเต็ม เช่น ข้อมูลจำนวนลูกค้าในร้านค้าแห่งหนึ่ง

2.2) การเตรียมข้อมูล (Data Preprocessing) เป็นการนำข้อมูลที่เลือกไว้แล้ว มาทำการตรวจสอบอีกครั้งโดยใช้กระบวนการทางสถิติ เพื่อคัดเลือกให้เหลือเฉพาะข้อมูลถูกต้องและพร้อมสำหรับการทำดาต้าไมนิงเท่านั้น ซึ่งสิ่งที่ต้องทำการตรวจสอบคือ ความถูกต้องครบถ้วนของข้อมูล นั่นคือข้อมูลไม่มีการสูญหาย ไม่มีข้อมูลรบกวน (Noisy Data) หรือมีข้อมูลที่มีค่าผิดปกติแตกต่างไปจากที่ควรจะเป็น ซึ่งอาจเกิดจากการป้อนข้อมูลผิดพลาด เช่น บันทึกค่าส่วนสูงของนักเรียนเป็นค่าติดลบ

2.3) การแปลงข้อมูล (Data Transformation) เป็นการแปลงข้อมูลที่ต้องนำมาใช้ให้เหมาะกับอัลกอริทึม (Algorithm) แต่ละแบบของดาต้าไมนิง เช่น การปรับอัตราส่วนตัวเลขให้อยู่ในช่วง 0 – 1 เพื่อใช้กับอัลกอริทึมในโครงข่ายประสาทเทียม (Neural Network)

3) การทำดาต้าไมนิง (Data Mining) เป็นการประมวลผลข้อมูลตามอัลกอริทึมที่ได้กำหนดไว้ ในขั้นตอนนี้จะมีความสัมพันธ์กับการวิเคราะห์ข้อมูลและขั้นตอนที่ผ่านมา โดยเมื่อทำในส่วนของดาต้าไมนิงแล้วอาจจะต้องย้อนกลับไปทำในขั้นตอนที่ 2 ใหม่ ในการพัฒนาในส่วนของดาต้าไมนิงจะเกี่ยวข้องกับการใช้อัลกอริทึมหลาย ๆ แบบ ซึ่งการเลือกใช้อัลกอริทึมนั้นจะพิจารณาจากปัญหาและลักษณะของข้อมูล

4) การวิเคราะห์ผลลัพธ์ (Analysis of Result) เป็นการวิเคราะห์และแปลความหมายผลลัพธ์ที่ได้จากการทำดาต้าไมนิง ซึ่งถ้าไม่สามารถยอมรับได้ อาจแก้ไขโดยเลือกใช้ข้อมูลที่มีปริมาณมากขึ้น หรือเลือกใช้อัลกอริทึมอื่นแทน ซึ่งการทำดาต้าไมนิงนี้ต้องมีการทำซ้ำอยู่ตลอดเวลา เนื่องจากสิ่งแวดล้อมมีการเปลี่ยนแปลงอยู่เสมอ

5) การนำมาประยุกต์ใช้กับธุรกิจ (Assimilation of Knowledge) การนำผลลัพธ์ที่ได้มาประยุกต์ใช้ รวมกับส่วนความรู้ทางธุรกิจเพื่อนำไปใช้ให้เกิดประโยชน์ต่อไป

2.2 การวิเคราะห์อนุกรมเวลา (Time Series Analysis)

เป็นการพยากรณ์ค่าที่จะได้มาจากความสัมพันธ์ของค่าก่อนหน้า ลักษณะของข้อมูลที่แสดงความสัมพันธ์กับเวลานั้นโดยทั่วไปจะอยู่ในรูปดังนี้

$t_{(1)}$	$t_{(2)}$	$t_{(3)}$	...	$t_{(i-1)}$	$t_{(i)}$	$t_{(i+1)}$

โดยที่ $t_{(i)}$ เป็นค่าล่าสุดของข้อมูลในช่วงเวลาใด ๆ

$t_{(i+1)}$ เป็นค่าที่จะพยากรณ์

ในการวิเคราะห์ชุดของข้อมูลที่มีความสัมพันธ์กันในอนุกรมเวลานั้น เป็นการสร้างแบบจำลองอีกแบบหนึ่งที่มีการนำไปประยุกต์ใช้ในเชิงธุรกิจอย่างกว้างขวาง เนื่องจากมีการพิจารณาว่าค่าของข้อมูลหรือลักษณะต่าง ๆ ที่เกิดขึ้นกับข้อมูลนั้นมีรูปแบบเฉพาะที่มีความสัมพันธ์กันในเชิงเวลา ดังนั้นจึงมีการนำข้อมูลจากช่วงเวลาที่ผ่านมาก่อนหน้า มาทำการวิเคราะห์เพื่อให้สามารถหาวิธีพยากรณ์ค่าข้อมูลที่จะเกิดขึ้นในอนาคตได้ และเมื่อพิจารณาลักษณะของการวิเคราะห์ข้อมูลแบบที่สัมพันธ์กับเวลานี้แล้ว สามารถแบ่งได้เป็น 2 ลักษณะ

2.2.1 การวิเคราะห์โดยพิจารณาตัวแปรเดียว (Univariate time-series analysis) คือ เป็นการนำข้อมูลของ

ตัวแปรที่สนใจเพียงตัวเดียวมาใช้ในการวิเคราะห์ เช่น การพยากรณ์ปริมาณการสั่งใช้ยาในอนาคต โดยใช้ข้อมูลในอดีตของปริมาณการสั่งใช้ยาที่ต้องการพยากรณ์มาพิจารณาเพียงตัวแปรเดียว

2.2.2 การวิเคราะห์โดยพิจารณาจากหลายตัวแปร (Multivariate time-series analysis) เป็นการพัฒนาจากการพยากรณ์โดยใช้ตัวแปรเดียว โดยแทนที่จะพิจารณาเพียงตัวแปรเดียว ก็จะมีการนำตัวแปรอื่น ๆ ณ ช่วงเวลาเดียวกันมาพิจารณาร่วมด้วย

เมื่อพิจารณาจากวัตถุประสงค์ในการวิเคราะห์ข้อมูลลักษณะนี้แล้ว จะพบว่า ในบางครั้งไม่จำเป็นที่จะต้องสร้างแบบจำลองที่สามารถรู้ค่าที่แน่นอนในช่วงเวลาที่กำหนดในอนาคตได้อย่างชัดเจน แต่ต้องการใช้เพียงแค่สังเกตแนวโน้มของข้อมูลว่ามีโอกาสเพิ่มขึ้นหรือลดลงมากน้อยแค่ไหนเท่านั้น โดยใช้ข้อมูลในอดีตมาเป็นพื้นฐาน ซึ่งตัวอย่างของงานที่มีการวิเคราะห์ข้อมูลในลักษณะนี้ได้แก่

2.2.2.1 การพยากรณ์ค่าข้อมูลที่น่าจะเกิดขึ้นอย่างแน่นอนของตัวแปรหนึ่ง ๆ ณ ช่วงเวลาใด ๆ ที่กำหนดในอนาคต

2.2.2.2 การพยากรณ์ข้อมูลเพื่อดูแนวโน้มที่จะเกิดขึ้นในอนาคต

2.2.2.3 การพยากรณ์ค่าความเปลี่ยนแปลงของข้อมูลในช่วงเวลาใด ๆ

ดังนั้นการสร้างแบบจำลองจึงต้องพิจารณาจากวัตถุประสงค์ที่ต้องการเป็นสำคัญ สำหรับในการวิจัยนี้จะสร้างโปรแกรมเพื่อการพยากรณ์ปริมาณการสั่งใช้ยา ซึ่งเป็นลักษณะของหัวข้อที่ 2.2.2.1 คือเป็นการพยากรณ์ค่าของปริมาณการสั่งใช้ยาในช่วงเวลาถัดไปของอนุกรมเวลา โดยใช้ค่าข้อมูลปริมาณการสั่งใช้ยาในช่วงที่ผ่านมาระยะเวลาหนึ่งเป็นเกณฑ์ในการสร้างแบบจำลองให้กับโปรแกรม

2.3 โครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม (Neural Network) [5] คือฟังก์ชันประมาณค่าทางคณิตศาสตร์รูปแบบพิเศษ ที่มีวิวัฒนาการเริ่มต้นจากการพยายามที่จะเลียนแบบโครงสร้างและความเข้าใจในฟังก์ชันการทำงานของระบบสมองของมนุษย์ ถึงแม้ว่าโครงสร้างของโครงข่ายประสาทเทียมมีรูปแบบที่แตกต่างกัน แต่โดยทั่วไปจะประกอบขึ้นจากฟังก์ชันทางคณิตศาสตร์อย่างง่ายจำนวนมากมาเชื่อมต่อกันในลักษณะโครงข่ายที่เป็นชั้น โดยค่าแต่ละอินพุตของโครงข่ายจะผ่านเข้ามาในชั้นแรกสุดซึ่งเรียกว่า ชั้นอินพุต (Input Layer) ซึ่งแต่ละอินพุตของโครงข่ายจะเชื่อมต่อกันโดยการคูณด้วยพารามิเตอร์หรือที่เรียกว่า ตัวแปรน้ำหนัก (Weight) และบวกรวมกันทางคณิตศาสตร์ที่จุดรวม (node) โดยแต่ละจุดรวม จะประกอบด้วยฟังก์ชันกระตุ้น (Activation Function) ซึ่งโดยทั่วไปจะเป็นฟังก์ชันที่มีอนุพันธ์แบบต่อเนื่อง และผลลัพธ์ของแต่ละจุดรวมจะเป็นอินพุตของโครงข่ายชั้นถัดมา โดยแต่ละชั้นถัดมาเรียกว่า ชั้นซ่อน (Hidden Layer) เนื่องจากไม่มีการเชื่อมโยงโดยตรงกับอินพุตจากภายนอก และแต่ละชั้นซ่อน จะต่อเชื่อมกันในลักษณะทำนองเดียวกัน โดยชั้นสุดท้ายของโครงข่ายซึ่งเรียกว่า ชั้นเอาต์พุต (Output Layer) ให้ค่าผลลัพธ์ของโครงข่ายประสาทเทียมออกมา ซึ่งโครงข่ายประสาทเทียมที่มีโครงสร้างดังกล่าวนี้เรียกว่า Feedforward Multilayer Neural Network

การฝึกสอนในช่วงแรก ๆ ใช้ เพอร์เซพตรอนเพียง 1 ชั้น ซึ่งประสบความสำเร็จมาก แต่ต่อมาได้ประสบปัญหาที่ลักษณะของเพอร์เซพตรอน 1 ชั้น ไม่สามารถที่จะแก้ปัญหาเกี่ยวกับการเอกคลูซีฟออร์ (XOR Problem) ซึ่งปัญหานี้สามารถแก้ได้โดยใช้ลักษณะของเพอร์เซพตรอนหลายชั้น (Multilayer Perceptron) แต่วิธีการที่จะได้มาซึ่งค่าน้ำหนัก (Weight) ของตัวโครงข่ายเป็นลักษณะของการกำหนดค่าโดยใช้ความคิดขึ้นมาเอง ไม่ใช่ลักษณะของการฝึกสอน เนื่องจากสมัยนั้นยังไม่มีการคิดวิธีการฝึกสอนที่ได้ผลในส่วนของเพอร์เซพตรอนหลายชั้น

นอกจากนี้โครงข่ายประสาทเทียมยังถือว่าเป็นฟังก์ชันประมาณค่าประเภทโมเดลอิสระ (Free model) ซึ่งหมายความว่าโครงสร้างของโครงข่ายประสาทเทียม เช่น จำนวนจุดรวมและจำนวนชั้น เป็นต้น สามารถเลือกได้อิสระและโครงข่ายประสาทเทียมสองฟังก์ชันที่ประมาณค่าปัญหาเดียวกันและให้ค่าระดับความถูกต้องเท่ากันไม่จำเป็นต้องมีโครงสร้างเหมือนกันเสมอไป และโดยทั่วไป ถึงแม้ว่าโครงสร้างของโครงข่ายประสาทเทียมสองฟังก์ชันจะเหมือนกัน แต่ค่าของพารามิเตอร์หรือตัวแปรน้ำหนักที่เชื่อมต่อระหว่างชั้นอาจจะมีค่าแตกต่างกันได้

ค่าพารามิเตอร์หรือตัวแปรน้ำหนักของโครงข่ายประสาทเทียมได้มาจากขั้นตอนที่เรียกว่าการสอน ซึ่งเป็นการใช้ชุดข้อมูลการสอน (Training data) กฎการปรับแต่งค่าตัวแปรน้ำหนักและวิธีการสอน (Training approach) ที่เหมาะสมในการปรับแต่งค่าตัวแปรน้ำหนักจนกระทั่งโครงข่ายประสาทเทียม

จะมีพฤติกรรมให้ผลลัพธ์หรือผลตอบสนองตามที่ต้องการ

2.3.1 วิธีการสอนโครงข่ายประสาทเทียมตามลักษณะการเรียนรู้

2.3.1.1 การเรียนรู้แบบมีการควบคุม (Supervised Learning) เป็นการเรียนรู้ซึ่งต้องมีชุดข้อมูลสำหรับฝึกอบรม (Training Data) ซึ่งมีทั้งชุดข้อมูลที่เป็นอินพุตและเอาต์พุต การสอนโครงข่ายจะใช้ชุดข้อมูลในการสอน โดยในระหว่างการสอนนั้นโครงข่ายจะให้ผลลัพธ์จริงที่ได้จากการคำนวณ (Actual Output) และจะมีการเปรียบเทียบกันระหว่างผลลัพธ์จริงที่ได้จากการคำนวณกับชุดข้อมูลเอาต์พุตเป้าหมาย (Target Output) โดยผลต่างจากการเปรียบเทียบนั้นคือค่าความผิดพลาด หรือค่าความคลาดเคลื่อน ตัวอย่างอัลกอริทึมการเรียนรู้แบบ Supervised นิยมใช้สำหรับกรณีที่มีหลาย Layer และมีระบบการเชื่อมโยงแบบ Feed Forward คือ Back Propagation ซึ่งการเรียนรู้แบบแบ็คพรอพพาเกชันจะมีการป้อนค่าความผิดพลาดกลับเข้าสู่โครงข่ายเพื่อใช้ในการปรับค่าน้ำหนัก (Weight) ในการเชื่อมโยงระหว่าง Layer เพื่อเพิ่มประสิทธิภาพในการคำนวณของโครงข่าย โดยการฝึกสอนโครงข่ายจะสิ้นสุดเมื่อค่าความผิดพลาดอยู่ในระดับที่ยอมรับได้หรือน้อยกว่าค่า Error acceptance ($E_{\max}$) ที่กำหนดไว้

2.3.1.2 การเรียนรู้แบบไม่มีการควบคุม (Unsupervised Learning) การเรียนรู้แบบไม่มีการควบคุมเป็นการเรียนรู้โดยไม่ต้องอาศัยชุดข้อมูลเอาต์พุตเป้าหมาย (Target Output) จะมีเพียงชุดข้อมูลอินพุตให้กับโครงข่ายเท่านั้น แต่โครงข่ายจะสามารถปรับค่าน้ำหนัก และสามารถจัดกลุ่มข้อมูลตามความสัมพันธ์ของข้อมูลนั้นได้ การเรียนรู้แบบไม่มีการควบคุมนี้มักจะใช้กับงานจัดกลุ่ม (Clustering) เป็นการปรับตัวเองโดยไม่ต้องการความช่วยเหลือกจากภายนอก ไม่ต้องมีตัวอย่างผลลัพธ์เพื่อให้ Neurons ได้เรียนรู้ กระบวนการเรียนรู้แบบนี้เรียกว่า Learning by Doing

2.3.1.3 การเรียนรู้แบบเชิงบังคับ (Reinforcement Learning) การเรียนรู้แบบเชิงบังคับ เป็นการเรียนรู้ที่มีการควบคุมเป็นกรณีพิเศษคือ โครงข่ายจะมีตัวคริติค (Critic) เป็นตัวประเมินค่าให้กับเอาต์พุตแทนการกำหนดชุดข้อมูลเอาต์พุตเป้าหมาย (Target Output) โดยระหว่างการสอนโครงข่ายนั้น ตัวคริติคจะบอกความถูกต้องของผลลัพธ์ที่ได้ แต่ไม่ได้บอกว่าข้อมูลที่ถูกต้องคือค่าใด ตัวอย่างเช่น Genetic Algorithm [6]

2.4 งานวิจัยที่เกี่ยวข้อง

จรรยารัตน์ (2541) เปรียบเทียบประสิทธิภาพของการพยากรณ์ที่ได้จากโครงข่ายประสาทเทียมกับวิเคราะห์อนุกรมเวลาโดยใช้โครงข่ายประสาทเทียมที่มีโครงสร้างแบบมัลติเลเยอร์เพอร์เซพตรอนและการเรียนรู้แบบแบ็คพรอพพาเกชัน จำนวน 10 เน็ตเวิร์คโดยแต่ละเน็ตเวิร์คประกอบด้วยนิวรอนในชั้นอินพุตจำนวน 31 นิวรอนและในชั้นเอาต์พุตจำนวน 1 นิวรอน สำหรับพยากรณ์อุณหภูมิที่วัดจากไฮกรอมิเตอร์ ผลจากการพยากรณ์ปรากฏว่า โครงข่ายประสาทเทียมและการวิเคราะห์อนุกรมเวลามีความสามารถในการพยากรณ์อุณหภูมิกระเปาะแห้งได้ดีพอกันแต่การพยากรณ์อุณหภูมิกระเปาะเปียกนั้น โครงข่ายประสาทเทียมให้ผลการพยากรณ์ที่แม่นยำกว่าการวิเคราะห์อนุกรมเวลา

ทิพวรรณ์ [8] ได้ใช้วิธีการพยากรณ์เปรียบเทียบกัน 2 วิธี คือ การหาเส้นแนวโน้ม และการถดถอยแบบพหุ สิ่งที่ศึกษาคือเปรียบเทียบผลของการพยากรณ์จำนวนเลขหมายโทรศัพท์โดยใช้ตัวแปรคือ จำนวนประชากร จำนวนบ้าน จำนวนธุรกิจ จำนวนเลขหมายโทรศัพท์ และความหนาแน่นการใช้โทรศัพท์ และผลการพยากรณ์จำนวนครั้งที่เรียกโทรศัพท์ ซึ่งได้ผลคือ วิธีการถดถอยแบบพหุ จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงมากกว่าในการพยากรณ์จำนวนเลขหมายโทรศัพท์ทุกประเภท และจำนวนครั้งการเรียกโทรศัพท์ทุกประเภท ส่วนวิธีการหาเส้นแนวโน้ม (Trend method) เหมาะกับการพยากรณ์จำนวนเลขหมายโทรศัพท์แยกประเภท และจำนวนครั้งที่เรียกโทรศัพท์แยกประเภท ผู้วิจัยได้เสนอว่าเมื่อมีข้อมูลใหม่ ควรนำมาปรับปรุงการพยากรณ์ให้ทันสมัยยิ่งขึ้น

พัชรี [9] ได้ใช้วิธีการพยากรณ์เปรียบเทียบกัน 2 วิธีคือ ริดจ์รีเกรสชัน (Ridge regression) และเครือข่ายประสาทเทียม โดยข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โล (Monticaro) ด้วยโปรแกรมฟอร์แทรน 77 ในการวิเคราะห์ด้วยเครือข่ายประสาทเทียมใช้โปรแกรมสำเร็จรูป Trajan เวอร์ชัน 2.1 ทำการศึกษาภายใต้สถานการณ์ต่าง ๆ คือ กรณีการแจกแจงของความคลาดเคลื่อนเป็นปกติ การแจกแจงปลอมปน และการแจกแจงลอกนอร์มอล พบว่าวิธีเครือข่ายประสาทเทียมจะใช้ในการพยากรณ์ได้ดีกว่าวิธีริดจ์รีเกรสชัน เมื่อความคลาดเคลื่อนมีการแจกแจงลอกนอร์มอล พบว่าวิธีเครือข่ายประสาทเทียมจะใช้ในการพยากรณ์ได้ดีกว่าวิธีริดจ์รีเกรสชันเมื่อความคลาดเคลื่อนมีการแจกแจงลอกนอร์มอล และการแจกแจงปกติปลอมปน

เรียงตามลำดับจากมากไปน้อย และเมื่อขนาดตัวอย่างระดับสัมประสิทธิ์การแปรผันของความคลาดเคลื่อน จำนวนตัวแปรอิสระ ระดับความสัมพันธ์ของตัวแปรอิสระ สเกลแฟคเตอร์ และเปอร์เซ็นต์การปลอมปนมีค่ามากขึ้น โดยเรียงลำดับอิทธิพลจากมากไปน้อย และวิธีริดจ์รีเกรสชัน จะใช้ในการพยากรณ์ได้ดีกว่าเมื่อความคลาดเคลื่อนมีการแจกแจงปกติ และได้เสนอว่าควรทำการศึกษาข้อมูลที่ไม่ใช้ตัวแบบเชิงเส้น

จากงานวิจัยต่าง ๆ ผู้วิจัยได้พบปัญหาหลักที่สำคัญอยู่สองประการคือ ประการแรก การพยากรณ์ของทิพวรรณ์ [8] เป็นการพยากรณ์แบบการศึกษาความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม คือ ถ้าต้องการพยากรณ์ข้อมูลตัวแปรตามในปีถัดไปจำเป็นต้องทราบข้อมูลตัวแปรต้นในปีถัดไปก่อน แต่ในความเป็นจริงนั้นเราไม่สามารถทราบข้อมูลตัวแปรต้นล่วงหน้าได้ จึงทำการออกแบบรูปแบบของตัวแปรต้นและตัวและตัวแปรตามใหม่ให้สามารถพยากรณ์อนาคตได้โดยทราบเพียงข้อมูลในปัจจุบัน ประการที่สอง เมื่อเวลาผ่านไปข้อมูลได้มีการเปลี่ยนแปลงไปมาก ซึ่งจะส่งผลต่อโมเดลการพยากรณ์เดิม สำหรับงานวิจัยนี้ได้ทำการปรับพยากรณ์ให้เป็นการพยากรณ์ล่วงหน้าแบบวัน เดือน หรือปี โดยใช้ข้อมูลในปัจจุบัน และข้อมูลที่ได้นำมาใช้สำหรับงานวิจัยนี้พบว่ามีลักษณะไม่เป็นเชิงเส้น ซึ่งเป็นการศึกษาตามข้อเสนอแนะที่ต้องการให้ศึกษาต่อของ พัชรี [9] และนำโครงข่ายประสาทเทียมมาสร้างโมเดลในการพยากรณ์ ตามข้อเสนอของจรรยารัตน์ [7]

3. วิธีการดำเนินงาน

ผู้วิจัยได้แบ่งวิธีการดำเนินงานและพัฒนาโปรแกรมออกเป็น 6 ขั้นตอน ซึ่งแต่ละขั้นตอนมีรายละเอียดดังต่อไปนี้

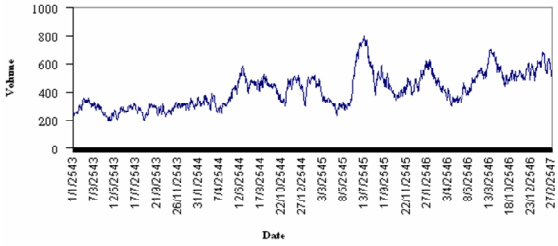
3.1 ศึกษาโครงข่ายประสาทเทียมและเทคนิควิธีแบ็คพรอพพาเกชัน

จากการทำการศึกษาโครงข่ายประสาทเทียมมัลติเลเยอร์เพอร์เซพตรอนโดยเทคนิควิธีแบ็คพรอพพาเกชัน พบว่าเป็นเทคนิคที่มีการนำไปประยุกต์ใช้หลายด้าน ทั้งการพยากรณ์ประมาณค่าเชิงปริมาณและจัดหมวดหมู่ โดยจะสามารถกำหนดโครงสร้างของโครงข่ายประสาทเทียมซึ่งจะมีจำนวนชั้นซ่อนตั้งแต่ 1 ชั้นขึ้นไป และแต่ละชั้นจะมีกี่โหนดหรือกี่นิวรอนก็ได้ ส่วนการคำนวณหาค่าเอาต์พุตจะคำนวณจากฟังก์ชันต่าง ๆ ซึ่งได้แก่ ฟังก์ชันเชิงเส้น (Linear) และฟังก์ชันไม่เชิงเส้น (Non Linear) ซึ่งแต่ละฟังก์ชันนั้นก็จะมีวิธีการหาเอาต์พุตและการปรับแต่งค่าน้ำหนัก (Weight) ที่แตกต่างกันด้วย

3.2 ทำการทดลองหาโมเดลด้วยโปรแกรมแมตแล็บ

3.2.1 วิเคราะห์ลักษณะของชุดข้อมูล

ข้อมูลที่ใช้เป็นข้อมูลปริมาณการใช้ยาของตัวยา Amoxy ขนาด 500 มิลลิกรัม ในโรงพยาบาลแห่งหนึ่ง มีจำนวน 1,521 เรคคอร์ด ตั้งแต่วันที่ 1 มกราคม พ.ศ. 2543 ถึงวันที่ 2 กุมภาพันธ์ พ.ศ. 2546 เป็นข้อมูลอนุกรมเวลามีลักษณะไม่เป็นเชิงเส้น และมีความเคลื่อนไหวไม่หยุดนิ่งดังภาพที่ 1 และทำการแบ่งชุดข้อมูลออกเป็นสองส่วนดังนี้ 1) ชุดข้อมูลในการเรียนรู้ 80% 2) ชุดข้อมูลในการทดสอบ 20%



ภาพที่ 1 ลักษณะของข้อมูลปริมาณการใช้ยา Amoxy ขนาด 500 มิลลิกรัม

3.2.2 การแปลงค่าข้อมูล (Data Transformation)

เป็นกระบวนการในการปรับขอบเขตของข้อมูลให้อยู่ในช่วงที่เหมาะสมต่อการนำไปใช้งานในการสอนให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ สำหรับในโมเดลที่ใช้หลักการของโครงข่ายประสาทเทียมนั้น กระบวนการแปลงค่าข้อมูลที่ต้องนำมาใช้คือ การนอมัลไลซ์ข้อมูล (Normalization) ซึ่งเป็นการลดค่าของข้อมูลให้อยู่ในขอบเขตที่น้อยลงเพื่อให้เหมาะสมกับฟังก์ชันที่ใช้งานของโครงข่ายประสาทเทียม ซึ่งวิธีการนอมัลไลซ์ข้อมูลนั้นมีหลายวิธี แต่ที่มีการนำมาใช้กันอย่างกว้างขวางได้แก่ การแปลงค่าข้อมูลในลักษณะเป็นเชิงเส้น (Min-max Normalization) ดังสมการที่ 1

$$V' = \frac{V - \min_{old}}{\max_{old} - \min_{old}} \qquad (1)$$

โดยที่ V' คือ ค่าของข้อมูลที่ได้หลังจากผ่านสมการ

V คือ ค่าของข้อมูลก่อนผ่านสมการ

$\min_{old}$ คือ ค่าของข้อมูลที่มีค่าต่ำที่สุดก่อนผ่านสมการ

$\max_{old}$ คือ ค่าของข้อมูลที่มีค่าสูงที่สุดก่อนผ่านสมการ

3.2.3 จัดเตรียมข้อมูลเพื่อสร้างรูปแบบของอินพุตและเอาต์พุต (Input and Output Pattern) สำหรับ time series

วิธีการโดยทั่วไปในการสร้างรูปแบบของอินพุตและเอาต์พุต สำหรับข้อมูลอนุกรมเวลา คือการแบ่งหน้าต่างย่อย (Windowing) และทำการคำนวณแบบเลื่อนหน้าต่าง (Sliding windows) โดยแบ่งเป็นหน้าต่างย่อยตามจำนวนวัน จำนวนเดือน หรือจำนวนปี และให้มีการซ้อนทับกัน ซึ่งมีวิธีการดังสมการที่ 2 และ 3

$$X_1 = \begin{bmatrix} x_{1,i-L-1} \\ \vdots \\ x_{1,i-2} \\ x_{1,i-1} \\ x_{1,i} \end{bmatrix} \quad X_2 = \begin{bmatrix} x_{2,i-L-1} \\ \vdots \\ x_{2,i-2} \\ x_{2,i-1} \\ x_{2,i} \end{bmatrix} \qquad (2)$$

$$Y_1 = \begin{bmatrix} y_{1,i+1} \\ y_{1,i+2} \\ y_{1,i+3} \\ \vdots \\ y_{1,i+n} \end{bmatrix} \quad Y_2 = \begin{bmatrix} y_{2,i+1} \\ y_{2,i+2} \\ y_{2,i+3} \\ \vdots \\ y_{2,i+n} \end{bmatrix} \quad \cdots \qquad (3)$$

$$p_j = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_z \end{bmatrix} \text{ โดยที่ } \quad j = 1,2,3,...,Q \ ; \qquad (4)$$

$$t_k = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_z \end{bmatrix} \quad k = 1,2,3...,Q$$

เพราะฉะนั้นชุดข้อมูลอินพุตและเอาต์พุตที่ได้คือ

$P = [p_1 \ \ p_2 \ \ p_3 ... \ p_L]$ และ $T = [t_1 \ \ t_2 \ \ t_3 ... \ t_N]$

หมายเหตุ

L คือ จำนวนข้อมูลหน้าต่าง (Window length)

N คือ จำนวนเอาต์พุตที่พยากรณ์

Q คือ รูปแบบข้อมูล (Patterns)

i คือ ลำดับที่ของข้อมูล

3.2.4 การออกแบบโครงข่ายประสาทเทียม

ก) ใช้โครงข่ายประสาทเทียมชนิด Feed-forward backpropagation เนื่องจากสามารถที่จะเรียนรู้ได้โดยการปรับค่าน้ำหนักเพื่อให้ค่าความคลาดเคลื่อนระหว่างค่าเอาต์พุตกับค่าเป้าหมายลดลง

ข) ใช้การสอน (Training function) แบบ TrainLM (Levenberg Marquardt Algorithm) เนื่องจากประมวลผลได้ค่อนข้างเร็ว แต่มีความต้องการหน่วยความจำมากพอสมควร

ค) ใช้การเรียนรู้ (Learning function) แบบ LearnGDM (Grad. descent w/momentum weight/bias learning function)

ง) ใช้ Mean square error (MSE) วัดความผิดพลาดเฉลี่ยกำลังสอง

จ) ใช้จำนวน Layer 2 ชั้น ในการหาเอาต์พุต

ฉ) ใช้ทรานส์เฟอร์ฟังก์ชัน binary sigmoid (logsig) ในชั้นซ่อน (Hidden Layer) และ linear (purelin) ในชั้นเอาต์พุต (Output Layer)

ช) รูปแบบโครงสร้างของชั้นอินพุต ชั้นซ่อน และชั้นเอาต์พุตเขียนอยู่ในรูปแบบของ จำนวนอินพุต –จำนวนโหนดในชั้นซ่อน – จำนวนเอาต์พุต

การทดลองในครั้งนี้มีโครงสร้างและรูปแบบของข้อมูลที่ใช้ในการทดลองดังนี้

1) 3-5-1 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 2 จุดข้อมูล
2) 3-5-2 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 2 จุดข้อมูล
3) 4-5-1 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 3 จุดข้อมูล
4) 4-5-2 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 3 จุดข้อมูล
5) 5-5-1 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 4 จุดข้อมูล
6) 5-5-2 โดยมีการซ้อนทับของหน้าต่าง (Overlap) 4 จุดข้อมูล

3.3 ออกแบบโปรแกรม

การออกแบบโปรแกรมมีขั้นตอนทั้งหมดดังแสดงในแผนผังการทำงานของโปรแกรม ภาพที่ 2

3.4 วางแผนและออกแบบหน้าจอโปรแกรม

โปรแกรมจะประกอบด้วย 2 ส่วนดังนี้

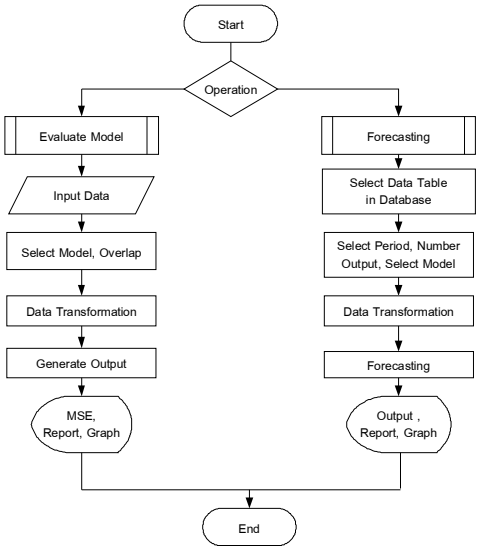
3.4.1 ส่วนของการประเมินประสิทธิภาพของโมเดล (Evaluate Model)

3.4.2 ส่วนของการพยากรณ์ (Forecasting)

3.5 พัฒนาโปรแกรม

เมื่อทำการวางแผนและออกแบบหน้าจอแล้ว ขั้นต่อมาคือ

ขั้นตอนของการพัฒนาโปรแกรมโดยใช้โปรแกรมไมโครซอฟท์วิชวลเบสิก 6.0 (Microsoft Visual Basic 6.0) เป็นเครื่องมือสำหรับพัฒนาโปรแกรม ซึ่งสามารถเขียนเป็นลำดับได้ดังนี้



ภาพที่ 2 แผนผังการทำงานของโปรแกรม

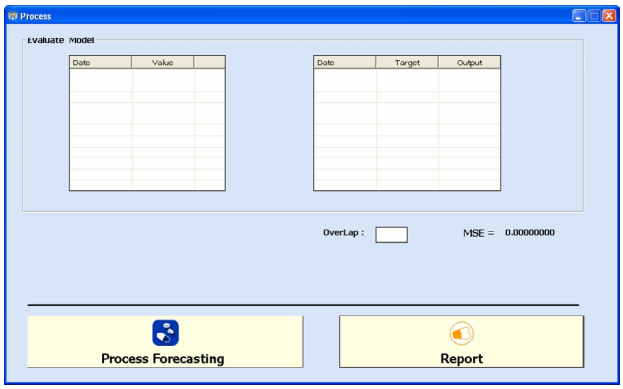
3.6 ทดสอบและปรับปรุงแก้ไขโปรแกรม

เมื่อพัฒนาโปรแกรมเสร็จเรียบร้อยแล้ว จึงทำการทดสอบความถูกต้องของโปรแกรมโดยใช้ข้อมูล 1,521 เรคคอร์ดที่จัดเตรียมไว้ทดสอบกับโมเดลต่าง ๆ ที่ได้กำหนดในหัวข้อที่ 3.4

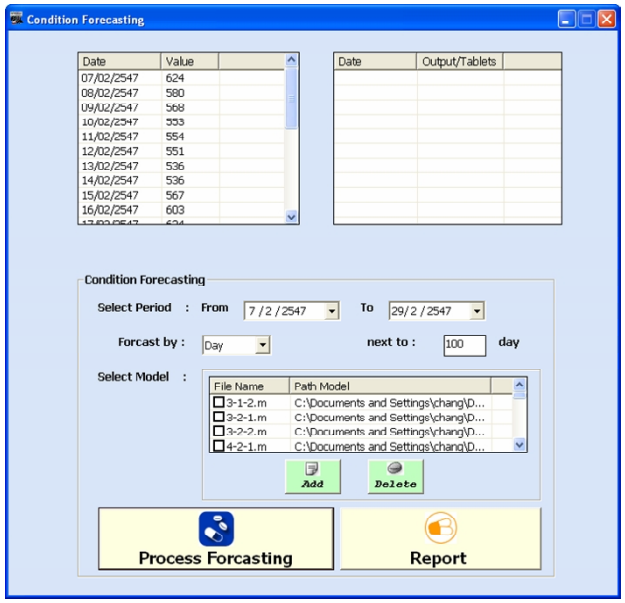
4. ผลการดำเนินงาน

4.1 ผลการพัฒนาโปรแกรมพยากรณ์ปริมาณการสั่งใช้ยา

ระบบที่พัฒนาขึ้นมามีตัวอย่างหน้าจอแสดงดังภาพที่ 3 และภาพที่ 4



ภาพที่ 3 หน้าจอสำหรับแสดงชุดข้อมูลที่นำเข้าและเปรียบเทียบค่าเอาต์พุตกับค่าเป้าหมาย



ภาพที่ 4 หน้าจอสำหรับการพยากรณ์

4.2 ผลการทดลองที่ได้

ทำการทดลองโดยกำหนดให้โครงข่ายมีจำนวน 2 ชั้น โดยนับจากชั้นซ่อนไปยังชั้นเอาต์พุต ใช้ฟังก์ชันส่งต่อแบบ Logsig ในชั้นซ่อน และใช้ฟังก์ชันส่งต่อแบบ Linear ในชั้นเอาต์พุต กำหนดให้มีโหนดในชั้นซ่อนจำนวน 5 โหนด และมีอัตราการเรียนรู้ (Learning rate) 0.1 ป้อนชุดข้อมูลการสอนให้โครงข่ายเพื่อทำการเรียนรู้ 80% และป้อนชุดข้อมูลการทดสอบเพื่อทำการทดสอบโครงข่าย 20% ทำการสอนโดยกำหนดค่าความผิดพลาดเป้าหมายที่จะให้โครงข่ายเรียนรู้ไปจนถึงที่ 0.001

หลังจากทำการสร้างโมเดลสำหรับพยากรณ์ปริมาณการใช้ยาแบบรายวันและรายสัปดาห์แต่ละรูปแบบแล้ว จึงนำโมเดลที่ได้มาประเมินประสิทธิภาพด้วยโปรแกรม ซึ่งได้ผลการเปรียบเทียบประสิทธิภาพของโมเดลดังตารางที่ 1 และ 2

ตารางที่ 1 ผลการเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ข้อมูลรายวัน

จำนวนอินพุต	จำนวนโหนดอินพุต	จำนวนโหนดเอาต์พุต	จำนวนซ้อนทับ (Overlap)	ค่าความผิดพลาด (MSE)
3	5	1	2	0.011
3	5	2	2	0.019
4	5	1	3	0.014

ตารางที่ 1 ผลการเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ข้อมูลยารายวัน (ต่อ)

จำนวนอินพุต	จำนวนโหนดอินพุต	จำนวนโหนดเอาต์พุต	จำนวนซ้อนทับ (Overlap)	ค่าความผิดพลาด (MSE)
4	5	2	3	0.112
5	5	1	4	0.103
5	5	2	4	0.107

ผลการเปรียบเทียบประสิทธิของโมเดลการพยากรณ์ปริมาณการใช้ยาแบบรายวัน โดยกำหนดให้มีโหนดในชั้นซ่อนจำนวน 5 โหนด ในกรณีที่มีอินพุตจำนวน 3 อินพุต มีการซ้อนทับของหน้าต่าง 2 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุตจะให้ค่าความผิดพลาด 0.011 ซึ่งน้อยกว่าค่าความผิดพลาดเมื่อพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต ในกรณีที่มีอินพุตจำนวน 4 อินพุตมีการซ้อนทับของหน้าต่าง 3 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุตจะให้ค่าความผิดพลาด 0.014 ซึ่งน้อยกว่าค่าความผิดพลาดเมื่อพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต และในกรณีที่มีอินพุตจำนวน 5 อินพุต มีการซ้อนทับของหน้าต่าง 4 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุตจะให้ค่าความผิดพลาด 0.111 ซึ่งน้อยกว่าการพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต และจากการเปรียบเทียบประสิทธิภาพของโมเดล ค่าความผิดพลาดของโมเดลที่ทำการพยากรณ์ปริมาณการใช้ยาแบบรายวัน พบว่า โมเดลที่มีอินพุตจำนวน 3 อินพุต จำนวนซ้อนทับของหน้าต่าง 2 จุด และจำนวนเอาต์พุต 1 เอาต์พุต ให้ค่าความผิดพลาด (MSE) น้อยที่สุดที่ 0.011

ผลการเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ปริมาณการใช้ยาแบบรายสัปดาห์ โดยกำหนดให้มีโหนดในชั้นซ่อนจำนวน 5 โหนด ในกรณีที่มีจำนวนอินพุตจำนวน 3 อินพุต มีการซ้อนทับของหน้าต่าง 2 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุตจะให้ค่าความผิดพลาด 0.052 ซึ่งน้อยกว่าค่าความผิดพลาดเมื่อพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต ในกรณีที่มีอินพุตจำนวน 4 อินพุต มีการซ้อนทับของหน้าต่าง 3 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุตจะให้ค่าความผิดพลาด 0.058 ซึ่งน้อยกว่าค่าความผิดพลาดเมื่อพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต และในกรณีที่มีอินพุตจำนวน 5 อินพุต มีการซ้อนทับของหน้าต่าง 4 จุด การพยากรณ์เอาต์พุตจำนวน 1 เอาต์พุต จะให้ค่าความผิดพลาด 0.103 ซึ่งน้อยกว่าการพยากรณ์เอาต์พุตจำนวน 2 เอาต์พุต และจากการเปรียบเทียบประสิทธิภาพของโมเดล ค่าความผิดพลาดของโมเดลที่ทำการพยากรณ์ปริมาณการใช้ยาแบบรายสัปดาห์ พบว่า โมเดลที่มีอินพุตจำนวน 3 อินพุต จำนวนซ้อนทับของหน้าต่าง 2 จุด และจำนวนเอาต์พุต 1 เอาต์พุต ให้ค่าความผิดพลาด (MSE) น้อยที่สุดที่ 0.052

ตารางที่ 2 ผลการทดสอบประสิทธิภาพของการพยากรณ์ข้อมูลยารายสัปดาห์ของโมเดล

จำนวนอินพุต	จำนวนโหนดอินพุต	จำนวนโหนดเอาต์พุต	จำนวนซ้อนทับ (Overlap)	ค่าความผิดพลาด (MSE)
3	5	1	2	0.052
3	5	2	2	0.059
4	5	1	3	0.058
4	5	2	3	0.112
5	5	1	4	0.103
5	5	2	4	0.107

5. สรุปผลและข้อเสนอแนะ

จากการพัฒนาโปรแกรมการพยากรณ์ปริมาณการใช้ยา ซึ่งใช้ข้อมูลปริมาณการใช้ยาในอดีตมาทำการทดลองหาโมเดลที่เหมาะสม โดยโปรแกรมจะทำการนำเข้าโมเดลที่สร้างจากโปรแกรมแมตแล็บเข้ามายังโปรแกรมที่พัฒนาขึ้น เพื่อทำการพยากรณ์ปริมาณการใช้ยาในอนาคตได้ตามต้องการ

5.1 สรุปผล

จากการเปรียบเทียบประสิทธิภาพของโมเดลทั้งการพยากรณ์ปริมาณการใช้ยาแบบรายวันและรายสัปดาห์ ซึ่งทำการทดลองโมเดลแบบ 3-5-1 และ 3-5-2 มีการซ้อนทับของหน้าต่าง 2 จุด โมเดลแบบ 4-5-1 และ 4-5-2 มีการซ้อนทับของหน้าต่าง 3 จุด โมเดลแบบ 5-5-1 และ 5-5-2 มีการซ้อนทับของหน้าต่าง 4 จุด ผลจากการเปรียบเทียบประสิทธิภาพของโมเดล การพยากรณ์ปริมาณการใช้ยาแบบรายวันนั้น โมเดล 3-5-1 ซึ่งมีอินพุตจำนวน 3 อินพุต โหนดในชั้นซ่อนจำนวน 5 โหนด จำนวนเอาต์พุต 1 เอาต์พุต และมีการซ้อนทับของหน้าต่าง 2 จุด ให้ค่าความผิดพลาด (MSE) น้อยที่สุดที่ 0.011

และผลจากการเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ปริมาณการใช้ยาแบบรายสัปดาห์นั้นโมเดล 3-5-1 ซึ่งมีอินพุตจำนวน 3 อินพุต โหนดในชั้นซ่อนจำนวน 5 โหนด จำนวนเอาต์พุต 1 เอาต์พุต และ มีการซ้อนทับของหน้าต่าง 2 จุด ให้ค่าความผิดพลาด (MSE) น้อยที่สุดที่ 0.052 จึงสามารถสรุปได้ว่าการพยากรณ์ปริมาณการใช้ยาแบบรายวันมีความถูกต้องมากกว่าการพยากรณ์ปริมาณการใช้ยาแบบรายสัปดาห์

5.2 อภิปรายผลที่ได้

การพยากรณ์ปริมาณการใช้ยาแบบรายวันมีความถูกต้องมากกว่าเนื่องจากข้อมูลรายวันเป็นข้อมูลที่มีความถี่และความละเอียดกว่าข้อมูลรายสัปดาห์ ซึ่งการพยากรณ์ข้อมูลอนุกรมเวลา โดยใช้วิธีการกำหนดความกว้างของหน้าต่างและใช้เทคนิคการเลื่อนหน้าต่าง (Sliding Window) โดยมีการจัดรูปแบบข้อมูล หรือการซ้อนทับของหน้าต่าง จะช่วยให้ข้อมูลมีความเป็นเชิงเส้นมากขึ้น เนื่องจากข้อมูลยามีลักษณะไม่เป็นเชิงเส้น เหมาะกับการนำโครงข่ายประสาทเทียมมาใช้ ซึ่งข้อดีของโครงข่ายประสาทเทียมคือ มีความยืดหยุ่น สามารถพยากรณ์ได้ทั้งข้อมูลที่เป็นเชิงเส้นและไม่เป็นเชิงเส้น โดยมีการเรียนรู้ที่เลียนแบบสมองของมนุษย์ และมีการปรับปรุงการเรียนรู้ได้ด้วยตัวเองเพื่อให้ได้ค่าความผิดพลาดที่น้อยลง แต่ทั้งนี้จะขึ้นอยู่กับค่าพารามิเตอร์ที่กำหนดในการสร้างแบบจำลอง เช่น ฟังก์ชันส่งต่อ อัลกอริทึมที่ใช้ในการสอน ฟังก์ชันการเรียนรู้ จำนวนข้อมูล จำนวนรอบในการสอน และค่าความคลาดผิดพลาดเป้าหมายที่จะให้โครงข่ายเรียนรู้ไปจนถึงเป้าหมาย แต่ข้อเสียโครงข่ายประสาทเทียมคือไม่สามารถที่จะอธิบายกระบวนการทำงานด้วยเหตุผลได้

5.3 ข้อเสนอแนะ

5.3.1 ในการพยากรณ์ครั้งนี้เป็นการนำค่าตัวแปรของปริมาณการใช้ยามาพิจารณาเพียงแค่ตัวแปรเดียวโดยไม่ได้พิจารณาถึงตัวแปรอื่น สำหรับการปรับปรุงการพยากรณ์ในครั้งต่อไปอาจนำปัจจัยหรือตัวแปรอื่นที่ส่งผลต่อปริมาณการใช้ยามาพิจารณาร่วมด้วย โดยลดความคลาดเคลื่อนลงจากเดิม

5.3.2 ปรับปรุงในส่วนของโปรแกรมให้สามารถเรียนรู้ได้เมื่อประสิทธิภาพของการพยากรณ์ลดลงถึงระดับหนึ่ง

5.3.3 เนื่องจากโมเดลการพยากรณ์ที่ทำเสร็จในครั้งนี้อาจล้าสมัยและนำไปใช้ไม่ได้ จึงควรติดตามหาข้อมูลและปรับปรุงโมเดลให้ทันสมัยอยู่เสมอ

เอกสารอ้างอิง

[1] โกศล ดีศีลธรรม. **"คลังข้อมูลสำหรับการสนันสนุนการตัดสินใจ."** กรุงเทพฯ. Industrial Technology Review, ฉบับที่ 105, มกราคม 2546.

[2] Aubrey E. Hill, Warren T.Jones, **Retrospective Case Base Browsing : A Data Mining Process Enhancement.** ACM Southeast Regional Conference, 1999.

[3] Cristina Olaru, Louis Wehenkel, University of Liege. **Data mining.** IEEE Computer Application in Power, 1999: 19-25.

[4] U. Fayyad, G. P.-Shapiro, and P. Smyth. **Knowledge Discovery and Data Mining: Towards a Unifying Framework**." Proc. Of the Second Intl. Conf. On Knowledge Discovery and Data Mining (KDD-96), Aug. 1996.

[5] R.Hecht-Nielson. **Neurocomputing.** Addison-Wesley Pub.1990.

[6] Simon Haykin. **Neural Network a Comprehensive Foundation** : Prentice-Hall International Inc. USA, 1998.

[7] จรรยารัตน์ พฤกษานันท์. "**การประยุกต์ใช้นิวรอลเน็ตเวิร์คในการพยากรณ์.** มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2541.

[8] ทิพวรรณ์ วุฒิสาร. **"การพยากรณ์ปริมาณการใช้โทรศัพท์ในเขตนครหลวง."** วิทยานิพนธ์ปริญญาพาณิชยศาสตร์มหาบัณฑิต, แผนกวิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2517.

[9] พัชรี คุณะสารพันธ์. **"การเปรียบเทียบวิธีการพยากรณ์ในการวิเคราะห์ความถดถอยพหุคูณโดยใช้วิธีริดจ์รีเกรสชันและวิธีที่ใช้หลักการของโครงข่ายประสาทเทียมในกรณีที่เกิดพหุสัมพันธ์ระหว่างตัวแปรอิสระ."** วิทยานิพนธ์ปริญญาสถิติศาสตรมหาบัณฑิต, สาขาวิชาสถิติ ภาควิชาสถิติ จุฬาลงกรณ์มหาวิทยาลัย, 2541.