# แบบจำลองการพยากรณ์ราคามันสำปะหลัง โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น

# A Cassava Price Forecasting Model Using a Multi-Layer Artificial Neural Network

ผุสดี บุญรอด* และกรวัฒน์ พลเยี่ยม

*ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ*

*มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ แขวงวงศ์สว่าง เขตบางซื่อ กรุงเทพมหานคร 10800*

Pudsadee Boonrawd* and Korawat Phonyiam

*Department of Information Technology, Faculty of Information Technology,*

*King Mongkut's University of Technology North Bangkok, Wongsawang, Bangsue, Bangkok 10800*

## บทคัดย่อ

งานวิจัยนี้นำเสนอแบบจำลองการพยากรณ์ราคามันสำปะหลัง เนื่องจากมันสำปะหลังเป็นพืชอาหารที่สำคัญของโลกประกอบกับราคามันสำปะหลังมีความผันผวนอยู่ตลอดเวลา โดยใช้ข้อมูลจากสมาคมแป้งมันสำปะหลังและสำนักงานเศรษฐกิจการเกษตร จำนวน 10 ปี (พ.ศ. 2549-2558) ซึ่งประยุกต์ใช้โครงข่ายประสาทเทียมแบบหลายชั้น (multi-layer artificial neural network) ที่ใช้การเรียนรู้แบบแพร่กระจายย้อนกลับในการหารูปแบบโครงข่ายสำหรับพยากรณ์ราคามันสำปะหลัง จากการประเมินประสิทธิภาพแบบจำลองการพยากรณ์ที่พัฒนาขึ้นโดยหาค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ (mean absolute percentage error, MAPE) โดยนำไปเปรียบเทียบกับเคเนียเรสเนเบอร์ (k-Nearest Neighbor, k-NN) และการวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) พบว่าโครงข่ายประสาทเทียมแบบหลายชั้นมีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 3.96 เคเนียเรสเนเบอร์มีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 7.11 และการวิเคราะห์การถดถอยเชิงเส้นมีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 11.10 จึงสามารถสรุปได้ว่าแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นมีประสิทธิภาพดีที่สุด ซึ่งสามารถนำไปประยุกต์ใช้และต่อยอดในการพยากรณ์ราคามันสำปะหลังได้ดีกว่าวิธีอื่น ๆ

**คำสำคัญ :** โครงข่ายประสาทเทียมแบบหลายชั้น; เคเนียเรสเนเบอร์; การวิเคราะห์การถดถอยเชิงเส้น; ราคามันสำปะหลัง; การพยากรณ์

## Abstract

This paper presents a model to forecast the price of cassava. The cassava is a world's important crops. The prices fluctuates over time. The research used the data from the Thai Tapioca

***ผู้รับผิดชอบบทความ :** pudsadee.b@it.kmutnb.ac.th

Starch Association and the Office of Agricultural Economics for 10 years from 2006 to 2015. The model is developed based on the multi-layer artificial neural network technique using back propagation learning forecast for the price of cassava. The performance of forecasting model evaluation employs the mean absolute percentage error (MAPE), compared to the k-Nearest Neighbor (k-NN) technique and the linear regression analysis technique. The result showed that the multi-layer artificial neural network technique has a MAPE of 3.96, the k-Nearest Neighbor technique has a MAPE of 7.11 and the linear regression analysis technique has a MAPE of 11.10. Therefore, it can be concluded that the predictive model using the multi-layer artificial neural network technique has the best performance. It can be applied and further researched to predict the price of cassava better than other methods.

**Keywords:** multi-layer artificial neural networks; k-Nearest Neighbor; linear regression analysis; cassava price; forecasting

## 1. บทนำ

มันสำปะหลังเป็นพืชอาหารที่สำคัญอันดับ 5 ของโลก ซึ่งเป็นแหล่งวัตถุดิบแป้งที่มีราคาถูกกว่าพืชผลิตแป้งชนิดอื่น ๆ ไม่มีสี ไม่มีกลิ่น และไม่ทำให้เกิดภูมิแพ้ จึงมีการนำมันสำปะหลังไปแปรรูปเพื่อใช้ในอุตสาหกรรมยา อาหาร เครื่องสำอางและเคมีภัณฑ์ [1] โดยปริมาณผลผลิตที่ได้ในแต่ละปีใช้เป็นอาหารของมนุษย์ร้อยละ 60 ใช้ทำอาหารสัตว์ร้อยละ 27.5 และใช้ประโยชน์ในด้านอื่น ๆ ร้อยละ 12.5 ประเทศไทยเป็นผู้ผลิตมันสำปะหลังอันดับสอง รองจากประเทศไนจีเรียที่มีการผลิตมากที่สุดในโลก ซึ่งในปี พ.ศ. 2559 จะมีการเปิดเสรีทางการค้าอาเซียน (AEC, Asean economic community) ประเทศไทยและเวียดนามถือเป็นคู่แข่งรายสำคัญ [2] ในการผลิตมันสำปะหลัง และยังเป็นพืชที่ทำรายได้ให้เกษตรกรไทยมากเป็นอันดับที่ 4 รองจากยางพารา อ้อย และข้าว ปัจจุบันราคามันสำปะหลังในประเทศไทยมีความผันผวนอยู่ตลอดเวลา เนื่องจากสภาวะของระบบเศรษฐกิจหรือเกิดการแทรกแซงทางราคา [3] ซึ่งความไม่แน่นอนของระดับราคาสินค้าและบริการนับเป็นปัญหาสำคัญของผู้ผลิต [2] หากสามารถคาดการณ์หรือพยากรณ์ราคารับซื้อมันสำปะหลัง จะส่งผลดีต่อการวางแผนเกิดการได้เปรียบทางธุรกิจ และมีการจัดการที่ดีกว่าประเทศคู่แข่งทางการค้า

จากปัญหาดังกล่าว ดังนั้นวัตถุประสงค์ของงานวิจัยนี้จึงได้นำเสนอแบบจำลองการพยากรณ์ราคามันสำปะหลัง โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น (multi-layer artificial neural network) ซึ่งแบบจำลองการพยากรณ์ที่พัฒนาขึ้นจะสามารถนำไปประยุกต์ใช้ และต่อยอดในการพัฒนาเครื่องมือสำหรับวิเคราะห์ คาดการณ์ราคามันสำปะหลัง และช่วยสนับสนุนการตัดสินใจในอนาคต โดยบทความวิจัยนี้แบ่งเนื้อหาออกเป็น 5 หัวข้อหลัก ประกอบด้วยบทนำ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง วิธีดำเนินการวิจัย ผลการวิจัย และสรุปผลการวิจัย

## 2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

### 2.1 มันสำปะหลัง (Cassava)

มันสำปะหลังเป็นพืชที่เก็บสะสมอาหารในรูปของคาร์โบไฮเดรตหรือแป้งไว้ในราก เป็นพืชที่ให้

พลังงานแก่คนและสัตว์ ทนแล้ง สามารถปลูกได้ในดินทุกชนิด แม้แต่ดินที่มีความอุดมสมบูรณ์ต่ำ ในประเทศไทยสามารถปลูกมันสำปะหลังได้ทุกภาคและตลอดทั้งปี ส่วนใหญ่จะนำหัวมันสำปะหลังไปใช้เป็นวัตถุดิบที่สำคัญในอุตสาหกรรมมันเส้น มันอัดเม็ด และอุตสาหกรรมแป้งมันสำปะหลัง [4] โดยชนิดของมันสำปะหลังสามารถแบ่งออกได้ดังนี้

2.1.1 มันสำปะหลังชนิดหวาน (sweet type) มีปริมาณกรดไฮโดรไซยานิคต่ำ ไม่มีรสขม ใช้เพื่อการบริโภคของมนุษย์ มีทั้งชนิดเนื้อร่วน นุ่ม และแน่นเหนียว ซึ่งประเทศไทยไม่มีการปลูกในพื้นที่ขนาดใหญ่ เนื่องจากมีตลาดจำกัด

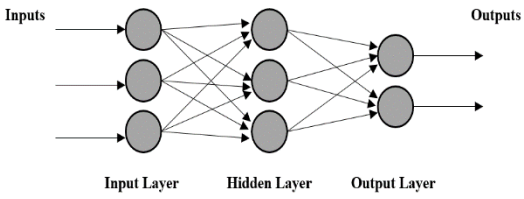
2.1.2 มันสำปะหลังชนิดขม (bitter type) มีปริมาณกรดไฮโดรไซยานิคสูง ไม่เหมาะสำหรับการบริโภคของมนุษย์ หรือใช้เลี้ยงสัตว์โดยตรง เหมาะสำหรับอุตสาหกรรมแปรรูปต่าง ๆ เช่น แป้งมัน มันอัดเม็ด และแอลกอฮอล์ เนื่องจากมีปริมาณแป้งสูง [4] ในประเทศไทยนิยมปลูกมันชนิดนี้ เนื่องจากนำไปใช้ในการแปรรูปเป็นผลิตภัณฑ์ใหม่และการส่งออก

**2.2 การพยากรณ์ข้อมูลอนุกรมเวลา (time series forecasting)**

ข้อมูลอนุกรมเวลา (time series) เป็นข้อมูลที่ถูกจัดเรียงตามเวลาที่บันทึกข้อมูล [5] ซึ่งการพยากรณ์ข้อมูลอนุกรมเวลาเป็นการศึกษาถึงการเปลี่ยนแปลงของข้อมูลตามระยะเวลาที่สนใจ เพื่อคาดการณ์หรือพยากรณ์ข้อมูลที่จะเกิดขึ้นในอนาคต โดยอาศัยวิธีการต่าง ๆ มาช่วยในการหารูปแบบข้อมูลลักษณะความสัมพันธ์ระหว่างค่าที่สังเกตที่เก็บตามลำดับเวลา และใช้รูปแบบมาพยากรณ์ค่าในอนาคต [6] ตัวอย่างวิธีการที่นำมาใช้ประยุกต์ใช้ เช่น เหมืองข้อมูล (data mining) เป็นกระบวนการสกัดหรือวิเคราะห์กลั่นกรองสารสนเทศที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ ซึ่งอาศัยหลักการเรียนรู้ของเครื่อง หรือการรู้จำรูปแบบ การเรียนรู้ที่แบ่งออกได้เป็น 2 ประเภท ดังนี้

2.2.1 การเรียนรู้แบบมีการสอน (supervised learning) เป็นการเรียนรู้ที่มีการตรวจคำตอบหรือพิจารณาผลลัพธ์ที่ได้เทียบกับข้อมูลฝึกฝน (training data) หากมีค่าไม่ตรงกับข้อมูลฝึกฝนจะปรับค่าที่มีความคลาดเคลื่อน เพื่อให้ได้ผลลัพธ์ที่ดีขึ้น [7] ซึ่งวิธีการเรียนรู้แบบมีการสอน ตัวอย่าง เช่น

(1) โครงข่ายประสาทเทียม (artificial neural network) เป็นแบบจำลองทางคณิตศาสตร์ที่ใช้หลักการคำนวณแบบคอนเนคชันนิสต์ (connectionnist) เพื่อจำลองลักษณะการทำงานให้คล้ายกับโครงข่ายประสาทของมนุษย์ ประกอบด้วย 3 ส่วนสำคัญ คือ เซลล์ประสาท (neuron หรือ soma) ใยประสาท (never fiber) และไซแนปส์ (synapse) เพื่อใช้สร้างเครื่องมือที่จะเรียนรู้หรือจดจำรูปแบบ [7] โดยหลักการทำงานแบ่งออกได้เป็น 3 ชั้น ได้แก่ ชั้นนำเข้าข้อมูล (input layer) จำนวนหนึ่งชั้น สำหรับนำเข้าข้อมูลเพื่อเรียนรู้ ชั้นแอบแฝง (hidden layer) เป็นชั้นตรงกลางโดยภายในจะประกอบไปด้วยโหนดสำหรับใช้ในการประมลผล และชั้นผลลัพธ์ (output layer) จำนวนหนึ่งชั้น ซึ่งได้จากการประมวลผลในชั้นแอบแฝง [8] ดังรูปที่ 1




**รูปที่ 1** โครงสร้างแบบจำลองโครงข่ายประสาทเทียม

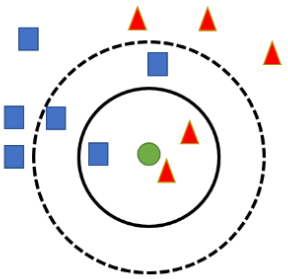
ซึ่งรูปแบบของสมการสามารถแสดงได้ดังสมการที่ (1)

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

โดย n คือ ผลรวมที่ได้จากฟังก์ชันผลรวม; $x_i$ คือ ค่า

ข้อมูลเข้าตัวที่ i; $w_i$ คือ ค่าถ่วงน้ำหนักของนิวรอนตัวที่ I; z คือ จำนวนนิวรอลชั้นข้อมูลเข้า; b คือ ค่าไบแอส; i คือ ค่าตั้งแต่ 1 ถึง z

(2) เคเนียเรสเนเบอร์ (k-Nearest Neighbor, k-NN) วิธีการหนึ่งสำหรับการทำเหมืองข้อมูลโดยใช้วิธีการจัดแบ่งกลุ่ม ซึ่งมีการแบ่งออกเป็นชุดข้อมูลฝึกฝน (training set) และชุดข้อมูลทดสอบ (testing set) โดยกระบวนการทำงานในแต่ละชุดข้อมูลจะหาระยะห่างของแต่ละแอตทริบิวต์ (attribute) ในข้อมูลที่เรียกว่าระยะทางแบบยุคลิด (euclidean distance) กับชุดข้อมูลฝึกฝน ดังนั้นจึงจัดกลุ่มข้อมูลที่อยู่ใกล้เข้าไว้ด้วยกัน [9] ดังรูปที่ 2



**รูปที่ 2** โครงสร้างเคเนียเรสเนเบอร์

(3) การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis)

การถดถอยเชิงเส้น (linear regression) เป็นวิธีการทางคณิตศาสตร์ที่อาศัยความสัมพันธ์เชิงเส้น ซึ่งเกี่ยวข้องกับการศึกษาและวิเคราะห์ตัวแปรอิสระที่มีอิทธิพลต่อการเปลี่ยนแปลงของตัวแปรตาม ซึ่งมีผลทำให้ค่าเปลี่ยนไปในรูปแบบใดรูปแบบหนึ่ง สามารถจะอธิบายลักษณะความสัมพันธ์ด้วยรูปแบบการถดถอย (regression model) [10] ดังสมการที่ (2)

$$Y = \alpha + \beta x \qquad (2)$$

โดย Y คือ ตัวแปรตาม; x คือ ตัวแปรอิสระหรือค่าคงที่ ; α คือ สัมประสิทธิ์ของความถดถอยที่เป็นจุดตัดบนแกน Y; β คือ สัมประสิทธิ์ของความถดถอยที่เป็นค่าความชันของเส้นตรง

2.2.2 การเรียนรู้แบบไม่มีการสอน (unsupervised learning) เป็นการเรียนรู้ที่ไม่ใช้ข้อมูลฝึกฝน ไม่มีการระบุผลลัพธ์ที่ต้องการไว้ก่อนล่วงหน้า เป็นการสร้างโมเดลหรือแบบจำลองที่เหมาะสมกับข้อมูล โดยพิจารณาจากข้อมูลนำเข้าที่เป็นเซตของตัวแปรสุ่มและหาความหนาแน่นร่วมของชุดข้อมูล ซึ่งอาจนำไปประยุกต์ใช้ในการอนุมาน (inference) การแบ่งกลุ่ม (clustering) หรือการจัดหมวดหมู่ข้อมูล (classification) ตัวอย่างวิธีการเรียนรู้แบบไม่มีการสอน [7,8] ได้แก่ วิธีการแบ่งกลุ่มข้อมูลแบบเคมีน (K-means clustering) ขั้นตอนวิธีอพริโอริ (Apriori algorithm) และการจัดกลุ่มแบบดีบีสแกน (density-based spatial clustering of application with noise, DBSCAN) เป็นต้น

**2.6 วรรณกรรมที่เกี่ยวข้อง**

การศึกษาเอกสารและงานวิจัยต่าง ๆ พบว่ามีหลายงานวิจัยที่เกี่ยวข้องหรือสามารถนำมาต่อยอดในการพยากรณ์ราคามันสำปะหลัง ตัวอย่าง เช่น การพยากรณ์ราคามันสำปะหลังโดยใช้แบบจำลองอารีมา (ARIMA) และการพยากรณ์โดยวิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลอย่างง่าย (simple exponential smoothing) ร่วมกับวิธีการของโฮลท์-วินเตอร์ (Holt-Winters model) ซึ่งเป็นวิธีการพยากรณ์ทางเศรษฐศาสตร์ที่นิยมนำมาประยุกต์ใช้ในการพยากรณ์ที่หลากหลาย [4,11] โดยวิธีการดังกล่าวเหมาะสำหรับการพยากรณ์ระยะสั้น [6] จึงมีการศึกษาหาแนวทางหรือวิธีการเพื่อเพิ่มประสิทธิภาพในการพยากรณ์ให้มีความแม่นยำมากยิ่งขึ้น ตัวอย่าง เช่น การพยากรณ์ราคาส่งออกยางพารา [5] การพยากรณ์ราคาหุ้น [7] และการพยากรณ์ราคาสินค้าเกษตร [6] จากผลการศึกษาวิจัยพบว่าวิธีการโครงข่ายประสาท

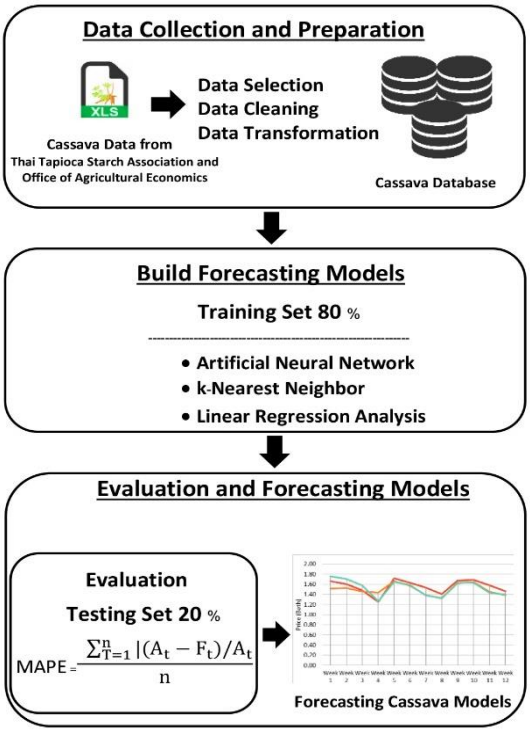
เทียมสามารถพยากรณ์ค่าข้อมูลที่เป็นค่าตัวเลขได้ดี ถูกนำมาประยุกต์ใช้ในการพยากรณ์ที่หลากหลาย และมีประสิทธิภาพในการทำนายสูง โดยสามารถพยากรณ์ในระยะยาวได้ดีกว่าวิธีการพยากรณ์ทางสถิติที่กล่าวมาข้างต้น [6,8,9] จึงเหมาะกับข้อมูลหลากหลายรูปแบบและสามารถวิเคราะห์ข้อมูลที่มีหลายตัวแปรได้ดี

## 3. วิธีดำเนินงานวิจัย

การศึกษาปัญหา ทฤษฎี และงานวิจัยที่เกี่ยวข้อง พบว่าจากปัญหาเรื่องความผันผวนของราคามันสำปะหลัง การเพิ่มขึ้นของต้นทุนการผลิตประกอบกับการลดลงของราคาผลผลิต หากสามารถคาดการณ์หรือพยากรณ์ราคามันสำปะหลังล่วงหน้า จะส่งผลดีต่อการวางแผนการผลิตมันสำปะหลัง โดยคณะผู้วิจัยจึงได้ออกแบบกรอบแนวคิดการวิจัยที่อิงตามแนวคิดการค้นหาฐานความรู้ในฐานข้อมูล (knowledge discovery in database, KDD) ที่ใช้ในการสกัดความรู้จากฐานข้อมูลขนาดใหญ่ โดยกรอบแนวคิดการวิจัยของแบบจำลองการพยากรณ์ราคามันสำปะหลังโดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นสามารถแบ่งเป็น 3 กระบวนการหลัก ดังแสดงได้ดังรูปที่ 3

### 3.1 การรวบรวมและเตรียมข้อมูล (data collection and preparation)

ข้อมูลที่ใช้ในการวิจัยมาจากสมาคมแป้งมันสำปะหลังไทยและสำนักงานเศรษฐกิจการเกษตร จำนวน 10 ปี ที่เก็บรวบรวมข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2549 ถึงเดือนธันวาคม พ.ศ. 2558 ซึ่งเป็นข้อมูลอนุกรมเวลาแบบรายสัปดาห์ที่ถูกจัดเรียงตามเวลาของการบันทึก โดยแสดงตัวอย่างข้อมูลดังตารางที่ 1




**รูปที่ 3** กรอบแนวคิดของการวิจัย

**ตารางที่ 1** ตัวอย่างรายละเอียดของข้อมูลที่นำมาใช้ในการวิจัย

| ชื่อแอตทริบิวต์ | ความหมาย |
|---|---|
| Date | วันที่บันทึกข้อมูลราคามันสำปะหลัง |
| Industrial | เขตพื้นที่ของโรงงานที่รับซื้อหัวมันสำปะหลัง |
| Avg_Percentage | ปริมาณเปอร์เซ็นต์แป้งในหัวมันสำปะหลังที่ 25 เปอร์เซ็นต์ |
| Avg_Root | ราคารับซื้อหัวมันมันสำปะหลัง (หน่วยเป็นบาทต่อกิโลกรัม) |
| Avg_Price | ราคาเฉลี่ยของหัวมันสำปะหลังแบบคละรวม (หน่วยเป็นบาทต่อกิโลกรัม) |
| Avg_Price_Percentage | ราคาเฉลี่ยที่ปริมาณเปอร์เซ็นต์แป้งเฉลี่ย (หน่วยเป็นบาทต่อกิโลกรัม) |

กระบวนการรวบรวมและเตรียมข้อมูลสามารถแบ่งออกได้เป็น 3 กระบวนการย่อย ได้แก่

(ก) การคัดเลือกข้อมูล (data selection) โดยข้อมูลที่นำมาใช้ในการวิจัย ได้แก่ วันที่บันทึกข้อมูล ราคามันสำปะหลัง เขตพื้นที่ของโรงงานที่รับซื้อมันสำปะหลัง ปริมาณเปอร์เซ็นต์แป้งในหัวมันสำปะหลังที่ 25 เปอร์เซ็นต์ ราคารับซื้อหัวมันสำปะหลัง ราคาเฉลี่ยของหัวมันสำปะหลังแบบคละรวม และราคาเฉลี่ยที่ปริมาณเปอร์เซ็นต์แป้งเฉลี่ย และเขตพื้นที่ของโรงงานที่รับซื้อหัวมันสำปะหลังซึ่งประกอบไปด้วย 9 เขตพื้นที่ ได้แก่ เขตพื้นที่โคราชตะวันออกและข้างเคียง เขตพื้นที่โคราชตะวันตกและข้างเคียง เขตพื้นที่กาฬสินธุ์และข้างเคียง เขตพื้นที่ภาคอีสานตอนบน เขตพื้นที่สระแก้ว/จันทบุรี/ฉะเชิงเทรา เขตพื้นที่ชลบุรี/ระยอง เขตพื้นที่ภาคเหนือตอนล่าง เขตพื้นที่ราชบุรี/กาญจนบุรีตอนใต้ และเขตพื้นที่อุทัยธานี

(ข) การทำความสะอาดข้อมูล (data cleaning) โดยกรองข้อมูลที่ไม่ถูกต้องหรือไม่สมบูรณ์ ซึ่งพบว่ามีข้อมูลบางเรคคอร์ดที่มีค่าว่างหรือมีค่าสูญหาย (missing value) จากการสอบถามผู้เชี่ยวชาญที่เกี่ยวข้องจึงนำค่าเฉลี่ยของราคามันสำปะหลังในแต่ละปีแทนที่ค่าข้อมูลที่มีค่าว่างหรือค่าสูญหาย

(ค) การแปลงรูปแบบข้อมูล (data transformation) โดยแปลงแฟ้มข้อมูลจากรูปแบบ spreadsheet file ให้อยู่ในรูปแบบ comma separate values (CSV) เพื่อนำไปใช้ในการพัฒนาแบบจำลองการพยากรณ์ต่อไป

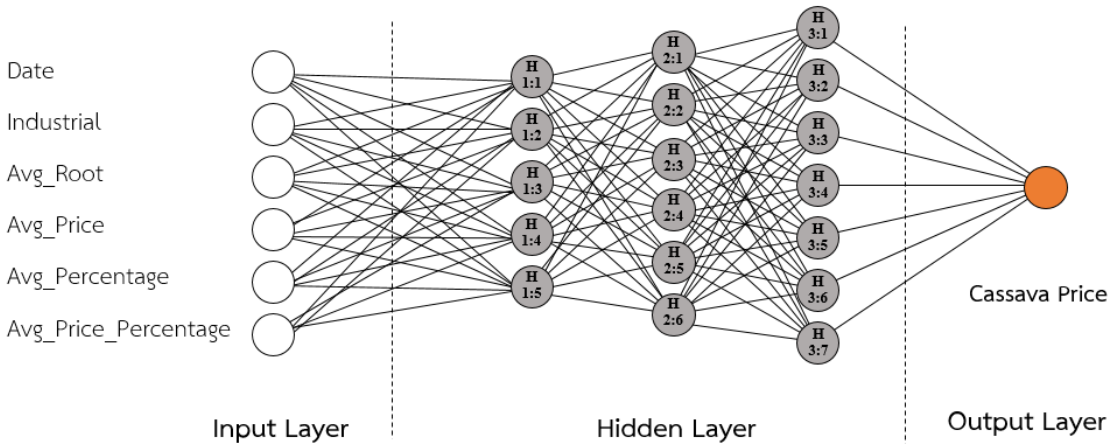
**3.2 การพัฒนาแบบจำลองการพยากรณ์ (build forecasting model)**

หลังจากผ่านกระบวนการรวบรวมและเตรียมข้อมูล คณะผู้วิจัยได้ใช้วิธี split test ของโปรแกรม RapidMiner Studio 7 ซึ่งเป็นการแบ่งข้อมูลด้วยการสุ่ม โดยแบ่งชุดข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลฝึกฝน (training set) คิดเป็นร้อยละ 80 จากข้อมูลทั้งหมด เพื่อใช้สำหรับพัฒนาแบบจำลองการพยากรณ์โดยประยุกต์ใช้ 3 วิธีการ ได้แก่ โครงข่ายประสาทเทียมแบบหลายชั้น เคเนียเรสเนเบอร์ และการวิเคราะห์การถดถอยเชิงเส้น ส่วนชุดข้อมูลทดสอบคิดเป็นร้อยละ 20 จากข้อมูลทั้งหมด เพื่อใช้ในการทดสอบประสิทธิภาพแบบจำลองการพยากรณ์ ซึ่งใช้ข้อมูลนำเข้าจำนวน 6 แอตทริบิวต์ โดยใช้โปรแกรม RapidMiner Studio 7 ที่เป็นโปรแกรมสำเร็จรูปในการทำเหมืองข้อมูลสำหรับการพัฒนาแบบจำลองการพยากรณ์

3.2.1 การพัฒนาแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น

พัฒนาแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นที่ใช้ลักษณะการเรียนรู้แบบแพร่กระจายย้อนกลับ (back propagation) โดยกำหนดค่าพารามิเตอร์ learning rate อยู่ในช่วง 0 ถึง 1 ค่า momentum อยู่ในช่วง 0 ถึง 1 และค่า training cycle ที่ 300, 500, 800, 1,000 และ 1,500 จากการพัฒนาแบบจำลองการพยากรณ์พบว่าแบบจำลองการพยากรณ์ที่พัฒนาขึ้น กำหนดค่า learning rate เท่ากับ 0.3 ค่า momentum เท่ากับ 0.2 และค่า training cycle เท่ากับ 1,000 ให้ค่าความผิดพลาดน้อยที่สุด โดยตัวอย่างแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นดังรูปที่ 4

เมื่อมีข้อมูลนำเข้าจะประมวลผล โดยคูณกับค่าถ่วงน้ำหนัก (weight, W) ซึ่งแทนความสำคัญของข้อมูลนำเข้าแต่ละตัวรวมกับค่าไบแอส (bias, b) ซึ่งผลรวมของข้อมูลที่ได้ปรับค่าถ่วงน้ำหนักและปรับค่าไบแอสแล้วจะถูกส่งต่อไปยังฟังก์ชันถ่ายโอน (transfer function) ซึ่งมีสมการสำหรับการคำนวณค่าผลลัพธ์ [12,13] ดังสมการที่ (3)




**รูปที่ 4** ตัวอย่างแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้นที่ได้พัฒนาขึ้น

$$a^3 = f^3(LW^{3,2} f^2(LW^{2,1} f^1(IW^{1,1} p + b^1) + b^2) + b^3) \quad (3)$$

โดย $W^{1,1}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นข้อมูลนำเข้าและชั้นซ่อนที่ 1; $W^{2,1}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นซ่อนที่และชั้นซ่อนที่ 2; $W^{3,2}$ คือ ค่าถ่วงน้ำหนักที่เชื่อมระหว่างชั้นชั้นซ่อนที่ 2 และชั้นข้อมูลผลลัพธ์; $b^1$ คือ ค่าไบแอสในชั้นซ่อนที่ 1; $b^2$ คือ ค่าไบแอสในชั้นซ่อนที่ 2; $b^3$ คือ ค่าไบแอสในชั้นข้อมูลผลลัพธ์

โดยค่าถ่วงน้ำหนักและค่าไบแอสสามารถปรับค่าได้ตามการเรียนรู้ ซึ่งผลลัพธ์ที่ได้จะขึ้นอยู่กับฟังก์ชันการถ่ายโอน [12] ดังสมการที่ (4)

$$n = W^{11}X^1 + W^{12}X^2 + \ldots W^{1R}X^R + b \quad (4)$$

เขียนให้อยู่ในรูปเมตริกได้ดังสมการที่ (5)

$$n = W_x + b \quad (5)$$

3.2.2 การพัฒนาแบบจำลองการพยากรณ์โดยใช้วิธีการเคเนียเรสเนเบอร์

พัฒนาแบบจำลองการพยากรณ์โดยใช้เคเนียเรสเนเบอร์ ซึ่งกำหนดค่าพารามิเตอร์ K ตั้งแต่ 1 ถึง 10 และค่า measure type แบบ mixed measure, nominal measure, numerical measure และ bregman divergence โดยตัวอย่างผลการพัฒนาแบบจำลองการพยากรณ์โดยใช้เคเนียเรสเนเบอร์ที่กำหนดค่า measure type แบบ mixed measure แสดงได้ดังตารางที่ 2

**ตารางที่ 2** ตัวอย่างผลการพัฒนาแบบจำลองการพยากรณ์โดยใช้เคเนียเรสเนเบอร์

| ค่าพารามิเตอร์ K | Measure type | Root mean square error |
|---|---|---|
| 1 | Mixed Measure | 1.319 |
| 2 | Mixed Measure | 0.446 |
| 3 | Mixed Measure | 0.413 |
| 4 | Mixed Measure | 0.440 |
| 5 | Mixed Measure | 0.395 |
| 6 | Mixed Measure | 0.389 |
| 7 | Mixed Measure | 0.435 |
| 8 | Mixed Measure | 0.467 |
| 9 | Mixed Measure | 0.501 |
| 10 | Mixed Measure | 0.536 |

3.2.3 การพัฒนาแบบจำลองการพยากรณ์โดยใช้การวิเคราะห์การถดถอยเชิงเส้น

พัฒนาแบบจำลองการพยากรณ์โดยใช้การวิเคราะห์ถดถอยเชิงเส้น ซึ่งมีตัวแปรนำเข้าทั้งหมด 6 ตัวแปร ซึ่งกำหนดค่า min tolerance อยู่ในช่วง 0.01 ถึง 1 ซึ่งเป็นการกำหนดค่าความคลาดเคลื่อนที่ยอมรับได้ โดยสามารถแสดงตัวอย่างผลการพัฒนาแบบจำลองการพยากรณ์โดยใช้การวิเคราะห์การถดถอยเชิงเส้นได้ดังตารางที่ 3

**ตารางที่ 3** ตัวอย่างผลการพัฒนาแบบจำลองการพยากรณ์โดยใช้การวิเคราะห์การถดถอยเชิงเส้น

| Min tolerant | Root mean square error |
|---|---|
| 0.01 | 0.040 |
| 0.02 | 0.040 |
| 0.03 | 0.040 |
| 0.04 | 0.040 |
| 0.05 | 0.040 |

**3.3 การประเมินประสิทธิภาพและแบบจำลองการพยากรณ์ (evaluation and forecast-ting model)**

ประเมินประสิทธิภาพแบบจำลองการพยากรณ์โดยหาค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ (mean absolute percentage error, MAPE) ซึ่งใช้ชุดข้อมูลทดสอบคิดเป็นร้อยละ 20 จากข้อมูลทั้งหมด โดยค่า MAPE ที่คำนวณได้หากมีค่าน้อยจะแสดงถึงความถูกต้องหรือมีความแม่นยำในการพยากรณ์มาก [14,15] ซึ่งแสดงดังสมการที่ (6)

$$MAPE = \frac{\sum_{t=1}^{n} |(A_t - F_t)/A_t}{n} \qquad (6)$$

โดย $A_t$ คือ ค่าจริงที่ได้ ณ เวลา t; $F_t$ คือ ค่าพยากรณ์ ณ เวลา t; n คือ จำนวนกลุ่มตัวอย่างทั้งหมด

เกณฑ์ที่ใช้ในการแปลผลการประเมินประสิทธิภาพแบบจำลองการพยากรณ์โดยใช้ค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ [15] ดังตารางที่ 4

## 4. ผลการวิจัย

**4.1 ผลการประเมินประสิทธิภาพแบบจำลองการพยากรณ์**

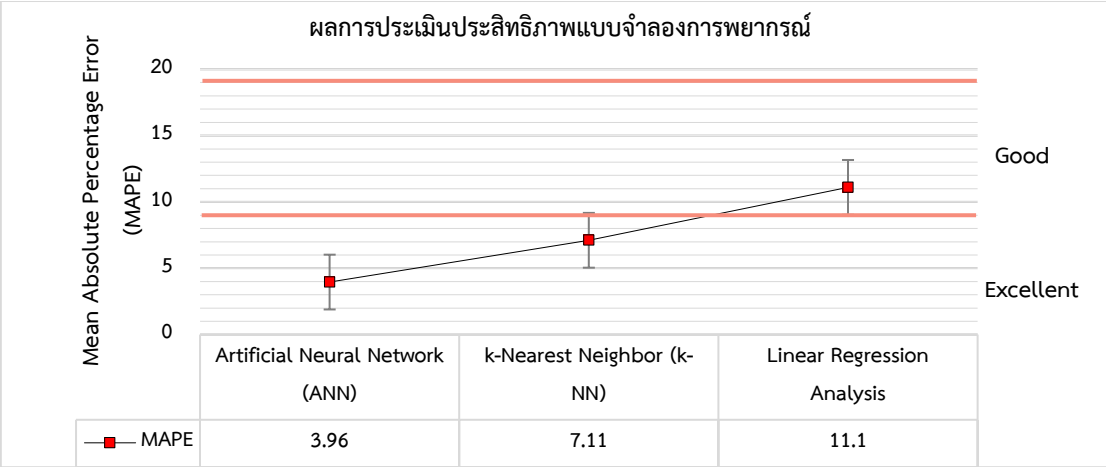
ผลการประเมินประสิทธิภาพแบบจำลองการพยากรณ์ที่พัฒนาขึ้นจากทั้ง 3 วิธีการ ได้แก่ วิธีการโครงข่ายประสาทเทียมแบบหลายชั้น เคเนียเรส-เนเบอร์ และการวิเคราะห์การถดถอยเชิงเส้น โดยคำนวณหาค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์หรือค่า MAPE ซึ่งพบว่าแบบจำลองการพยากรณ์ โดยใช้วิธีการโครงข่ายประสาทเทียมแบบหลายชั้น มีค่า MAPE น้อยที่สุดเท่ากับ 3.96 ซึ่งแบบจำลองการพยากรณ์ที่มีค่าความผิดพลาดน้อยที่สุดมีการกำหนดค่า learning rate เท่ากับ 0.3 ค่า momentum เท่ากับ 0.2 และค่า training cycle เท่ากับ 1,000 ส่วนวิธีการเคเนียเรสเนเบอร์มีค่า MAPE รองลงมาเท่ากับ 7.11 ซึ่งแบบจำลองการพยากรณ์ที่มีค่าความผิดพลาดน้อยที่สุดมีการกำหนดค่า k เท่ากับ 6 และค่า measure type แบบ mixed measure และการวิเคราะห์การถดถอยเชิงเส้นมีค่า MAPE มากที่สุดเท่ากับ 11.10 ซึ่งแบบจำลองการพยากรณ์ที่มีค่าความผิดพลาดน้อยที่สุดมีการกำหนด ค่า min tolerance อยู่ในช่วง 0.01 ถึง 0.05 สามารถแสดงดังรูปที่ 5

**4.2 ผลการพยากรณ์ราคามันสำปะหลัง**

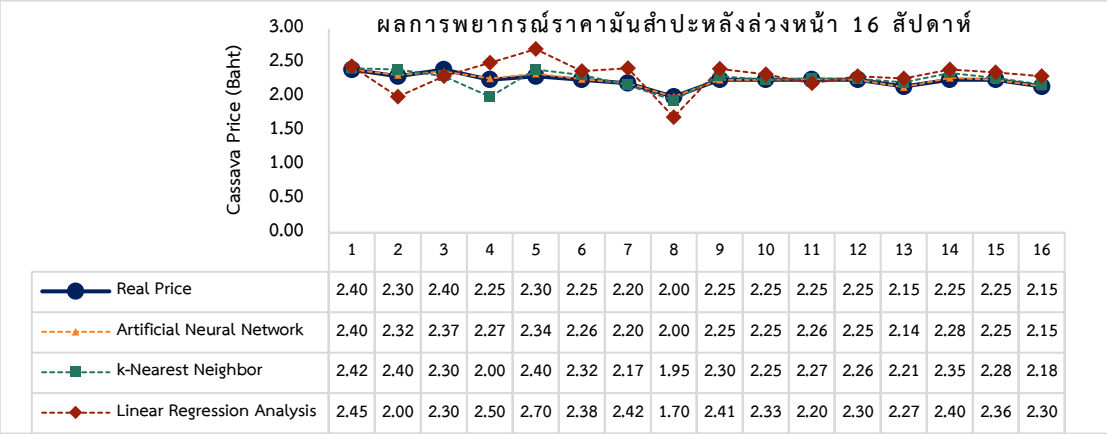
ผลการพยากรณ์ราคามันสำปะหลังล่วงหน้า 16 สัปดาห์ โดยเปรียบเทียบกับราคาจริง (real price) พบว่าโครงข่ายประสาทเทียมแบบหลายชั้นพยากรณ์ได้ใกล้เคียงกับราคาจริงมากที่สุดตามด้วยเคเนียเรสเนเบอร์และการวิเคราะห์การถดถอยเชิงเส้น ซึ่งแสดงดังรูปที่ 6

**ตารางที่ 4** เกณฑ์การแปลผลการประเมินประสิทธิภาพแบบจำลองการพยากรณ์โดยใช้ค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์

| ค่าร้อยละเฉลี่ย | การแปลผล |
|---|---|
| 0-10 % | แบบจำลองการพยากรณ์มีประสิทธิภาพอยู่ในเกณฑ์ดีเยี่ยม |
| 11-20 % | แบบจำลองการพยากรณ์มีประสิทธิภาพอยู่ในเกณฑ์ดี |
| 21-50 % | แบบจำลองการพยากรณ์มีประสิทธิภาพอยู่ในเกณฑ์ปานกลาง |
| 51-100 % | แบบจำลองการพยากรณ์มีความผิดพลาดมาก |




**รูปที่ 5** ผลการประสิทธิภาพแบบจำลองการพยากรณ์




**รูปที่ 6** ผลการพยากรณ์ราคามันสำปะหลังล่วงหน้า 16 สัปดาห์

## 5. สรุป

การพัฒนาแบบจำลองการพยากรณ์ราคามันสำปะหลังที่ปริมาณเปอร์เซ็นต์แป้งที่ 25 เปอร์เซ็นต์ ประยุกต์ใช้ 3 วิธีการ ได้แก่ โครงข่ายประสาทเทียมแบบหลายชั้น เคเนียเรสเนเบอร์ และการวิเคราะห์การถดถอยเชิงเส้น โดยแบ่งข้อมูลออกเป็น 2 ส่วน ได้แก่

ชุดข้อมูลฝึกฝนร้อยละ 80 สำหรับใช้ในการพัฒนาแบบจำลองการพยากรณ์ และชุดข้อมูลทดสอบร้อยละ 20 เพื่อใช้สำหรับการประเมินแบบจำลองการพยากรณ์ จากผลการวิจัยพบว่าแบบจำลองการพยากรณ์โดยใช้โครงข่ายประสาทเทียมแบบหลายชั้น สามารถพยากรณ์ราคามันสำปะหลังได้ใกล้เคียงกับราคาจริงที่สุดโดยมี ค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดเท่ากับ 3.96 ซึ่งแสดงให้เห็นว่าแบบจำลองการพยากรณ์ มีประสิทธิภาพอยู่ในเกณฑ์ที่ดีเยี่ยม และเหมาะสมการพยากรณ์ราคามันสำปะหลังในระยะยาวได้ดีกว่าเคเนียเรสเนเบอร์ที่มีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดเท่ากับ 7.11 และการวิเคราะห์การถดถอยเชิงเส้นที่มีค่าร้อยละเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์น้อยที่สุดเท่ากับ 11.10 ตามลำดับ โดยในการพัฒนาแบบจำลองคณะผู้วิจัยได้ใช้วิธี split test ซึ่งเป็นการแบ่งชุดข้อมูลโดยใช้การสุ่มเพื่อลดโอกาสของการเกิด overfitting ซึ่งแบบ จำลองการพยากรณ์ที่พัฒนาขึ้นนี้ สามารถนำไปประยุกต์ใช้ในการพยากรณ์ข้อมูลอนุกรมเวลาอื่น ๆ ซึ่งแนวทางการวิจัยในอนาคตอาจมีการเพิ่มชุดข้อมูลให้มีจำนวนมากขึ้น เพื่อนำมาใช้ในการพัฒนาและประเมินประสิทธิภาพแบบจำลอง และปรับค่าพารามิเตอร์ต่าง ๆ ให้เหมาะสม การศึกษาปัจจัยที่อาจจะส่งผลต่อการเกิด overfitting รวมไปถึงอาจศึกษาปัจจัยอื่น ๆ เพิ่มเติมและประยุกต์ใช้วิธีการแบบผสมผสาน (hybrid technique) ซึ่งอาจส่งผลต่อประสิทธิภาพของแบบจำลองการพยากรณ์ให้ดียิ่งขึ้น

## 6. กิตติกรรมประกาศ

งานวิจัยนี้ได้รับความอนุเคราะห์ข้อมูลจากสำนักงานเศรษฐกิจการเกษตร และสมาคมแป้งมันสำปะหลังไทย ขอขอบคุณ คณะเทคโนโลยีสารสนเทศและบัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ที่ให้การสนับสนุนงานวิจัยในครั้งนี้

## 7. รายการอ้างอิง

[1] สำนักงานเศรษฐกิจ, โอกาสสินค้าเกษตรสู่ประชาคมอาเซียน, กระทรวงเกษตรและสหกรณ์, แหล่งที่มา : http://www.oae.go.th/download/ownload_journal/2558/yearbook57.pdf, 5 พฤศจิกายน 2558.

[2] สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติปทุมธานี, 2554, ยุทธศาสตร์วิจัยและพัฒนาอุตสาหกรรมมันสำปะหลังประเทศไทย (พ.ศ. 2555-2559) และโปรแกรมวิจัยและพัฒนามันสำปะหลังภายใต้แผนกลยุทธ์การวิจัยและพัฒนา สวทช. ระยะที่ 2 พ.ศ. 2554-2559 (Strategic Planning Alliance II: SPA II), สำนักงานพัฒนาวิทยาศาสตร์และเทคโนโลยีแห่งชาติ, ปทุมธานี.

[3] สำนักงานเศรษฐกิจ, สถิติการเกษตรของประเทศไทยประจำปี2557, กระทรวงเกษตรและสหกรณ์, แหล่งที่มา : http://www.oae.go.th/download/download_journal/2558/yearbook57.pdf, 20 ตุลาคม 2558.

[4] Sunasuan, S. , Sombatvichathorn, Y. , Thanapongpipat, S., Duangmanee, K. and Woradit, S., 2011, The study of Agricultural products of cassava, Technical report, The Office of the Agricultural Futures Trading Commission of Thailand (AFTC).

[5] Pattranurakyothin, T. and Kumnungkit, K., 2012, Forecasting model for para rubbers export sales, KMITL Sci. Technol. J. 12(2): 198-202.

[6] Udomsri, N., Kengpol, A., Ishii, K. and

Shimada, Y., 2013, The design of a forecasting support models on demand of durian for export markets by time series and ANN, KMUTNB Int. J. Appl. Sci. Technol. 4(2): 49-65.

[7] Adebiyi, A.A., Adebiyi, A.O., Adewumi and Ayo, C.K., 2014, Comparison of ARIMA and artificial neural networks models for stock price prediction, J. Appl. Math. 2: 1-7.

[8] Pannakong, W., Huyun, V. and Sririviboonchitta, S., 2016, ARIMA versus artificial neural network for Thailand's cassava starch export forecasting, Casuals Interference Econometric 622: 255-227.

[9] Michel, B., van den Dirk, P., Nathalie, H. and Ruben, G., 2015, Evaluating multiple classifiers for stock price direction prediction, Exp. Syst. Appl. 42: 7046-7056.

[10] Weisent, J., Seaver, W., Odol, A. And Rohrbach, B., 2010, Comparison of three time-series models for predicting ampylobacteriosis Risk, Epidemiol. Infect 138: 898-906.

[11] Ahmad, S., 2015, Providing a measure for bullwhip effect in a two-product supply chain with exponential smoothing forecasts, Int. J. Prod. Econ. 169: 44-54.

[12] Chiulli, R.M., 1999, Quantitative Analysis: An Introduction, Gordon and Breach Science Publishers, imprint.

[13] Jyothi, P. and Mytri, V.D., 2013, A predictition model for population dynamics of cotton pest (*Thrips tabaci* Linde) using multilayer-perceptron neural network, Int. J. Comp. Appl. 67(4): 19-26.

[14] Neto, M.C.A., Tavares, G., Alves, V.M.O., Cavalcanti, G.D.C. and Tsang, I.R., 2010, Improving financial time series prediction using exogenous series and neural networks committees, In The 2010 International Joint Conference on Neural Networks (IJCNN): 1-8.

[15] Chen, R.J., Bloomfield, P. and Fu, J.S., 2003, An evaluation of alternative forecasting methods to recreation visitation, J. Leisure Res. 35: 441-454.