การประยุกต์ใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น เพื่อการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปี*

เกียรติศักดิ์ จันทร์แก้ว** สุพจน์ นิตย์สุวัฒน์***

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อสร้างโมเดลการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีของประเทศ โดยใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Multilayer Perceptron Polynomial Artificial Neural Network) ทำการเปรียบเทียบกับโครงข่ายประสาทแบบธรรมดา ใช้ข้อมูลอนุกรมเวลาของกรมสถิติแห่งชาติ ตั้งแต่ปี พ.ศ. 2544 – 2550 ผู้วิจัยได้ทำการแบ่งข้อมูลออกเป็น 2 ขบวนการคือ การเรียนรู้ชุดข้อมูล (Training) และการทดสอบ (Testing) ทำการสอนให้เกิดการเรียนรู้ด้วยเทคนิคการแพร่กระจายย้อนกลับ ฝึกสอนด้วยวิธี Levenberg-Maquardt algorithm (trainlm) และฟังก์ชันการปรับการเรียนรู้ (Learning Function) แบบ LearnD (Grad Descent) โดยใช้ค่าเฉลี่ยกำลังสองสัมบูรณ์ (Mean Square Error: MSE) ของผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมทั้งสองชนิดเป็นตัวชี้วัด ผลการวิจัยพบว่าโครงสร้างที่เหมาะสมที่สุดที่ใช้ในการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีคือ 10-10-1 ฟังก์ชันกระตุ้นคือ logsig tansig purelin โครงข่ายประสาทเทียมให้ค่าความผิดพลาดในการเรียนรู้ชุดข้อมูลเท่ากับ 0.0033 และค่าความผิดพลาดในการทดสอบ (Testing) เท่ากับ 0.0297 และโครงข่ายประสาทเทียมแบบพหุนามหลายชั้นให้ค่าความผิดพลาดในการเรียนรู้ชุดข้อมูลเท่ากับ 0.0010 และค่าความผิดพลาดในการทดสอบ (Testing) เท่ากับ 0.0259 ซึ่งแสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Multilayer Perceptron Polynomial Artificial Neural Network) เป็นโมเดลการพยากรณ์ที่ให้ค่าความผิดพลาดต่ำกว่า เมื่อทดสอบกับชุดข้อมูลจริง และเหมาะกับการนำไปใช้เพื่อการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีได้อย่างมีประสิทธิภาพ

---

* เป็นผลงานจากการประชุมและนำเสนอผลงานวิชาการ เทคโนโลยีภาคใต้วิจัย "สหวิทยาการเพื่อการพัฒนาอย่างยั่งยืน" เมื่อวันที่ 21 มกราคม 2554

** นักศึกษาปริญญาโท สาขาวิชาระบบสารสนเทศเพื่อการจัดการ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

*** ผู้ช่วยศาสตราจารย์ ดร. ภาควิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

คำสำคัญ : โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น การแพร่กระจายย้อนกลับ การพยากรณ์ราคา

The Application of a Predictions Model for Amount of Rice Crop Exports by Multilayer Perceptron Polynomial Neural network *

Keattisak Chankaew ** **Supot Nitsuwat** ***

Abstract

This study aimed to create forecasting models for predicting the exported amount of rice crop. The model applied Multilayer Perceptron Polynomial Artificial Neural Networks (MLPPANN, PANN) and Artificial Neural Networks (ANN). The data set from The National Department of Statistics since 2001 - 2007. This data were divided into two sets, learning and testing sets. In eaching step, the system was taught using the Levenberg-Maquardt algorithm, a variance of the back-propagation techniques. Learning function used in this step was the Grad Descent function (LearnD). We compared performance of our network with the conventional neural network using Mean square error (MSE). The results shown that the appropriate neural network structure for predicting the exported amount of rice crop was the 10-10-1 structure – using the logsig, tansig and purelin activated functions. The MSE for learning and testing steps of MLPPANN were 0.0010 and ANN ware 0.0259, respectively. Using the same data sets,

---

* The Work from Meetings and Presentations Scholarly Works Research of Southern Technology “Interdisciplinary for Sustainable Development” January 21, 2011

** Graduate students, Major of Management Information System, King Mongkut’s University of Technology North Bangkok

*** Asst. Prof. Dr., Faculty of Information Technology, King Mongkut’s University of Technology North Bangkok

The MSE for learning and testing steps of conventional ANN were 0.0033 and 0.0297, respectively. This indicated that the MLPPANN was the best model for this time series data with lowest MSE.

Key words : Multilayer Perceptron Polynomial Artificial Neural Network, Back Propagation, Predictions

บทนำ

ปัจจุบันระบบเศรษฐกิจได้เข้าสู่ระบบการค้าเสรีมากขึ้น โดยการเน้นการลดอุปสรรคทางการค้า การยกเลิกนโยบายและมาตรการช่วยเหลืออุดหนุนภาคเกษตร และเปิดตลาดในประเทศต่าง ๆ มากขึ้น ในระบบเศรษฐกิจเสรีราคายังคงมีบทบาทสำคัญในการตัดสินใจเกี่ยวกับปัจจัยทางด้านเศรษฐกิจ เช่น การบริโภค การจำหน่าย ของสินค้าต่างๆ เพราะราคาจะช่วยจัดสรรทรัพยากรที่มีอยู่จำกัดในระบบเศรษฐกิจ เพื่อบำบัดความต้องการที่มีไม่สิ้นสุด โดยปกติราคาดุลยภาพเป็นราคาที่ไม่ได้เกิดขึ้นง่าย ๆ อีกทั้งยังไม่สามารถอยู่คงที่ได้นาน แต่จะมีการเคลื่อนไหวตลอดเวลาตามอิทธิพลการเปลี่ยนแปลงของอุปสงค์และอุปทานในตลาด เพื่อเข้าไปหาจุดที่เป็นดุลยภาพใหม่ ราคาที่เสถียรภาพทำให้การใช้ทรัพยากรมีประสิทธิภาพและผู้บริโภคจะได้รับความพึงพอใจสูงสุด

ข้าวเจ้านาปีถือเป็นพืชเศรษฐกิจที่มีความสำคัญของประเทศ ทั้งในด้านเศรษฐกิจ สังคม สภาพแวดล้อม หากพิจารณาแล้วการส่งออกข้าวเจ้านาปีนั้นยังก่อให้เกิดรายได้เข้าสู่ประเทศจากการส่งออก อีกทั้งการส่งออกสินค้าถือเป็นเครื่องมือสำคัญในการตัดสินใจเลือกผลิตสินค้า ช่วยในการวางแผนการส่งออกทั้งในระยะสั้นและระยะยาว อีกทั้งการส่งออกข้าวในปีที่ผ่านมา ยังสามารถนำมาวางแผนการผลิต การพยากรณ์หรือคาดการณ์การปริมาณส่งออกในอนาคตนั้นถือเป็นจุดเริ่มต้นของการวางแผน การตัดสินใจ และการบริหารความเสี่ยง ช่วยในการประเมินการคาดการณ์ล่วงหน้า ทำให้ทราบถึงปริมาณการผลิต การเพิ่ม หรือการชะลอตัวการผลิต นอกจากการพยากรณ์ให้ผลดีกับภาครัฐและเกษตรกรแล้ว ด้านการบริหารจัดการองค์กร ผู้บริหารระดับสูงขององค์กรยังใช้การพยากรณ์ปริมาณการส่งออกเพื่อวางแผนกลยุทธ์ทั้งในระยะสั้นและระยะยาว โดยที่ผู้ลงทุนใช้ข้อมูลที่ได้จากการพยากรณ์หรือคาดการณ์ เพื่อดูแนวโน้มความเสี่ยงที่ไม่แน่นอน หน่วยงานของรัฐใช้การพยากรณ์เพื่อวางนโยบายระดับชาติ เช่น อัตราการเติบโตทางเศรษฐกิจ รายได้ประชาชาติ มูลค่าการส่งออกอีกด้วย สำหรับการพยากรณ์ด้วยวิธีการทางสถิติเป็นวิธีการที่ใช้กันมาจนถึงปัจจุบัน วิธีเหล่านี้บางวิธีก็สามารถใช้ได้กับสมการเส้นตรงเท่านั้น เช่น การวิเคราะห์สหสัมพันธ์ การวิเคราะห์การถดถอยพหุคุณ และบางวิธีก็มีหลายขั้นตอน ผู้ที่ทำการพยากรณ์ทางสถิติจะต้องมีความรู้ ความเข้าใจและความชำนาญเกี่ยวกับวิธีการที่จะเลือกใช้ อีกทั้งต้องทราบการแจกจง ตัวแปรที่สนใจ เช่น ตัวแปรตาม ตัวแปรอิสระ และหากมีค่านอกกลุ่มจะส่งผลให้การประมาณค่าพารามิเตอร์ในตัวแบบเอนเอียง

ได้ สิ่งที่สำคัญที่สุดของวิธีการเหล่านี้คือ ไม่สามารถเรียนรู้ข้อมูลที่อยู่ในรูปแบบอนุกรมเวลาได้ จึงไม่สามารถใช้งานได้อย่างมีประสิทธิภาพ ทำให้นักวิจัยจำนวนมากได้หันมาใช้เทคนิคโครงข่ายประสาทเทียม เพื่อเพิ่มประสิทธิภาพในการพยากรณ์ ในปัจจุบันได้มีการนำเทคนิคโครงข่ายประสาทเทียมมาประยุกต์ใช้ในงานต่างๆ เช่น การพยากรณ์อากาศ การพยากรณ์หุ้นในตลาดหลักทรัพย์ การพยากรณ์ราคาทอง เป็นต้น โดยที่โครงข่ายประสาทเทียมไม่สนใจว่าข้อมูลมีการแจกแจงอย่างไร ข้อมูลมีปัจจัย และสิ่งรบกวนภายนอกมากน้อยหรือไม่ วิธีการนี้มีความยืดหยุ่นสูงกว่าการใช้โครงสร้างทางสถิติและคณิตศาสตร์ อีกทั้งโครงข่ายยังสามารถปรับตัวเองให้ทำงานในสภาพที่เปลี่ยนแปลงไปได้ และที่สำคัญจะไม่มีข้อตกลงเบื้องต้นเหมือนวิธีทางสถิติทั่วไป โครงข่ายประสาทเทียม (Neural Networks) เป็นแบบจำลองทางคณิตศาสตร์ที่เลียนแบบการทำงานของโครงข่ายสมองมนุษย์ ทำให้เชื่อว่าผลที่ได้จากแบบจำลองจะคล้ายคลึงกับการคิดและตัดสินใจของมนุษย์ นอกจากนั้นด้วยความเป็นเครื่องมือทางอิเลคทรอนิคส์จึงทำให้สามารถทำงานได้เกินขอบเขตของความสามารถในการคำนวณของมนุษย์

โครงข่ายประสาทเทียมนิยมนำมาใช้ในการพยากรณ์ข้อมูลแต่การหาค่า Parameter ของข้อมูลนั้นจะทำได้ยาก หากชุดข้อมูลมีความแปรปรวนสูง และข้อมูลอยู่ในรูปแบบไม่เป็นเชิงเส้น (Non-linear) (Gomez, 2007) จากปัญหาที่กล่าวมาผู้วิจัยจึงมีแนวคิดที่จะนำโครงข่ายประสาทเทียมแบบพหุนาม (Polynomial Artificial Neural Network) ซึ่งมีข้อดีคือ ไม่มีข้อตกลงเบื้องต้น เหมือนวิธีทางสถิติทั่วไป ไม่สนใจว่าข้อมูลมีการแจกแจงแบบใด ไม่สนใจความแปรปรวนของชุดข้อมูล และสามารถปรับตัวเองให้ทำงานได้ในสภาพที่เปลี่ยนแปลงไป ถึงแม้ไม่เคยมีการเรียนรู้มาก่อน มาประยุกต์ใช้ในการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีของประเทศไทย

วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาระบบพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีของประเทศไทย โดยใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น

2. เพื่อทดสอบประสิทธิภาพการทำงานของระบบพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีของประเทศไทย โดยใช้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น

3. เพื่อเปรียบเทียบเทคนิคการพยากรณ์โดยใช้โครงข่ายประสาทเทียมและโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น

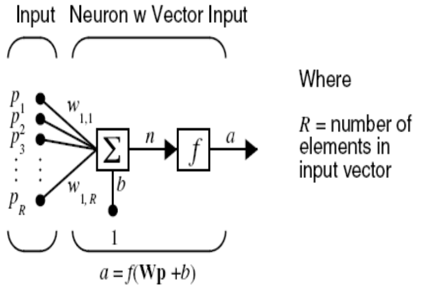
ทฤษฎีงานวิจัยที่เกี่ยวข้อง

1. โครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

โครงข่ายประสาทเทียม คือการคำนวณที่เลียนแบบการทำงานของระบบสมองมนุษย์ เพื่อใช้ประโยชน์ในการคาดคะเนเหตุการณ์จากข้อมูลที่มีอยู่

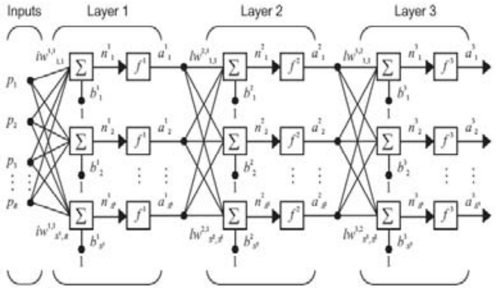
โครงข่ายประสาทเทียมจะประกอบด้วยเซลประสาท (Neural) ซึ่งจำลองมาจากการทำงานของระบบสมองมนุษย์ โดยใช้ฟังก์ชันถ่ายโอน (Transfer function, f) ค่าถ่วงน้ำหนัก (Weight, w) และค่าไบแอส (Bias, b) เป็นเครื่องมือจำลองคุณสมบัติของเซลล์ประสาท ดังรูปที่ 1

โครงสร้างการเชื่อมต่อโครงข่ายประสาทเทียมแบ่งออกเป็น โครงข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Neural Networks) (Holland, 1975) ดังรูปที่ 1 และโครงข่ายประสาทเทียมแบบหลายชั้น (Multi Layer Neural Networks) (Holland, 1975) ดังรูปที่ 2 ที่นิยมใช้กัน ซึ่งประกอบด้วยชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นแสดงผล (Output)



ภาพที่ 1 โครงข่ายประสาทเทียมแบบชั้นเดียว

ที่มา : Howard Demuth, 1996



ภาพที่ 2 โครงข่ายประสาทเทียมแบบหลายชั้น

ที่มา : Howard Demuth, 1996

เมื่อ W^1 คือ ค่าถ่วงน้ำหนัก (Weight) ที่เชื่อมระหว่างชั้นข้อมูลด้านอินพุตและชั้นซ่อนที่ 1

W^2 คือ ค่าถ่วงน้ำหนัก (Weight) ที่เชื่อมระหว่างชั้นซ่อนที่ 1 และชั้นซ่อนที่ 2

W^3 คือ ค่าถ่วงน้ำหนัก (Weight) ที่เชื่อมระหว่างชั้นซ่อนที่ 2 และชั้นเอาต์พุท (Output)

b^1 คือ ค่าไบเอส (Bias) ในชั้นซ่อน 1

b^2 คือ ค่าไบเอส (Bias) ในชั้นซ่อน 2

b^3 คือ ค่าไบเอส (Bias) ในชั้นซ่อนเอาต์พุท

การทำงานของเครือข่ายประสาทเทียม เมื่อมีข้อมูลนำเข้า (Input) เข้ามายังเครือข่ายประสาท (Network) ก็นำ Input มาคูณกับค่าน้ำหนัก (Weight) ของแต่ละขา ผลที่ได้จาก

Input ทุก ๆ ขาของ Neuron จะนำมารวมกัน แล้วก็นำมาเทียบกับระดับการกระตุ้นต่ำสุดที่ทำให้เกิดการตอบสนอง (Threshold) ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า Threshold แล้ว Neuron จะส่งผลลัพธ์ (Output) ออกไป และ Output ก็จะถูกส่งไปยัง Input ของ Neuron อื่น ๆ ที่เชื่อมกันใน Network จนได้เป็นผลลัพธ์ที่ต้องการ หลังจากนั้นจะต้องมีการตรวจสอบผลลัพธ์ที่ได้จากระบบว่ามีความคลาดเคลื่อนเพียงใด หากความคลาดเคลื่อนจากเป้าหมายมากเกินไป จะต้องมีการนำค่าความคลาดเคลื่อนนี้ไปปรับน้ำหนักการเรียนรู้ใหม่ (Weight) แบบแพร่กระจายย้อนกลับ (back-propagation) ซึ่งจะเป็นการปรับน้ำหนักความคลาดเคลื่อนจากชั้นผลลัพธ์ ไปยังชั้นก่อนหน้า และทำการปรับน้ำหนักไปเรื่อย ๆ จนกระทั่งถึงชั้นรับข้อมูล ซึ่งกระบวนการเรียนรู้แบบนี้จะต้องอาศัยการทำซ้ำหลายรอบ จนกว่าจะได้ผลลัพธ์ตามที่กำหนดหรือได้ค่าความคลาดเคลื่อนที่น้อยจนพอยอมรับได้ ซึ่งจำนวนรอบนี้ก็จะอยู่กับความยากง่ายของปัญหา ขนาดข้อมูล รวมไปถึงจำนวนชั้นของโครงสร้างของโครงข่ายที่เราสร้างไว้ด้วย ในการหาค่าความคลาดเคลื่อนจะใช้ค่าเฉลี่ยกำลังสองของความผิดพลาด (Mean Square Error: MSE) ดังสมการที่ (1)

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \quad (1)$$

เมื่อ t_i คือ ค่าจริง
a_i คือ ค่าพยากรณ์
e_i คือ ค่าความคลาดเคลื่อน
N คือ จำนวนค่าพยากรณ์

2. โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Polynomial Artificial Neural Network: PANN)

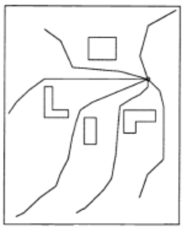
โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Multilayer Perceptron Polynomial Artificial Neural Network) ถือเป็นสถาปัตยกรรมที่ถูกพัฒนามาจากโครงข่ายประสาทเทียม (Artificial Neural Network) (Delivopoulos, 2004) ที่อยู่บนพื้นฐานของ Group Method of Data Handing (GMDH) (Byoung, 2004) และหลักการของ Polynomial Regression ถูกพัฒนาโดย Ivakhvenko ในปี 1960 และได้มีการพัฒนาประสิทธิภาพการทำงานมาอย่างต่อเนื่อง ทำให้โครงข่ายประสาทเทียมแบบพหุนามหลายชั้น มีความสามารถในการพยากรณ์ข้อมูลที่ถูกจัดเก็บในรูปแบบอนุกรมเวลาและสามารถแก้ไขปัญหาของข้อมูลที่ไม่อยู่ในรูปแบบเชิงเส้น (Non-linear) และรูปแบบเชิงเส้น (linear) และชุดข้อมูลที่มีความแปรปรวนร่วมแบบย้อนหลังไปในอดีต (Auto covariance Coefficients; γ_k) ของอนุกรมเวลาได้อย่างมีประสิทธิภาพซึ่งฟังก์ชันความแปรปรวนร่วมแบบย้อนหลังไปในอดีตของอนุกรมเวลาที่เวลาย้อนไป k จะถูกนิยามดังนี้

$$\gamma_k = \frac{1}{N}\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x}) \quad (2)$$

เมื่อ x_t คือ ค่า ณ ตำแหน่งของชุดข้อมูลนั้น ๆ

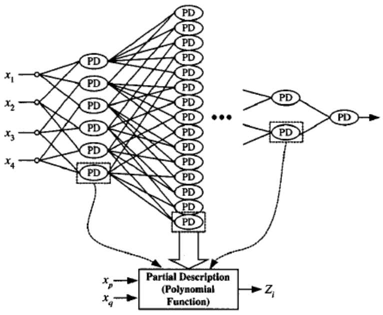
$\bar{x}$ คือ ค่าเฉลี่ยของชุดข้อมูล

โดยลักษณะของการจำแนกข้อมูลของฟังก์ชันพหุนามสามารถจำแนกลักษณะและแบ่งกลุ่มของข้อมูลได้ดีกว่าฟังก์ชันแบบเส้นตรง ดังแสดงในภาพที่ 3



ภาพที่ 3 การจำแนกลักษณะการกระจายข้อมูลโดยใช้ฟังก์ชันพหุนาม (Byoung, 2004)

โครงข่ายประสาทเทียมแบบพหุนามหลายชั้นยังสามารถลดเวลาและความผิดพลาดในกระบวนการเรียนรู้ชุดข้อมูล ซึ่งให้ผลเที่ยงตรงกว่าโครงข่ายประสาทเทียมแบบปกติ (Vasechkina, 2001) ลักษณะโครงข่ายสามารถแสดงได้ด้วยสถาปัตยกรรมโครงข่ายในชั้นของการคำนวณ ดังภาพที่ 4 ซึ่งประกอบด้วยส่วนหลักพื้นฐาน 3 ส่วนเหมือนโครงข่ายประสาทเทียมทั่วไปแต่ลักษณะการเชื่อมโยงของโหนดแต่ละโหนดแตกต่างกัน (Nikolay, 2003) กล่าวคือ โครงข่ายประสาทเทียมแบบพหุนามหลายชั้นไม่มีการเชื่อมต่อโหนดทุกโหนดเข้าด้วยกันระหว่างชั้นนำเข้าและชั้นซ่อน และโครงข่ายสามารถปรับเปลี่ยนเองได้ขึ้นอยู่กับชนิดของข้อมูล



ภาพที่ 4 สถาปัตยกรรมพื้นฐานของโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น

โดยค่า Input ของโครงข่ายประสาทเทียมจะใช้โครงสร้างในรูปแบบของ Polynomial Function ดังสมการที่ (3)

$$y_0 = w_0 + \sum_{i=1}^{m} w_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} w_{ij} x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} w_{ijk} x_i x_j x_k + \dots \quad (3)$$

เมื่อ a_0 คือ ค่าสัมประสิทธิ์ของตัวแปร

x_1, x_2 คือ ตัวแปรในการนำเข้า (input)

y_0 คือ ผลลัพธ์ (output)

โครงสร้างของโครงข่ายประสาทเทียมสามารถเขียนได้อีกรูปแบบ (Gomez, 1999) ซึ่งเป็นรูปแบบของอนุกรมเวลาย้อนหลังไปยังตำแหน่งใด ๆ สามารถอธิบายได้

$$\hat{y} = [\,\phi(x_{1,k}, x_{2,k}, \dots x_{n_i k}, x_{1,k-1}, x_{2,k-1}, \dots, x_{n,k-n}, \dots y_{k-1}, y_{k-2}, \dots y_{k-n_2})]_{\varphi\max}^{\varphi\min} \quad (4)$$

เมื่อ $\hat{y}$ คือ Estimated function

$\phi(x,y)$ คือ ฟังก์ชันสมการไม่เป็นเชิงเส้น (Non-linear Function)

x_i คือ ตัวแปรที่ใช้ในการนำเข้า

$i=1,...,n$ คือ จำนวนค่าข้อมูลนำเข้า (Number of Input)

y_{k-j} คือ คำตอบที่ได้จากประสาทเทียมในชั้นก่อนหน้า

$j=1,...,n$ คือ ค่าของข้อมูลก่อนหน้า (Number of delay)

x,y คือ ตัวแปรนำเข้าของข้อมูล

เพื่อให้ลดความซับซ่อน สามารถเขียนสมการอย่างง่ายของสมการพหุนามได้ดังนี้

$$z = \{x_{1,k}, x_{2,k}, x_{n1,k},..., y_{k-2},..., y_{k-n2}\} \quad (5)$$
$$= \{z_1, z_2, z_3,..., z_{n_v}\}$$

เมื่อ n_v คือ ผลรวมของในแต่ละสมาชิกของสมการ z สามารถอธิบายได้ด้วยสมการที่ (6)

$$n_v = n_i + n_1 n_i + n_2 \quad (6)$$

สามารถแจกแจงสมการ Polynomial ได้ดังต่อไปนี้

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i}^{n} x_i y_i & & \end{bmatrix} \begin{Bmatrix} a_0 \\ a_1 \end{Bmatrix} = \begin{Bmatrix} \sum_{i}^{n} y_i \end{Bmatrix}$$

$$\sum_{i=1}^{n} x_i^2 \quad \sum_{i=1}^{n} x_i^3 \quad \sum_{i=1}^{n} x_i^4 \qquad a_4$$

$$\sum_{i}^{n} x_i^2 y_i$$

การนิยามกระบวนเรียนรู้ของโครงข่ายประสาทเทียมแบบพหุนามหลายชั้นอธิบายแนวคิดบางอย่างที่มีการใช้กันอยู่อย่างแพร่หลาย

การประมาณค่าความผิดพลาด Error ของโครงข่ายจะสามารถอธิบายได้จากสมการ (7)

$$err_n(y^n, \varphi(z)) := \frac{1}{n}\sum (Y_k - \varphi(z_k))^2 \quad (7)$$

$$y^n = (y_1, y_2, ..., y_n)$$

เมื่อ $err_n(y^n, \varphi(z))$ คือ ผลลัพธ์ของโครงข่าย $\varphi(z_k)\varphi_p$ และ n คือ ค่าผิดพลาดที่ได้จากการพยากรณ์ข้อมูล

โดยค่าความผิดพลาดที่น้อยที่สุด (Optimal error) จะสามารถนิยามได้ดังสมการ (8)

$$opt\ err_n(y^n, \varphi\varphi(z)) = \min_{\varphi \in \varphi_p} err_n(y^n, \varphi\varphi(z)) = err_n(y^n, \phi^*(z)) \quad (8)$$

เมื่อ $\varphi^*(z) \in \varphi_p$ คือ ค่าความผิดพลาดน้อยที่สุดของโครงข่าย

การเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลปริมาณการส่งออกข้าวเจ้านาปี รวบรวมข้อมูลจากสำนักงานสถิติแห่งชาติ มีลักษณะเป็นอนุกรมเวลา มีจำนวน 3,352 ข้อมูล

ข้อมูลที่นำมาศึกษาจะอยู่ในช่วงเดือน มกราคม พ.ศ 2544 ถึงเดือน ธันวาคม พ.ศ 2550 เป็นข้อมูลแบบอนุกรมเวลาที่มีลักษณะเป็นเชิงเส้น แบ่งข้อมูลที่ใช้สำหรับการเรียนรู้ 2,346 ระเบียน คิดเป็น 70 เปอร์เซนต์ของข้อมูล และชุดข้อมูลในการทดสอบ 1,006 ระเบียน คิดเป็น 30 เปอร์เซนต์ของข้อมูล รายละเอียดดังแสดงในตารางที่ 1

ตารางที่ 1 แสดงจำนวนข้อมูลนำเข้าในการพยากรณ์

รายละเอียดข้อมูล	
1	ปริมาณการส่งออก
2	ราคาขายส่งข้าวเจ้านาปี
3	ราคาข้าว 5 เปอร์เซนต์
4	ปริมาณน้ำฝน
5	อัตราการว่างงาน
6	อัตราเงินเฟ้อ
7	อัตราแลกเปลี่ยน
8	ดัชนีราคาผู้บริโภค
9	อุปสงค์ภายในประเทศ
10	ราคาข้าวที่เกษตรขายได้

จากตารางที่ 1 เป็นการแสดงรายละเอียดของข้อมูลนำเข้า (Input layer) ที่ใช้สำหรับการเรียนรู้ (Training) ของโครงข่ายประสาทเทียม ซึ่งประกอบด้วยปัจจัยต่าง ๆ ใช้ข้อมูลทั้งหมด 10 ข้อมูล

ตารางที่ 2 แสดงจำนวนข้อมูลที่ใช้ในการทดสอบ

รายละเอียดข้อมูล	
1	ปริมาณการส่งออกข้าวเจ้านาปี

จากตารางที่ 2 เป็นการแสดงรายละเอียดของข้อมูลที่ใช้ในการทดสอบ (Testing) ของโครงข่ายคือ ปริมาณการส่งออกข้าวเจ้านาปี

การออกแบบงานวิจัย

เนื่องจากกลุ่มข้อมูลมีค่าความแตกต่างกันมากไม่เหมาะที่จะนำไปใช้ในแบบจำลอง จะต้องหาค่าสูงสุดและค่าต่ำสุดของกลุ่มข้อมูล และลดทอนขนาดข้อมูลให้อยู่ในช่วงที่เหมาะสมต่อการนำไปใช้สอนในโครงข่าย ซึ่งเป็นการแปลงค่าของข้อมูลให้อยู่ในช่วง (0 ถึง 1) โดยใช้สมการ (9)

$$P_n = \frac{P - P_{\min}}{P_{\max} - P_{\min}} \quad (9)$$

เมื่อ P_n คือ ค่าที่ได้จากสมการ
P คือ ค่าข้อมูลก่อนผ่านสมการ
$P_{\max}$ คือ ค่าข้อมูลสูงสุดก่อนผ่านสมการ
$P_{\min}$ คือ ค่าข้อมูลต่ำสุดก่อนผ่านสมการ

เมื่อ P, P_{min}, P_{max}, P_n เป็นค่าข้อมูลใด ๆ ได้แก่ ค่าข้อมูลต่ำสุด ค่าข้อมูลสูงสุด และค่า

ข้อมูลที่การลดทอนขนาดแล้วตามลำดับ หลังจากที่ได้ผลลัพธ์จะต้องถูกแปลงค่ากลับเป็นค่าจริงด้วยสมการที่ (7)

$$P = [P_n(P_{\max} - P_{\min})] + P_{\min} \quad (10)$$

การวิเคราะห์ข้อมูลที่ผู้วิจัยได้ใช้สำหรับการทดลองนั้นได้เตรียมข้อมูลสำหรับการเรียนรู้และทดสอบ โดยสามารถระบุค่าต่างๆ ที่เกี่ยวข้องเช่น ค่าการเรียนรู้ (Epoch) จำนวนโหนดในชั้นซ่อน (Hidden Node) ค่าคุณสมบัติต่าง ๆ หลังจากนั้นโปรแกรมจะทำการเรียนรู้โดยใช้ Back – Propagation training Algorithm โดยการปรับค่า Weight และ Bias ในทุก ๆ 1 รอบและจะปรับจนกระทั่ง Training Output และ Target Output มีค่าใกล้เคียงกัน ผู้วิจัยได้เตรียมข้อมูลไว้จำนวน 10 โหนด การกำหนดค่าพารามิเตอร์ต่าง ๆ ที่ใช้สำหรับการเรียนรู้คือ โหนดในชั้นซ่อน (Hidden Node) จำนวนโหนดในชั้นซ่อนมีความสำคัญอย่างหนึ่งของโครงข่ายประสาทเทียม ดังนั้นจึงได้พิจารณาเลือกใช้โหนดในการทดลองที่แตกต่างกัน โดยกำหนดค่าดังนี้ 10 20 30 40 50 70 100 120 และ 140 และทดลองใช้รอบในการเรียนรู้ 1,000 รอบ

การจัดรูปแบบข้อมูล (Pattern Data) ข้อมูลจะนำเข้ามาใช้คือ x_i และผลลัพธ์เป้าหมาย y_i เป็นข้อมูลอนุกรมเวลา (Time Series Data) จากตัวแปร 10 ตัว โดยข้อมูลนำเข้า x_i และผลลัพธ์เป้าหมาย y_i เป็นข้อมูลอนุกรมเดือนก่อน โครงข่ายประสาทเทียมฝึกสอนด้วย Levenberg-Maquardt algorithm (trainlm) และฟังก์ชันการปรับการเรียนรู้ (Learning Function) แบบ LearnGD (Grad Descent) โดยทดลองใช้ทรานเฟอร์ฟังก์ชัน (linear transfer function, tan-sigmoid transfer function และ log-sigmoid transfer function สลับกันใน f^1, f^2 และ f^3 เพื่อหาค่าเฉลี่ยความผิดพลาดน้อยสุด (Mean Square Error) หาได้จากสมการที่ (11) – (14)

linear transfer function (purelin)

$$a(n) = n \quad (11)$$

tan-sigmoid transfer function (tansig)

$$a(n) = \frac{1 - e^{-n}}{1 + e^{-n}} \quad (12)$$

log-sigmoid transfer function (logsig)

$$a(n) = \frac{1}{1 + e^{-n}} \quad (13)$$

Mean Square Error (MSE)

$$Mse = \frac{1}{n}\sum_{i=1}^{N}(t_i - a_i)^2 \quad (14)$$

เมื่อ t_i คือ ค่าจริง

a_i คือ ค่าพยากรณ์

e_i คือ ค่าความคลาดเคลื่อน

N คือ จำนวนค่าพยากรณ์ทั้งหมด

ผลการวิจัย

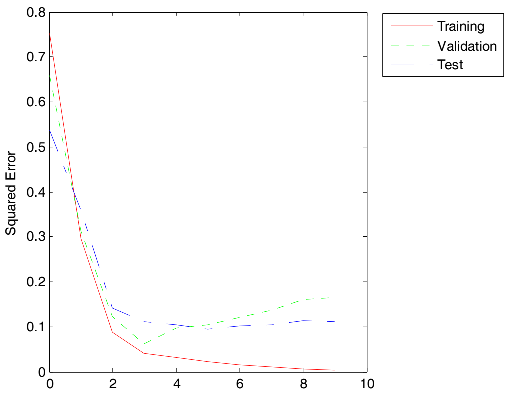
ผลการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีโดยใช้โครงข่ายประสาทเทียมและโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น โดยใช้ค่าเฉลี่ยกำลังสองสัมบูรณ์ (Mean Square Error: MSE) ในการวัดค่าความผิดพลาด

ซึ่งผู้วิจัยได้ทดลองโดยการใช้โครงข่ายประสาทเทียมเพื่อให้ได้โครงสร้างที่เหมาะสม หลังจากที่ได้ทำการทดลองหลายครั้งโดยการปรับเปลี่ยนจำนวนนิวตรอนในแต่ละชั้นซ่อน จากการลองผิดลองถูก (trial and error) ผลการทดลองแสดงดังตารางที่ 3

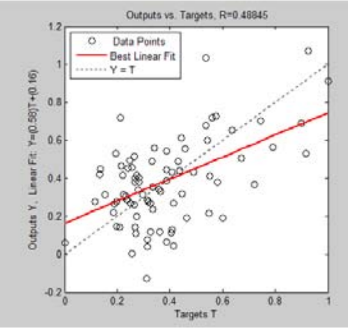
ตารางที่ 3 ผลการทดลองของโครงข่ายประสาทเทียม (Epoch = 1,000, TrainLM, goals = 0.0001)

จำนวนนิวรอน	MSE of Training	MSE of Testing
10-10-1	**0.0033**	**0.0297**
10-20-1	0.0050	0.0384
10-30-1	0.0202	0.0406
10-40-1	0.0089	0.4335
10-50-1	0.0133	0.0462
10-70-1	0.0108	0.0754
10-100-1	0.0081	0.0659
10-120-1	0.0122	0.0748
10-140-1	0.0333	0.0864

จากตารางที่ 3 เป็นการทดลองโครงข่ายประสาทเทียมหลาย ๆ ครั้ง เพื่อให้ได้โครงสร้างที่ดีที่สุด กล่าวคือ ความผิดพลาดจากการเรียนรู้ชุดข้อมูลจริงและชุดข้อมูลที่ใช้ในการพยากรณ์ จะได้โครงสร้างของโครงข่ายประสาทเทียมเพื่อการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีที่ดีที่สุด คือ 10-10-1 รูปแบบฟังก์ชันการกระตุ้นที่เกิดจากการทดลองหลาย ๆ ครั้ง (Activate Function) คือ logsig, tansig, purelin ผลการวิจัยโดยโครงข่ายประสาทเทียมให้ความผิดพลาดในการเรียนรู้ (Training) มีค่าเท่ากับ 0.0033 และค่าความผิดพลาดในการทดสอบ (Testing) มีค่าเท่ากับ 0.0297

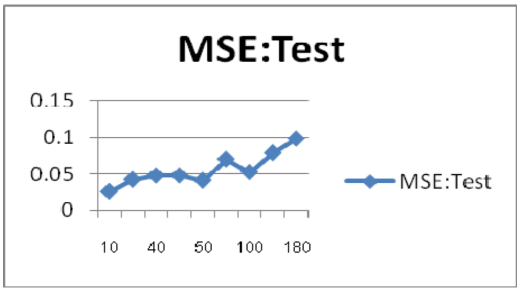


ภาพที่ 5 แสดงค่าความผิดพลาดในการเรียนรู้และทดสอบกับชุดข้อมูลในโครงข่ายประสาทเทียม



ภาพที่ 6 แสดงการกระจายตัวและแบ่งแยกกลุ่ม Target ของโครงข่ายประสาทเทียม

จากภาพที่ 7 แสดงให้เห็นถึงการกระจายตัวของกลุ่มข้อมูลใน Output ซึ่งเส้นสมการถดถอยที่ b = 1 (ความชัน 45 องศา) ลักษณะกลุ่มข้อมูลของผลลัพธ์ที่ได้จากการพยากรณ์กระจายตัวตามระนาบของเส้นสมการถดถอย (Best Linear) แสดงให้เห็นว่าชุดข้อมูลที่ได้จากการพยากรณ์ใน Output มีความคลาดเคลื่อนที่สูงเมื่อเทียบกับชุดของข้อมูลจริง



ภาพที่ 7 แสดงค่า MSE ที่ได้จากการทดสอบโครงข่ายประสาทเทียมแบบพหุนาม

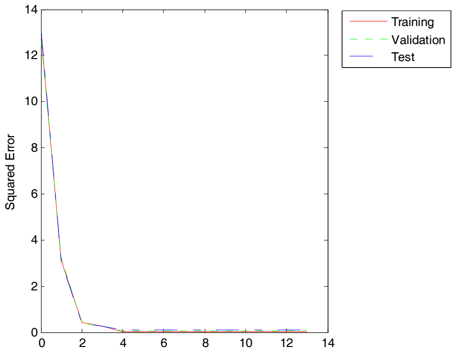
จากภาพที่ 7 พบว่าค่า MSE มีค่าต่ำสุดในจำนวนชั้นซ่อนที่ 10 คือ 0.0297 ซึ่งแสดงให้เห็นว่าจำนวนชั้นซ้อนที่เพิ่มมากขึ้น ไม่สามารถทำให้อัตราค่าเฉลี่ยกำลังสองสัมบูรณ์ลดลงตามไปด้วย

จากการผลทดลองของโครงข่ายประสาทเทียมเพื่อให้ได้ผลการทดลองที่ดีที่สุด และได้นำโครงสร้างที่ได้มาทำการทดลองกับโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น ดังแสดงในตารางที่ 4

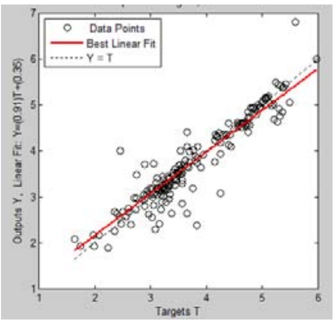
ตารางที่ 4 ผลการทดลองของโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น (Epoch = 1000, TrainLM, goals = 0.0001)

จำนวนนิวรอน	MSE of Training	MSE of Testing
10-10-1	0.0010	0.0259

จากการทดลองของโครงข่ายประสาทเทียม ผลการทดลองทำให้ได้โครงสร้างที่ดีที่สุด คือ จำนวนโหนดในชั้นนำเข้าเท่ากับ 10 จำนวนโหนดในชั้นซ่อนเท่ากับ 10 และจำนวนโหนดในชั้นแสดงผลเท่ากับ 1 รูปแบบฟังก์ชันการกระตุ้น (Activate Function) คือ logsig, tansig, purelin และนำโครงสร้างของโครงข่ายประสาทเทียมดังกล่าวมาใช้กับโครงข่ายประสาทเทียมแบบพหุนามเพื่อการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีที่ดีที่สุด ผลการวิจัยโดยโครงข่ายประสาทเทียมแบบพหุนามให้ความผิดพลาดในการเรียนรู้ (Training) เท่ากับ 0.0010 และค่าความผิดพลาดในการทดสอบ (Testing) มีค่าเท่ากับ 0.0259



ภาพที่ 8 แสดงค่าความผิดพลาดในการเรียนรู้และทดสอบกับชุดข้อมูลในโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น



ภาพที่ 9 แสดงการกระจายตัวและแบ่งแยกกลุ่ม Target ของโครงข่ายประสาทเทียมแบบพหุนามหลายชั้น

จากภาพที่ 9 แสดงให้เห็นถึงการกระจายตัวของกลุ่มข้อมูลใน Output ซึ่งเส้นสมการถดถอยที่ b = 1 (ความชัน 45 องศา) ลักษณะกลุ่มข้อมูลของผลลัพธ์ที่ได้จากการพยากรณ์กระจายตัวตามระนาบของเส้นสมการถดถอย (Best Linear) แสดงให้เห็นว่าชุดข้อมูลที่ได้จากการพยากรณ์ใน Output มีความคลาดเคลื่อนต่ำกว่าโครงข่ายประสาทเทียมเมื่อเทียบกับชุดของข้อมูลจริง

สรุปผลการวิจัย

จากผลการทดสอบโครงข่ายประสาทเทียมฟังก์ชันพหุนาม และโครงข่ายประสาทเทียม สามารถเปรียบเทียบการทดลองได้ดังตารางที่ 5

ตารางที่ 5 เปรียบเทียบค่าเฉลี่ยของโครงข่ายประสาทเทียมและโครงข่ายประสาทเทียมฟังก์ชันพหุนาม

Technical	MSE of Train	MSE of Test
ANN	0.0033	0.0297
PANN	0.0010	0.0259

การวิจัยครั้งนี้พบว่า การนำโครงข่ายประสาทเทียมแบบพหุนาม ซึ่งเป็นวิวัฒนาการมาจากโครงข่ายประสาทเทียมเป็นโครงสร้างที่ไม่สนใจลักษณะข้อมูล และเหมาะกับชุดข้อมูลที่มีความแปรปรวนสูงสามารถนำมาใช้พยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีได้มีประสิทธิภาพ โดยได้โครงสร้างที่เหมาะสมดังนี้ โครงสร้างโหนดของโครงข่ายที่ได้ คือ ชั้นรับข้อมูลเท่ากับ 10 โหนด ชั้นซ่อนเท่ากับ 10 โหนด ชั้นผลลัพธ์เท่ากับ 1 โหนด กำหนดค่าพารามิเตอร์คือ ฝึกสอนด้วย Levenberg-Maquardt algorithm (trainlm) และฟังก์ชันการปรับการเรียนรู้ (Learning Function) แบบ LearnGD (Grad Descent) Epoch = 1000, goals = 0.0001 ฟังก์ชันการเรียนรู้ (Activated Function) เท่ากับ logsig, tansig, purelin ให้ความผิดพลาดในการเรียนรู้ (Training) มีค่า

เท่ากับ 0.0010 และค่าความผิดพลาดในการทดสอบ (Testing) มีค่าเท่ากับ 0.0259 โดยที่โครงข่ายประสาทเทียมให้ความผิดในการเรียนรู้เท่ากับ 0.0033 และค่าความผิดพลาดในการทดสอบ เท่ากับ 0.0297 แสดงให้เห็นว่าโครงข่ายประสาทเทียมแบบพหุนามสามารถนำมาใช้ในการพยากรณ์ปริมาณการส่งออกข้าวเจ้านาปีได้อย่างมีประสิทธิภาพกว่าโครงข่ายประสาทเทียมแบบธรรมดา

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

แบบจำลองโครงข่ายประสาทเทียมทั้งสองชนิด ที่ทำการศึกษาไม่ใช่แบบจำลองที่ถาวรตายตัว สามารถทำการแก้ไข ปรับเปลี่ยนได้ เนื่องจากในแต่ละปีที่ผ่านไปจะมีการเปลี่ยนแปลงเหตุการณ์ต่าง ๆ มากมาย รวมถึงความสามารถหาแหล่งข้อมูลใหม่ที่สะดวกและมีความสมบูรณ์กว่าเดิมเพิ่มเข้ามา ดังนั้นตัวแบบจำลองก็ควรจะมีการปรับปรุงอย่างต่อเนื่องให้สะท้อนถึงเหตุการณ์ และตัวแปรใหม่ ๆ เหล่านั้น หรืออาจมีการนำขั้นตอนวิธีเชิงพันธุกรรมเข้ามาผสมผสานการทำงาน เพื่อเพิ่มความสามารถการพยากรณ์ได้สูงขึ้นต่อไป

เอกสารอ้างอิง

Byoung Jun Park, Witold Pedrycz & Sung Kwun Oh.(2004). *"Fuzzy Polynomial Neural Network: Hybrid Architectures of Fuzzy Modeling".*

Delivopoulos, E. & Theocharis,J.B. (2004). *"A modified PNN algorithm with optimal PD modeling using the orthogonal least squares method"*, Journal IEEE, 133-170.

Gomez, E. Najim K. & Ikonen, E. (2007). *"Forecasting time series with a new architecture for polynomial artificial neural network",* Journal IEEE, 1209-1216.

Holland. (1975). *"Adaptation in Natural and Artificial System",*The University of Michigan Press, ANN Arbor, Michigan.

Nikolay Y. Nikolaev & Hitoshi Iba, (2003). *"Learning Polynomial Feedforward Neural Network by Genetic Programming and* Back propagation*"*, Journal IEEE, 337-350.

Philip Chen, C.L. & Farid Ahmed. *(1993). "Polynomial Neural Networks Based Mobile Robot Path Planning"*, Journal IEEE, 1993, 320-32.

Vasechkina, & Yarin. (2001). *"Evolving polynomial neural network by means of genetic algorithm: some application examples",* Journal IEEE, 2001, 1-13.