

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การทำนายข้อมูลอนุกรมเวลาโดยใช้โครงข่ายประสาทเทียมแบบ
ซิมพลิไฟด์ฟัซซี่อาทแมพและแบคพรอพพาเกชั่น

HYBRID TIME SERIES FORECASTING SYSTEM USING SIMPLIFIED
FUZZY ARTMAP AND BACKPROPAGATION NEURAL NETWORKS



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

วพ.

จ 475 ก

๒๕๔๗

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2547

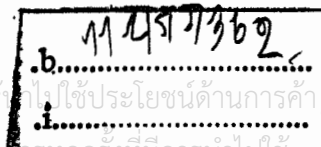
ISBN 974-15-1166-3

เลขหมู่.....

เลขทะเบียน..... 56671 <

ออกสารบัญ..... 12 ก.ค. 2548

วัน,เดือน,ปี.....



สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรณีไปใช้

HYBRID TIME SERIES FORECASTING SYSTEM USING SIMPLIFIED
FUZZY ARTMAP AND BACKPROPAGATION NEURAL NETWORKS



A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
SCHOOL OF GRADUATE STUDIES
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2004

ISBN 974-15-1166-3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2004

SCHOOL OF GRADUATE STUDIES

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์

การทำนายข้อมูลอนุกรมเวลาโดยใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซีอาทแมพและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

นักศึกษา

นางสาวจินดารัตน์ จันทร์อุบล

รหัสประจำตัว

44067028

ปริญญา

วิทยาศาสตรมหาบัณฑิต

สาขาวิชา

เทคโนโลยีสารสนเทศ

พ.ศ.

2547

อาจารย์ผู้ควบคุมวิทยานิพนธ์

รศ.ดร.อาริต ธรรมโน

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้ นำเสนอวิธีการทำนายข้อมูลอนุกรมเวลาซึ่งเป็นโมเดลแบบผสมที่ได้จากการนำเอาโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซีอาทแมพ (Simplified Fuzzy ARTMAP) ร่วมกับโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน (Backpropagation) มาใช้ร่วมกัน โดยใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซีอาทแมพ ซึ่งเป็นโครงข่ายประสาทเทียมที่มีความสามารถในการ classification ได้ดีมาทำนายทิศทางของข้อมูลอนุกรมเวลาก่อน จากนั้นจึงใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันทำนายค่าข้อมูลในอนาคต ซึ่งผลการทดลองปรากฏว่าวิธีที่พัฒนาขึ้นนี้ให้ความถูกต้องในการทำนายทิศทางสูงและมีความผิดพลาดในการทำนายต่ำกว่าเมื่อเปรียบเทียบกับการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันและการทำนายโดยวิธี ARIMA

Thesis Title Hybrid Time Series Forecasting System Using Simplified Fuzzy
ARTMAP and Backpropagation Neural Networks

Student Ms.Jindarat Juntauppalee

Student ID. 44067028

Degree Master Science

Program Information Technology

Year 2004

Thesis Advisor Assc.Prof.Dr. Arit Thammano

ABSTRACT

This thesis propose a hybrid time series forecasting system using a combination of Simplified Fuzzy ARTMAP and Backpropagation neural networks. Simplified Fuzzy ARTMAP which is suitable for classification, is used to find the direction of time series data and Backpropagation neural networks is used to forecast the value. The results show that this hybrid model outperforms basic Backpropagation neural network and ARIMA on both percent of direction correct and error measures.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลงได้อย่างดี ด้วยคำแนะนำและคำปรึกษาจาก รศ.ดร.อาทิตย์ ธรรมโน ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ผู้วิจัยรู้สึกซาบซึ้งในความอนุเคราะห์และขอขอบคุณ เป็นอย่างสูง

ขอบคุณทุกๆ คนในครอบครัวที่ให้โอกาสและให้กำลังใจในการเรียนมาตลอด
ขอบคุณเพื่อนร่วมชั้นเรียนที่ให้คำแนะนำและช่วยเหลือในการทำวิทยานิพนธ์ฉบับนี้
คุณงามความดีที่เกิดจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบให้กับคุณพ่อและคุณแม่อันเป็นที่รักของลูก รวมทั้งอาจารย์ทุกท่านที่ได้ถ่ายทอดความรู้และประสบการณ์แก่ข้าพเจ้า

จินดารัตน์ จันทอุปพี



สารบัญ

	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VIII
สารบัญรูป	XI
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย	2
1.3 ขอบเขตงานวิจัย	2
1.4 วิธีดำเนินงานวิจัย	2
1.5 การจัดเรียงหัวข้อในการเรียบเรียงวิทยานิพนธ์	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	4
2.1 การทำนาย (Forecasting)	4
2.2 เทคนิคการทำนาย (Forecasting Technique).....	5
2.3 ข้อมูลอนุกรมเวลา (Time Series Data).....	7
2.4 การวัดความถูกต้องในการทำนาย (Forecasting Measure).....	11
2.5 โครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ (Fuzzy ARTMAP Neural Network).....	12
2.5.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ	13
2.5.2 การเรียนรู้ของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ	14
2.6 โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ (Simplified Fuzzy ARTMAP)	19
2.6.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ.....	19
2.6.2 การเรียนรู้ของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ.....	21
2.6.3 การทำนายของซิมพลิไฟด์ฟัซซี่อาทแมพ (Simplified Fuzzy ARTMAP Classification).....	23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และ IV ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

2.7	โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน	23
2.7.1	สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน.....	24
2.7.2	การเรียนรู้ของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน	26
2.8	Autoregressive Integrated Moving Average (ARIMA).....	29
2.8.1	เครื่องหมาย B (Backshift Notation).....	29
2.8.2	ขั้นตอนการทำนายโดยวิธี ARIMA	30
2.8.3	อนุกรมเวลาที่เป็นสแตชันนารี (Stationary Time Series).....	31
2.8.4	การกำหนดรูปแบบ (Model Identification).....	35
2.8.5	การทดสอบความเหมาะสมของรูปแบบ.....	40
2.8.6	ตัวอย่างการทำนายโดยวิธี ARIMA.....	45
บทที่ 3	การทำนายข้อมูลอนุกรมเวลาโดยโมเดลแบบผสม	52
3.1	การแสดงรูปแบบข้อมูลอนุกรมเวลา	52
3.1.1	การนอร์มัลไลซ์ข้อมูลอนุกรมเวลา	52
3.1.2	แพทเทินอินพุตและแพทเทินเอาต์พุตสำหรับโครงข่ายประสาทเทียม แบบซิมพลิไฟด์พีชชี้อาทแมพ	54
3.1.3	แพทเทินอินพุตและแพทเทินเอาต์พุตสำหรับโครงข่ายประสาทเทียม แบบแบคพรอพพาเกชัน	55
3.2	สถาปัตยกรรมของโมเดลแบบผสม.....	55
3.3	การเรียนรู้ของโมเดลแบบผสม	57
3.3.1	การเรียนรู้ในส่วนทำนายทิศทาง	57
3.3.2	การเรียนรู้ส่วนทำนายค่า.....	57
3.4	การทำนายโดยโมเดลแบบผสม.....	58
บทที่ 4	การทดลองและผลการทดลอง.....	59
4.1	ข้อมูลที่ใช้ในการทดลอง.....	59
4.2	การทำนายโดยโมเดลแบบผสม.....	65
4.2.1	การกำหนดค่าพารามิเตอร์.....	65

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา IV นี้ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
4.2.2 ข้อมูลที่ใช้ในการทดลอง.....	67
4.2.3 การเรียนรู้.....	67
4.2.4 วิธีการทำนาย.....	68
4.2.5 ผลการทดลอง.....	68
4.2.6 สรุปผลการทดลอง.....	70
4.3 การทำนายโดยโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	71
4.3.1 การกำหนดค่าพารามิเตอร์.....	71
4.3.2 ข้อมูลที่นำมาใช้ในการทดลอง.....	72
4.3.3 การเรียนรู้.....	72
4.3.4 วิธีการทำนาย.....	72
4.3.5 ผลการทดลอง.....	73
4.3.6 สรุปผลการทดลอง.....	74
4.4 สรุปผลการเปรียบเทียบการทำนายโดยโมเดลแบบผสมและโครงข่ายประสาท เทียมแบบแบคพรอพพาเกชั่น	76
4.5 รูปแสดงผลการทำนายโดยโมเดลแบบผสมเปรียบเทียบกับการทำนายโดยใช้ โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	77
4.6 การทำนายโดยวิธี Autoregressive Integrated Moving Average (ARIMA).....	82
4.6.1 การกำหนดรูปแบบและค่าพารามิเตอร์.....	82
4.6.2 การทำนาย.....	82
4.6.3 ข้อมูลที่ใช้ทดลอง.....	82
4.6.4 ผลการทดลอง.....	82
4.7 สรุปผลการเปรียบเทียบการทำนายโดยโมเดลแบบผสมและวิธี ARIMA.....	83
4.8 รูปแสดงผลการทำนายโดยโมเดลแบบผสมเปรียบเทียบกับการทำนายโดยวิธี ARIMA.....	83
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	88

สารบัญ (ต่อ)

	หน้า
5.1 สรุปผลการวิจัย.....	88
5.2 ข้อเสนอแนะ.....	89
เอกสารอ้างอิง	90
ภาคผนวก	92
ประวัติผู้เขียน	98



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา VII นี้ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 แนวทางการกำหนดรูปแบบสำหรับอนุกรมเวลาที่ไม่มีฤดูกาล.....	38
2.2 แนวทางการกำหนดรูปแบบสำหรับอนุกรมเวลาที่มีฤดูกาล.....	39
3.1 การนอร์มัลไลซ์ข้อมูลอนุกรมเวลา.....	53
3.2 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ SFAM ก่อนทำ complement coding.....	54
3.3 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ SFAM หลังทำ complement coding.....	54
3.4 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ BPNN.....	55
4.1 แสดงค่าพารามิเตอร์ในการทำนายโดยโมเดลแบบผสม.....	66
4.2 แสดง period ของข้อมูลและขนาดแพทเทินอินพุท.....	67
4.3 ผลการทดลองที่ 1.....	68
4.4 ผลการทดลองที่ 2.....	69
4.5 ผลการทดลองที่ 3.....	69
4.6 สรุปผลการทำนายโดยโมเดลแบบผสม.....	71
4.7 แสดงค่าพารามิเตอร์ในการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	71
4.8 ผลการทดลองที่ 4.....	73
4.9 ผลการทดลองที่ 5.....	73
4.10 ผลการทดลองที่ 6.....	74
4.11 สรุปผลการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	75
4.12 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	76
4.13 ผลการทำนายโดยวิธี ARIMA.....	83

สารบัญรูป

รูปที่	หน้า
2.1 ข้อมูลอนุกรมเวลาที่ลักษณะแนวโน้ม.....	8
2.2 ข้อมูลอนุกรมเวลาที่ลักษณะผันแปรตามวัฏจักร.....	9
2.3 ข้อมูลอนุกรมเวลาที่ลักษณะผันแปรตามฤดูกาล.....	10
2.4 ข้อมูลอนุกรมเวลาที่ลักษณะไม่แน่นอน	11
2.5 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบพีชชี้อาทแมพ	14
2.6 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพ	20
2.7 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น.....	24
2.8 ฟังก์ชัน sigmoid.....	26
2.9 แสดงลักษณะอนุกรมเวลาที่ไม่เป็นสเตชันนารีเนื่องจากแนวโน้มและอิทธิพลฤดูกาล....	32
2.10 แสดงลักษณะอนุกรมเวลาที่ไม่เป็นสเตชันนารีเนื่องจากแนวโน้ม.....	33
2.11 แสดงลักษณะอนุกรมเวลาที่เป็นสเตชันนารีหลังจากลบแนวโน้ม.....	33
2.12 แสดงลักษณะอนุกรมเวลาที่เป็นสเตชันนารีหลังจากลบแนวโน้มและฤดูกาล.....	34
2.13 แสดงลักษณะของ ACF หรือ PACF ที่หายไป(Cut Off).....	36
2.14 แสดงลักษณะของ ACF หรือ PACF ที่ลดลงแบบ exponential แต่ไม่แกว่ง.....	36
2.15 แสดงลักษณะของ ACF หรือ PACF ที่ลดลงแบบ exponential แต่แกว่ง.....	37
2.16 แสดงลักษณะของ ACF หรือ PACF ที่มีรูปแบบ sine wave.....	37
2.17 กราฟ ACF ของ residuals ที่เป็นอิสระกัน.....	41
2.18 กราฟ PACF ของ residuals ที่เป็นอิสระกัน.....	41
2.19 แสดงลักษณะของข้อมูลจริง.....	45
2.20 แสดง ACF ของข้อมูลจริง.....	46
2.21 แสดง PACF ของข้อมูลจริง.....	46
2.22 แสดงลักษณะของข้อมูลหลังทำ first differencing.....	47
2.23 ACF ของข้อมูลหลังทำ first differencing.....	47
2.24 PACF ของข้อมูลหลังทำ first differencing.....	48
2.25 ACF ของ residuals.....	49
2.26 PACF ของ residuals.....	50
3.1 สถาปัตยกรรมของโมเดลแบบผสม.....	56
4.1 ลักษณะของข้อมูลชุดที่ 1.....	60

สารบัญญรูป (ต่อ)

รูปที่	หน้า
4.2 ลักษณะของข้อมูลชุดที่ 2.....	61
4.3 ลักษณะของข้อมูลชุดที่ 3.....	61
4.4 ลักษณะของข้อมูลชุดที่ 4.....	62
4.5 ลักษณะของข้อมูลชุดที่ 5.....	62
4.6 ลักษณะของข้อมูลชุดที่ 6.....	63
4.7 ลักษณะของข้อมูลชุดที่ 7.....	63
4.8 ลักษณะของข้อมูลชุดที่ 8.....	64
4.9 ลักษณะของข้อมูลชุดที่ 9.....	64
4.10 ลักษณะของข้อมูลชุดที่ 10.....	65
4.11 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 1.	77
4.12 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 2.	78
4.13 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 3.	78
4.14 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 4.	79
4.15 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 5.	79
4.16 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 6.	80
4.17 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 7.	80
4.18 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 8.	81
4.19 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 9.	81
4.20 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 10.....	82
4.21 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 3.....	84
4.22 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 4.....	84
4.23 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 5.....	85
4.24 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 6.....	85

สารบัญญรูป (ต่อ)

รูปที่	หน้า
4.25 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบ ชุดที่ 7.....	86
4.26 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบ ชุดที่ 8.....	86
4.27 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบ ชุดที่ 9.....	87
4.28 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบ ชุดที่ 10.....	87



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในเชิงการบริหารจัดการ การทำนายถือว่ามีผลสำคัญอย่างยิ่งต่อการวางแผนและการตัดสินใจในการดำเนินงานกิจกรรมต่างๆ ภายในองค์กร เช่น การตัดสินใจและการวางแผนในการลงทุน การตัดสินใจและการวางแผนการผลิตสินค้าและการขนส่งสินค้า การทำนายยอดขายในอนาคตขององค์กร เป็นต้น ดังนั้น หากผู้บริหารหรือนักวางแผนสามารถทำนายเพื่อค้นหาแนวโน้มที่เป็นไปได้และปัจจัยต่างๆ ที่เอื้อต่อการเกิดเหตุการณ์นั้นในอนาคตด้วยความเชื่อมั่นระดับหนึ่งก็จะทำให้ตัดสินใจหรือวางแผนงานได้ถูกต้องมากขึ้นซึ่งจะช่วยลดความเสี่ยงและลดต้นทุนอันจะส่งผลให้องค์กรทำงานได้อย่างมีประสิทธิภาพมากขึ้นและมีโอกาสประสบความสำเร็จในการทำงานมากขึ้น อย่างไรก็ตาม การเกิดเหตุการณ์หรือสภาพการณ์หนึ่งในอนาคตเป็นการเกิดภายใต้ความไม่แน่นอน ดังนั้น การทำนายที่ให้ความถูกต้องสูงจึงถือว่ามีผลสำคัญอย่างยิ่ง

โครงข่ายประสาทเทียม (Neural Network) ได้รับความสนใจและถูกพัฒนาไปประยุกต์ใช้อย่างแพร่หลายในงานหลายด้าน เช่น การจัดกลุ่มข้อมูล (Classification) การรู้จำรูปแบบ (Pattern Recognition) รวมทั้งการทำนายข้อมูลอนุกรมเวลา (Time Series Forecasting) โดยเฉพาะการทำนายข้อมูลอนุกรมเวลาได้ถูกวิจัยอย่างกว้างขวาง เช่น Sameer Singh และ Jonathan Fieldsend [1] ได้สร้างโมเดลแบบผสมที่ใช้ Pattern Matching Recognition System และโครงข่ายประสาทเทียมแบบแคพรอพพาเกชันทำนายข้อมูลอนุกรมเวลาเปรียบเทียบกับการทำนายโดยใช้ Pattern Matching Recognition System และโครงข่ายประสาทเทียมแบบแคพรอพพาเกชันเพียงอย่างเดียว Ricardo Bastos Cavalcante Prudencio และ Teresa Bernardo Ludermir [2] ได้ใช้เจเน็ติกอัลกอริทึมในการหาค่าพารามิเตอร์ที่เหมาะสมเพื่อใช้สร้างสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแคพรอพพาเกชัน และ Juan M. Corchado และ Jim Aiken [3] ได้ใช้ Case-Based Reasoning System ร่วมกับโครงข่ายประสาทเทียมแบบ Radial Basis Function ในการทำนายอุณหภูมิของมหาสมุทรแบบ Real time จากงานวิจัยเหล่านี้แสดงให้เห็นว่าการทำนายโดยใช้หลายวิธีผสมกันให้ผลการทำนายที่ถูกต้องมากกว่าการทำนายโดยใช้วิธีเดียวทำนาย ดังนั้น ผู้วิจัยจึงมีแนวคิดว่าถ้านำเอาหลายวิธีมารวมกันน่าจะให้ผลการทำนายดีกว่าการทำนายโดยใช้วิธีเดียว

เนื่องจากลักษณะของข้อมูลอนุกรมเวลามีการเปลี่ยนแปลงทิศทางขึ้นลงตลอดเวลาและจากการสังเกตวิธีการทำนายข้อมูลอนุกรมเวลาโดยทั่วไปไม่ว่าจะเป็นวิธีทางสถิติหรือวิธีทาง

ปัญหาประดิษฐ์พบว่าผลการทำนายมักจะ lag อยู่ 1 จุดเสมอทุกครั้งที่มีข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงทิศทาง ดังนั้น ผู้วิจัยจึงมีแนวคิดว่าจะทำนายหรือหาทิศทางของอนุกรมเวลาก่อน จากนั้นจึงค่อยทำนายค่าของอนุกรมเวลา และหากสามารถทำนายทิศทางของข้อมูลอนุกรมเวลาได้ถูกต้องก็น่าจะสามารถทำนายค่าได้ถูกต้องมากขึ้นด้วยและจะสามารถแก้ปัญหาวิธีการทำนายที่ lag ได้ด้วย ดังนั้น งานวิจัยฉบับนี้จึงได้นำเสนอโมเดลแบบผสมเพื่อใช้ในการทำนายข้อมูลอนุกรมเวลาที่ได้จากการนำเอาโครงข่ายประสาทเทียม 2 แบบมาใช้ร่วมกันนั่นคือ ใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพในการทำนายทิศทางของอนุกรมเวลาก่อน จากนั้นจึงใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันในการทำนายค่าที่จะเกิดขึ้นในอนาคต ซึ่งผู้วิจัยคาดว่าโมเดลแบบผสมนี้น่าจะทำนายผลได้ถูกต้องมากขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อสร้างโมเดลในการทำนายข้อมูลอนุกรมเวลาที่มีความสามารถในการทำนายทิศทางของอนุกรมเวลาได้และสามารถทำนายค่าที่จะเกิดขึ้นในอนาคตได้ถูกต้องมากขึ้น

1.3 ขอบเขตงานวิจัย

1. งานวิจัยนี้ใช้การทดลองแบบ off-line
2. โมเดลแบบผสมที่พัฒนาขึ้นใช้ทำนาย 1 step ล่วงหน้า
3. ในการทดลองนี้จะเปรียบเทียบความถูกต้องในการทำนายของโมเดลที่พัฒนาขึ้นกับการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันและการทำนายด้วยวิธี Autoregressive Integrated Moving Average (ARIMA)

1.4 ขั้นตอนการดำเนินงานวิจัย

1. ศึกษาวิธีการทำนายข้อมูลอนุกรมเวลาแบบต่างๆ เน้นวิธีที่ใช้ปัญญาประดิษฐ์
2. ศึกษาทฤษฎีที่เกี่ยวข้องกับการทำนายข้อมูลอนุกรมเวลา เช่น อนุกรมเวลา โครงข่ายประสาทเทียมแบบพีชชีอาทแมพ โครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพ โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน เป็นต้น
3. ออกแบบโมเดลในการทำนายข้อมูลอนุกรมเวลา
4. สร้างและพัฒนาโปรแกรมเพื่อใช้ในการทดสอบโมเดลที่ออกแบบไว้ โดยใช้ซอฟต์แวร์ MATLAB รุ่น 6.1
5. รวบรวมข้อมูลเพื่อนำมาทดสอบโมเดล
6. ทดลองและวัดประสิทธิภาพของโมเดลและสรุปผลการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. เรียบเรียงเอกสารประกอบวิทยานิพนธ์

1.5 การจัดเรียงหัวข้อในการเรียบเรียงวิทยานิพนธ์

โครงสร้างของวิทยานิพนธ์ฉบับนี้ประกอบด้วย 5 บทคือ

บทที่ 1 บทนำ กล่าวถึงประวัติและความเป็นมา วัตถุประสงค์ของงานวิจัย ข้อจำกัดและขอบเขตงานวิจัยและขั้นตอนวิธีดำเนินงานวิจัย

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง กล่าวถึงทฤษฎีที่เกี่ยวข้องซึ่งมีเนื้อหาเกี่ยวกับสถาปัตยกรรมและการเรียนรู้ของโครงข่ายประสาทเทียมแบบพีชชีอาทแมพ โครงข่ายประสาทเทียมแบบซิมพลีไฟด์พีชชีอาทแมพ โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันและการทำนายโดยวิธี Autoregressive Integrated Moving Average (ARIMA)

บทที่ 3 โมเดลทำนายข้อมูลอนุกรมเวลาแบบผสม กล่าวถึงสถาปัตยกรรม การเรียนรู้และการทำนายของโมเดลแบบผสมที่พัฒนาขึ้น

บทที่ 4 การทดลองและผลการทดลอง กล่าวถึงการออกแบบการทดลอง การกำหนดค่าพารามิเตอร์ ข้อมูลที่ใช้ในการทดลอง ผลการทดลองและสรุปผลการทดลอง

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1 การทำนาย (Forecasting)

การทำนาย คือการคาดคะเนการเกิดของเหตุการณ์หรือสภาพการณ์ต่างๆ ในอนาคต โดยการศึกษาแนวโน้มและรูปแบบการเกิดของเหตุการณ์จากข้อมูลในอดีตหรือใช้ความรู้และความสามารถ ประสบการณ์และวิจารณญาณของผู้ทำนาย

การทำนายมีบทบาทสำคัญต่อการตัดสินใจในองค์กร ทั้งองค์กรในภาครัฐบาลและภาคเอกชน ปัจจุบันความต้องการในการทำนายมีมากขึ้นเพราะผู้บริหารพยายามที่จะลดการขึ้นอยู่กับโอกาสหรือลดความเสี่ยงมาเป็นการจัดการอย่างมีเหตุผลแทน เช่น รัฐบาลต้องการการทำนายเกี่ยวกับคุณภาพน้ำ คุณภาพอากาศ อัตราการว่างงาน อัตราเงินเฟ้อ รายจ่าย รายได้จาก การจัดเก็บภาษีต่างๆ การส่งออก เป็นต้นเพื่อนำมากำหนดนโยบายของรัฐบาล มหาวิทยาลัย ต้องการการทำนายเกี่ยวกับจำนวนนักศึกษาที่จะลงทะเบียนเพื่อให้สามารถตัดสินใจเกี่ยวกับทรัพยากรที่ต้องใช้ในการเรียนการสอน ในองค์กรทางธุรกิจแล้วการทำนายมีบทบาทสำคัญในเรื่องต่อไปนี้ เช่น

1. ด้านการตลาด หากสามารถทำนายความต้องการสินค้าได้ก็จะทำให้สามารถกำหนดแผนการขาย แผนการสนับสนุนการขาย (Promotion) และถ้าหากสามารถทำนายได้ว่ากลุ่มเป้าหมายของสินค้าคือกลุ่มใดก็จะทำให้สามารถวางแผนการโฆษณาสินค้าได้ถูกต้องตรงกลุ่มเป้าหมายยิ่งขึ้น อันจะส่งผลดีต่อยอดขายสินค้า
2. ด้านการเงิน การทำนายอัตราดอกเบี้ยจะถูกใช้ในการวางแผนการให้ได้มาซึ่งเงินลงทุน การทำนายรายรับและรายจ่ายถูกใช้เพื่อคาดคะเนกระแสเงินสด เป็นต้น
3. ด้านการบริหารจัดการบุคลากร การทำนายเกี่ยวกับความต้องการบุคลากรในสาขาต่างๆ จะถูกนำไปใช้ในการวางแผนการรับสมัครพนักงานใหม่และการฝึกอบรม เป็นต้น
4. ด้านการวางแผนการผลิต การทำนายความต้องการสินค้าแต่ละชนิดในแต่ละเดือนหรือสัปดาห์จะทำให้สามารถวางแผนการผลิตและการควบคุมสินค้าคงคลังได้เหมาะสม รวมทั้งสามารถนำไปวางแผนการซื้อวัตถุดิบที่ต้องใช้ในการผลิตได้อีกด้วย
5. ด้านการบริหารจัดการยุทธศาสตร์ในการทำงาน การทำนายสภาพทางเศรษฐกิจ การเปลี่ยนแปลงราคาและต้นทุน การเปลี่ยนแปลงเทคโนโลยี การเติบโตของตลาดจะช่วยในการวางแผนอนาคตของบริษัทในระยะยาวได้ เช่นการวางแผนการซื้อเครื่องจักรและส่วนประกอบที่ต้องใช้ในอนาคต เป็นต้น

2.2 เทคนิคการทำนาย (Forecasting Technique)

ปัจจุบันเทคนิคการทำนายได้รับการพัฒนาไปมากมายแต่สามารถแบ่งได้ 2 กลุ่มคือ เทคนิคทางสถิติและเทคนิคด้านปัญญาประดิษฐ์

2.2.1 เทคนิคทางสถิติ (Statistical Technique)

การทำนายทางสถิติสามารถแบ่งออกเป็น 2 กลุ่มย่อย คือ การทำนายเชิงปริมาณและการทำนายเชิงคุณภาพ [4]

1. การทำนายเชิงปริมาณ (Quantitative Forecasting)

การทำนายเชิงปริมาณเน้นที่การนำข้อมูลในอดีตมาเป็นแนวทางการทำนายโดยนำข้อมูลที่เก็บรวบรวมไว้มาสร้างเป็นโมเดลทางคณิตศาสตร์ การทำนายแบบนี้จะกระทำได้อาจต้องมีเงื่อนไขคือ ต้องมีข้อมูลในอดีตและข้อมูลนั้นต้องสามารถวัดออกมาเป็นตัวเลขได้และสันนิษฐานว่ารูปแบบของข้อมูลในอดีตจะเกิดขึ้นอีกในอนาคต ซึ่งสามารถแบ่งโมเดลที่ใช้ทำนายได้ 2 แบบ ดังนี้

1.1 Explanatory Model

การทำนายแบบนี้ตั้งบนสมมติฐานที่ว่าตัวแปรที่จะทำนายมีความสัมพันธ์กับตัวแปรอิสระอย่างน้อย 1 ตัวแปร ดังนั้นโมเดลนี้จะทำนายโดยการหาความสัมพันธ์ระหว่างตัวแปรที่ต้องการทำนายกับตัวแปรอื่นๆ โดยมีจุดมุ่งหมายเพื่อค้นหารูปแบบของความสัมพันธ์ของตัวแปรที่ต้องการทำนายกับตัวแปรที่เกี่ยวข้องและใช้รูปแบบนั้นทำนายค่าในอนาคต ตัวอย่างเช่น

$$GNP = f(\text{นโยบายด้านการเงิน, รายได้แผ่นดิน, ภาวะเงินเฟ้อ, รายจ่ายของประเทศ, การส่งออก, การนำเข้าและค่าความผิดพลาด}) \quad (2.1)$$

จากเงื่อนไขที่ (2.1) จะเห็นว่า ผลิตภัณฑ์มวลรวมประชาชาติ (Gross National Product : GNP) ขึ้นอยู่กับ นโยบายด้านการเงิน รายได้แผ่นดิน ภาวะเงินเฟ้อ รายจ่ายของประเทศ การส่งออก การนำเข้าและค่าความผิดพลาด การรวมค่าความผิดพลาดเข้าไปด้วย เนื่องจากความสัมพันธ์นี้ไม่ได้เป็นความสัมพันธ์ที่แน่นอนอาจจะมีตัวแปรอื่นที่มีผลกระทบต่อเปลี่ยนแปลงค่าผลิตภัณฑ์มวลรวมประชาชาติด้วย การทำนายด้วยโมเดลแบบนี้ถ้ามีการเปลี่ยนแปลงใดๆ ของตัวแปรที่เป็นอินพุทจะมีผลกระทบต่อเอาต์พุทที่ได้จากการทำนาย โดยสันนิษฐานว่าความสัมพันธ์ของตัวแปรนั้นไม่เปลี่ยนแปลง

1.2. Time Series Model

การทำนายแบบนี้จะมองระบบเหมือนเป็นกล่องดำ (Black Box) และไม่พยายามค้นหาปัจจัยที่มีผลกระทบต่อตัวแปรที่ต้องการทำนาย การทำนายแบบนี้จึงใช้เพียงข้อมูลในอดีตและ

หรือค่าความผิดพลาดที่ผ่านมาโดยจะไม่พิจารณาตัวแปรอื่นที่มีผลกระทบต่อตัวแปรที่ต้องการทำนาย จุดประสงค์ของการทำนายโดยโมเดลนี้คือค้นหารูปแบบข้อมูลในอดีตและใช้รูปแบบข้อมูลในอดีตนั้นประมาณรูปแบบหรือค่าของข้อมูลในอนาคต เช่น ค่าผลิตภัณฑ์มวลรวมประชาชาติในอนาคตขึ้นอยู่กับค่าผลิตภัณฑ์มวลรวมประชาชาติในอดีต ดังเงื่อนไขที่ (2.2)

$$GNP_{t+1} = f(GNP_t + GNP_{t-1} + GNP_{t-2} + Error) \quad (2.2)$$

สมมติ t คือเดือนปัจจุบัน จากเงื่อนไขอธิบายได้ว่า ค่าผลิตภัณฑ์มวลรวมประชาชาติของเดือนหน้า ($t+1$) ขึ้นอยู่กับค่าผลิตภัณฑ์มวลรวมประชาชาติของเดือนที่ผ่านมา การพิจารณาเฉพาะตัวเลขในอดีตแบบนี้ทำให้การทำนายทำได้ง่ายขึ้นเพราะไม่ต้องระบุตัวแปรที่เป็นอินพุทอย่างในเงื่อนไขที่ (2.1) เหตุผลของการมองระบบเป็นกล่องดำเนื่องจากการยากที่เข้าใจหรือวัดความสัมพันธ์ของตัวแปรแต่ละตัวและการทำงานมีจุดประสงค์อยู่ที่ต้องการทำนายว่าจะอะไรจะเกิดขึ้นไม่ได้ต้องการหาเหตุผลว่าทำไมจึงเกิดขึ้น

2. การทำนายเชิงคุณภาพ (Qualitative Forecasting)

การทำนายเชิงคุณภาพจะกระทำก็ต่อเมื่อไม่มีข้อมูลที่อยู่ในรูปของตัวเลขในอดีตหรือมีข้อมูลน้อยมากแต่มีความรู้เพียงพอที่จะทำนายได้ การทำนายวิธีนี้ผู้ทำนายต้องมีความรู้ความสามารถและประสบการณ์ในเรื่องที่จะทำนายได้เป็นอย่างดี ตัวอย่างการทำนายเชิงคุณภาพ เช่น การทำนายผลกระทบของราคาน้ำมันที่มีต่อความต้องการใช้น้ำมัน การทำนายความเร็วของการสื่อสารในปี 2020 เป็นต้น

ขั้นตอนพื้นฐานในการทำนายเชิงปริมาณ มีดังนี้

1. กำหนดปัญหา เป็นการทำความเข้าใจและกำหนดว่าต้องการทำนายในเรื่องใด การทำนายจะถูกใช้อย่างไร ผู้ที่ต้องการใช้การทำนายและจะใช้การทำนายนั้นให้เหมาะสมกับองค์กรได้อย่างไร เช่น องค์กรมีระบบควบคุมสินค้าคงคลังด้วยคอมพิวเตอร์ แต่ต้องการทำนายว่าในอนาคตควรจะมีผลิตสินค้าเท่าไรและมีความต้องการเท่าไรเพื่อลดต้นทุนในการเก็บสินค้า ซึ่งผู้ทำนายต้องระบุปัญหาให้เหมาะสม

2. รวบรวมข้อมูล เป็นการรวบรวมข้อมูลต่างๆ ที่เกี่ยวข้องกับปัญหาที่ได้กำหนดเอาไว้เพื่อใช้ในการทำนาย ข้อมูลมีอย่างน้อย 2 ชนิดคือข้อมูลที่เป็นตัวเลขและข้อมูลที่เป็นความคิดเห็นของผู้เชี่ยวชาญหรือผู้รอบรู้ การรวบรวมข้อมูลที่เป็นตัวเลขในอดีตเป็นสิ่งสำคัญเพราะข้อมูลเหล่านั้นจะถูกนำมาสร้างเป็นโมเดลเพื่อใช้ในการทำนายในอนาคต

3. การวิเคราะห์ข้อมูลเบื้องต้น การทำนายโดยใช้วิธีทางสถิติอาจจะนำข้อมูลมาสร้างเป็นกราฟเพื่อให้เห็นภาพโดยรวมของข้อมูล จุดประสงค์ของการสร้างกราฟเพื่อดูว่าข้อมูลมี

ลักษณะเป็นอย่างไร เช่น มีแนวโน้ม มีความผันแปรตามฤดูกาล มีความผันแปรตามวัฏจักร หรือ มีความผันแปรที่ไม่แน่นอนหรือไม่และคำนวณค่าทางสถิติของแต่ละชุดข้อมูล เช่น ความเฉลี่ย ค่าสูงสุด ค่าต่ำสุด ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์สหสัมพันธ์แบบออกโต ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบออกโต เป็นต้น เพื่อนำมาเป็นส่วนประกอบในการพิจารณาว่าจะเลือกใช้วิธีการทำนายแบบใด

4. การเลือกโมเดลในการทำนายที่เหมาะสม โดยพิจารณารูปแบบกราฟของชุดข้อมูล และค่าสถิติที่ได้จากขั้นตอนที่ 3 มาประกอบการตัดสินใจ

5. การใช้และประเมินโมเดลนั้น เมื่อเลือกโมเดลและกำหนดค่าพารามิเตอร์ที่เหมาะสมสำหรับโมเดลนั้นแล้ว จากนั้นก็นำโมเดลไปใช้และผู้ใช้ต้องประเมินข้อดีและข้อเสียของโมเดลเมื่อนำโมเดลนั้นไปใช้

2.2.2 เทคนิคด้านปัญญาประดิษฐ์ (Artificial Intelligent Technique)

เนื่องด้วยลักษณะของข้อมูลอนุกรมเวลาส่วนมากแล้วจะอยู่ในรูปแบบที่ไม่เป็นเชิงเส้น (Nonlinear) จึงได้เกิดแนวคิดในการนำเอาเทคนิคทางปัญญาประดิษฐ์มาใช้ในการประมาณฟังก์ชันของข้อมูลอนุกรมเวลาที่มีลักษณะไม่เป็นเชิงเส้นดังกล่าวและปัจจุบันเทคนิคนี้ได้รับความสนใจในการทำวิจัยอย่างกว้างขวาง วิธีทางปัญญาประดิษฐ์ที่ถูกนำมาใช้ในการทำนายข้อมูลอนุกรมเวลาเช่น โครงข่ายประสาทเทียม (Neural Network) ฟัซซี่ลอจิก (Fuzzy Logic) เจนเนติกอัลกอริทึม (Genetic Algorithm) ซึ่งอาจจะทำนายโดยใช้วิธีเดียวหรือใช้หลายๆ วิธีรวมกัน เทคนิคการทำนายเหล่านี้จะอาศัยการเรียนรู้ เช่น การทำนายโดยใช้โครงข่ายประสาทเทียมซึ่งเป็นเทคนิคที่จำลองการเรียนรู้ของสมองมนุษย์ กระบวนการในการเรียนรู้ก็จะรับข้อมูลเข้ามาและทำการคำนวณแล้วส่งค่าที่ได้ไปยังขั้นต่อไปจนได้ค่าเอาท์พุทออกมา และทำการปรับโครงข่ายตามค่าความผิดพลาดที่เกิดขึ้น เมื่อโครงข่ายหรือโมเดลเรียนรู้จนได้พารามิเตอร์ที่เหมาะสมแล้วก็จะนำโมเดลนั้นไปใช้ทำนายต่อไป

2.3 ข้อมูลอนุกรมเวลา (Time Series Data)

ข้อมูลอนุกรมเวลา (Time Series Data) คือ กลุ่มของค่าสังเกตที่เก็บรวบรวมตามเวลาที่มีระยะห่างเท่าๆ กันและต่อเนื่องกัน ดังนั้น ข้อมูลอนุกรมเวลาจึงเป็นค่าที่แสดงการเปลี่ยนแปลงของตัวแปรตัวใดตัวหนึ่งที่เปลี่ยนแปลงไปตามเวลา หน่วยของเวลาอาจเป็น ชั่วโมง วัน สัปดาห์ เดือน ปี เช่น ยอดขายรถยนต์ ตัวแปรคือจำนวนรถยนต์ที่ขายได้ของแต่ละเดือนที่อาจไม่เท่ากัน หรือราคาหุ้นของธนาคารกรุงไทยที่เปลี่ยนแปลงไปในแต่ละวัน เป็นต้น

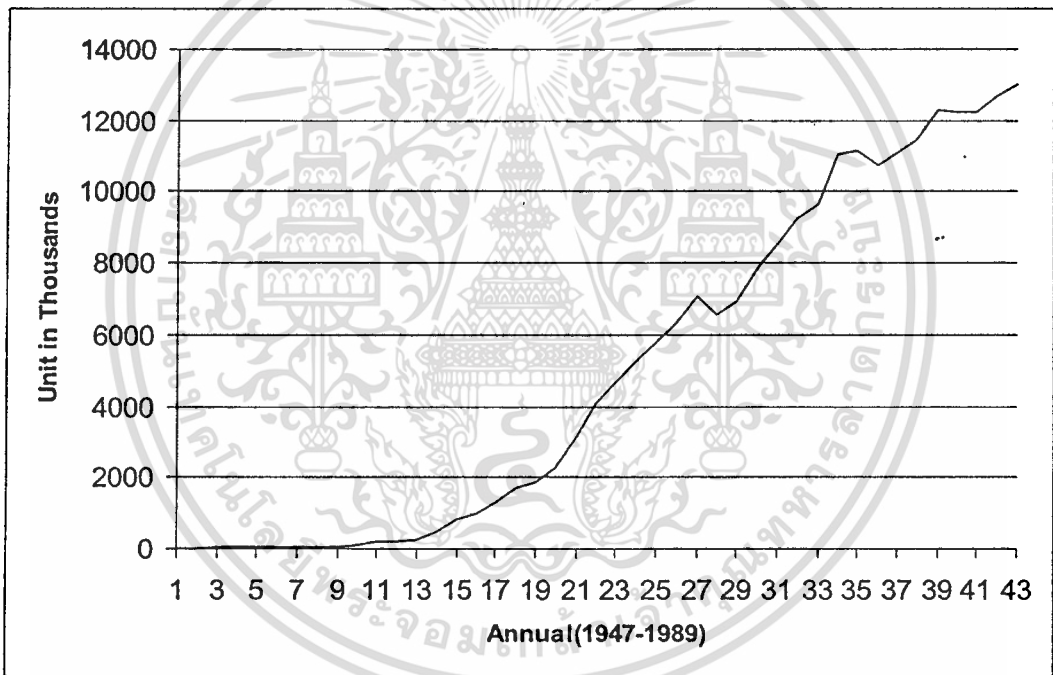
ดังนั้น การทำนายข้อมูลอนุกรมเวลา (Time Series Forecasting) ก็คือการศึกษาหารูปแบบการเปลี่ยนแปลงของตัวแปรที่เปลี่ยนไปตามเวลาตั้งแต่อดีตจนถึงปัจจุบัน แล้วนำรูปแบบ

นั้นมาวิเคราะห์เพื่อทำนายค่าของตัวแปรนั้นในอนาคตโดยจะถือเสมือนว่าสิ่งแวดล้อมหรือปัจจัยที่เกี่ยวข้องในอนาคตเหมือนหรือใกล้เคียงกับสิ่งแวดล้อมในอดีตและจะไม่พิจารณาตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ทำนายหรือที่มีผลต่อพฤติกรรมของระบบ

ถ้านำข้อมูลอนุกรมเวลามาวิเคราะห์จะพบว่าข้อมูลอนุกรมเวลามีส่วนประกอบ 4 ส่วน ดังนี้คือ

1. แนวโน้ม (Trend : T)

แนวโน้ม คือลักษณะของข้อมูลอนุกรมเวลาที่มีการเปลี่ยนแปลงในระยะเวลาที่นานพอที่จะเห็นแนวโน้มของข้อมูลว่าในอนาคตจะเพิ่มขึ้นหรือลดลง และระยะเวลาที่ทำให้สามารถเห็นแนวโน้มจะต้องนานกว่า 1 ปี ตัวอย่างเป็นข้อมูลการผลิตรถยนต์ของญี่ปุ่น ตั้งแต่ปี ค.ศ. 1947 ถึงปี ค.ศ.1989 หน่วยเป็นพันคัน [5] ซึ่งมีแนวโน้มที่เพิ่มขึ้น

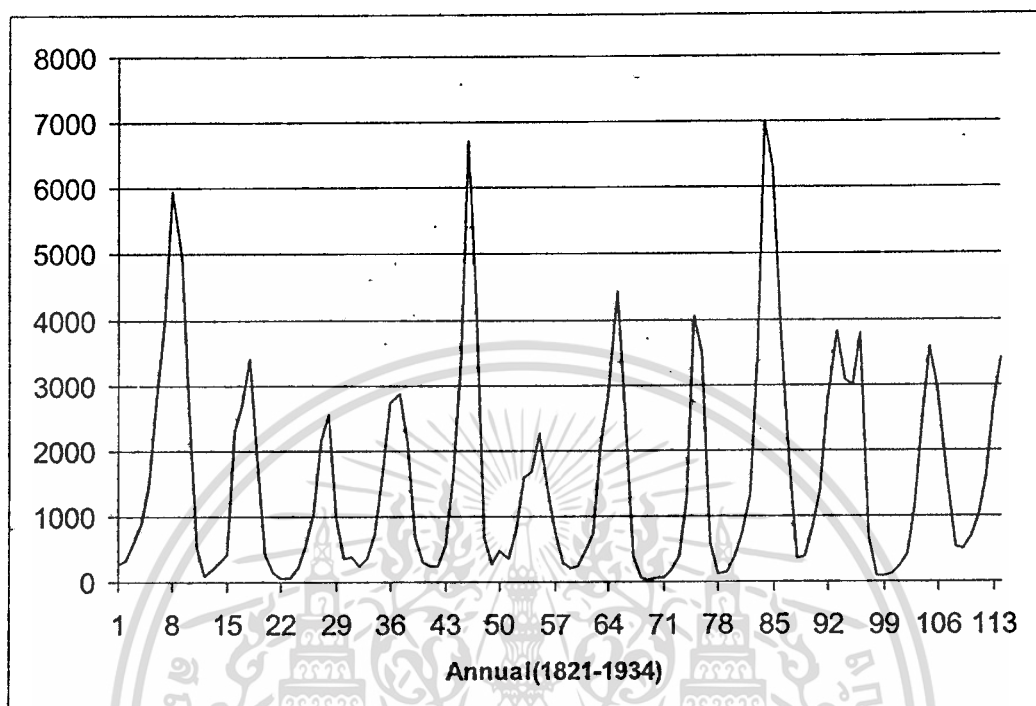


รูปที่ 2.1 ข้อมูลอนุกรมเวลาที่ลักษณะแนวโน้ม

2. ความผันแปรตามวัฏจักร (Cyclical Variation : C)

ความผันแปรตามวัฏจักร หมายถึงการเคลื่อนไหวของข้อมูลซ้ำๆ กันในระยะเวลายาวมากกว่า 1 ปี เช่น วัฏจักรทางธุรกิจที่อาจจะเป็นช่วงธุรกิจรุ่งเรือง 3 ปี เศรษฐกิจตกต่ำ 4 ปีและระยะฟื้นฟู 2 ปีเป็นเช่นนี้ไปเรื่อยๆ แต่ละรอบของวัฏจักรอาจเท่ากันหรือไม่เท่ากันก็ได้ ดังรูปที่ 2.2 แสดงกราฟของข้อมูลรายปีของจำนวนตัวลิงคซ (Lynx) ที่ถูกตัดขนในย่านแม่น้ำ mckenzie ภาค

ตะวันตกเฉียงเหนือของประเทศแคนาดา ตั้งแต่ปี ค.ศ.1821 ถึงปี ค.ศ.1934 [5] ซึ่งมีลักษณะการผันแปรตามวัฏจักร



รูปที่ 2.2 ข้อมูลอนุกรมเวลาที่ลักษณะผันแปรตามวัฏจักร

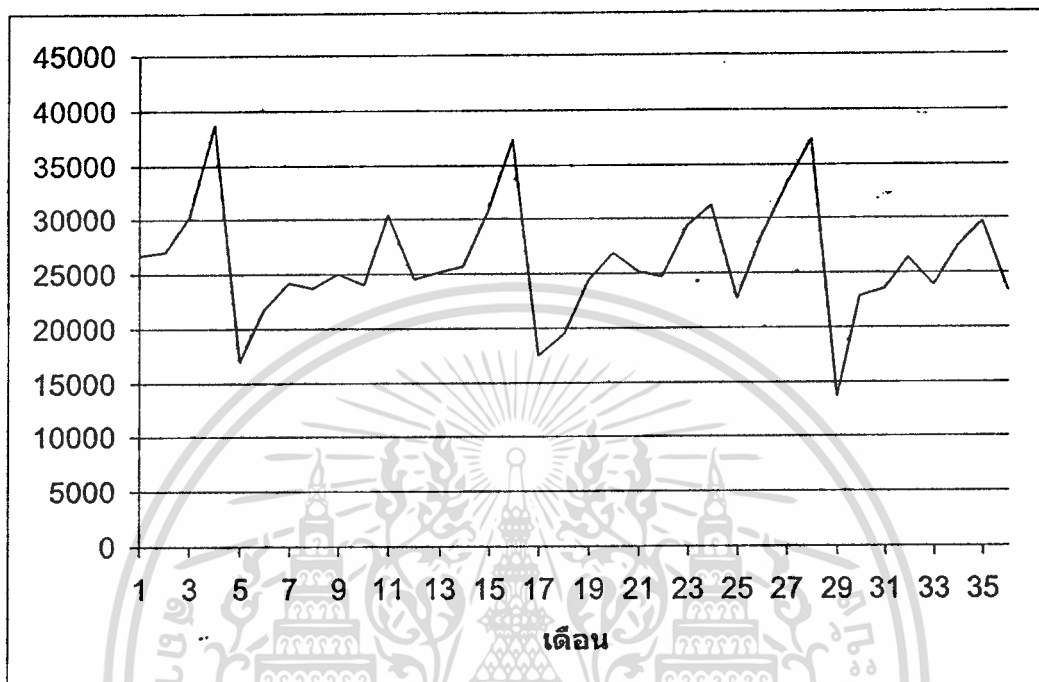
3. ความผันแปรตามฤดูกาล (Seasonal Variation : S)

ความผันแปรตามฤดูกาลหมายถึงการเปลี่ยนแปลงของข้อมูลที่เกิดขึ้นเนื่องจากอิทธิพลของฤดูกาล ซึ่งอาจเกิดขึ้นซ้ำๆ กันในช่วงเวลาเดียวกันของแต่ละปี ฤดูหนึ่งๆ จะสั้นกว่า 1 ปี เช่น อาจเป็น 3 เดือน 4 เดือนหรือ 6 เดือน เป็นต้น ตัวอย่างของความผันแปรตามฤดูกาล เช่น ยอดขายเสื้อกันหนาวจะสูงในฤดูหนาว ยอดขายการ์ด ส.ศ. จะสูงในเดือนธันวาคมและมกราคม เป็นต้น จากตัวอย่างเป็นข้อมูลรายเดือนจำนวน 36 เดือนแสดงยอดขายไวน์ของผู้ผลิตไวน์ที่ขายในประเทศออสเตรเลียตั้งแต่เดือนกันยายน ปี ค.ศ.1991 ถึงเดือนสิงหาคม ปี ค.ศ.1994 [6] จะเห็นว่าไวน์จะขายได้ดีในช่วงปลายปีคือเดือนพฤศจิกายนและเดือนธันวาคมซึ่งเป็นช่วงเทศกาลฉลองวันคริสต์มาสและวันปีใหม่ซึ่งเป็นลักษณะการผันแปรตามฤดูกาล

ข้อแตกต่างระหว่างความผันแปรตามฤดูกาลและความผันแปรตามวัฏจักร คือ

1. ช่วงเวลาของการเกิดความผันแปรตามวัฏจักร 1 รอบจะนานกว่าความผันแปรตามฤดูกาล

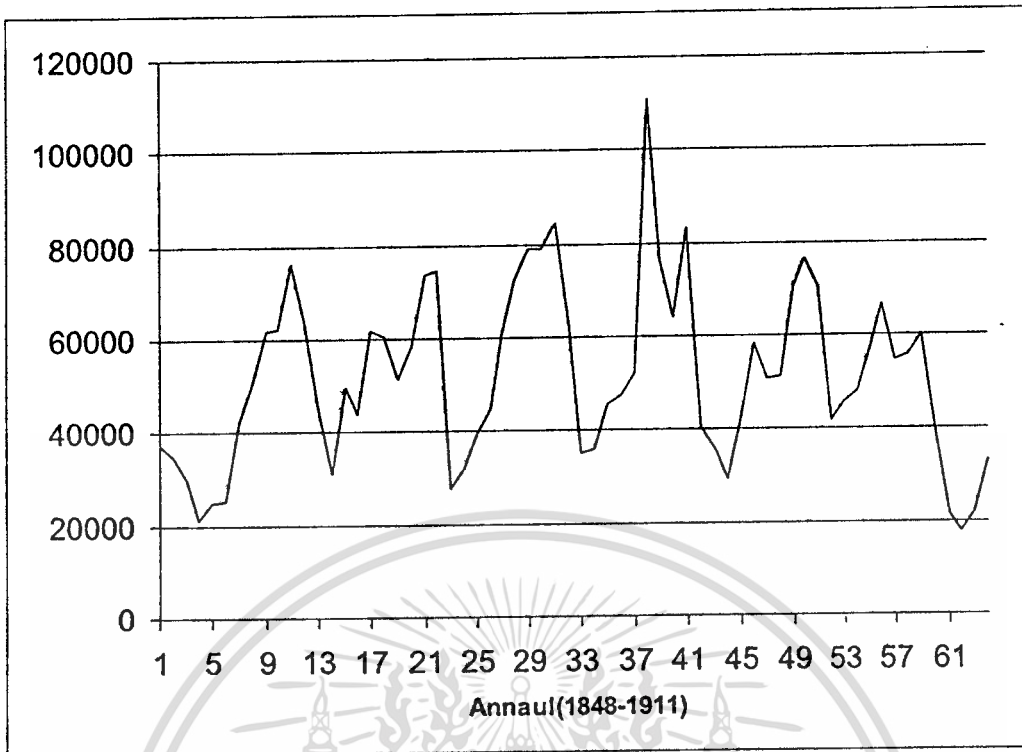
2. ช่วงเวลาการเกิดความผันแปรตามฤดูกาลจะแน่นอนกว่าช่วงการเกิดความผันแปรตามวัฏจักร เช่น ยอดขายเครื่องปรับอากาศจะขายได้ดีในช่วงเดือนมีนาคมถึงเมษายนของทุกปี เป็นต้น



รูปที่ 2.3 ข้อมูลอนุกรมเวลาที่ลักษณะผันแปรตามฤดูกาล

4. ความผันแปรที่ไม่แน่นอน (Irregular Variation : I)

ความผันแปรที่ไม่แน่นอนหรือความผันแปรอย่างสุ่ม (Random Variable) หมายถึงการเปลี่ยนแปลงของข้อมูลที่ไม่มีรูปแบบที่แน่นอน ซึ่งเกิดจากปัจจัยหรือสาเหตุที่ไม่สามารถคาดคะเนล่วงหน้า เช่น น้ำท่วม สงคราม การเลือกตั้ง เป็นต้น ซึ่งจะทำให้ตัวแปรต่างๆ ผิดปกติไป เช่น ยอดขายสินค้าฟุ่มเฟือยในช่วงเกิดสงครามจะต่ำกว่าปกติ หรือราคาสินค้าอุปโภคบริโภคในช่วงสงครามจะสูงกว่าราคาในภาวะปกติ เป็นต้น ตัวอย่างรูปที่ 2.4 เป็นข้อมูลรายปีแสดงจำนวนตัวมิงค์ (Mink) ที่ถูกตัดขนในย่านแม่น้ำ mckenzie ภาคตะวันตกเฉียงเหนือของประเทศแคนาดา ตั้งแต่ปี ค.ศ.1848 ปี ค.ศ.1911 [5] จะเห็นว่าข้อมูลมีลักษณะการเปลี่ยนแปลงที่ไม่แน่นอนซึ่งเกิดขึ้นในช่วงเดือนที่ 37 ถึง 41



รูปที่ 2.4 ข้อมูลอนุกรมเวลาที่ลักษณะไม่แน่นอน

2.4 การวัดความถูกต้องในการทำนาย (Forecasting Measure)

การวัดความถูกต้องในการทำนายทำได้โดยการคำนวณหาค่าความคลาดเคลื่อนในการทำนาย (Forecast Error) ซึ่งเป็นผลต่างของค่าเป้าหมายและค่าทำนายโดยคำนวณได้จาก [7]

$$e_t = Y_t - \hat{Y}_t \quad (2.3)$$

ค่าสถิติที่ใช้ในการวัดความคลาดเคลื่อนในการทำนายมีหลายวิธี แต่ในงานวิจัยฉบับนี้ใช้ 2 วิธีที่เป็นที่นิยมคือ

2.4.1 Mean Square Error (MSE)

MSE เป็นค่าเฉลี่ยกำลังสองของค่าความคลาดเคลื่อนในการทำนาย ค่า MSE จะวัดต่อความคลาดเคลื่อนที่มีขนาดใหญ่เพราะได้จากการนำค่าความคลาดเคลื่อนมายกกำลังสอง ซึ่งสามารถคำนวณได้ดังเงื่อนไขที่ (2.4)

$$MSE = \left(\sum_{t=1}^n e_t^2 \right) / n \quad (2.4)$$

โดยที่ n คือจำนวนข้อมูลทั้งหมด

2.4.2 Root Mean Square Error (RMSE)

RMSE เป็นรากที่สองของ MSE ซึ่งคำนวณได้ดังเงื่อนไขที่ (2.5)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2.5)$$

2.5 โครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ (Fuzzy ARTMAP Neural Network)

เนื่องจากโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพได้รับการพัฒนามาจากโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ ดังนั้น จะอธิบายสถาปัตยกรรมและการเรียนรู้ของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพก่อน แล้วจึงจะอธิบายสถาปัตยกรรมและการเรียนรู้ของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ สถาปัตยกรรมและการเรียนรู้ของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นและการทำนายข้อมูลอนุกรมเวลาโดยวิธีทางสถิติคือ Autoregressive Integrated Moving Average ตามลำดับ

โครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพถูกพัฒนาโดย Gail Carpenter, Stephen Grossberg และ David Rosen ในปี พ.ศ. 2535 โดยพัฒนามาจากฟัซซี่อาท (Fuzzy ART) และอาทแมพ (ARTMAP) รวมกัน ฟัซซี่อาทแมพถูกออกแบบให้มีการเรียนรู้แบบมีผู้สอน (Supervised) และเป็นโครงข่ายประสาทเทียมที่มีความสามารถในการจัดกลุ่มข้อมูลได้ดี โดยฟัซซี่อาทแมพจะเรียนรู้เพื่อจับกลุ่มอินพุทเพื่อให้ตอบสนองต่อเอาต์พุทที่ต้องการ

โครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพยังคงรักษาคุณสมบัติเด่นของอาท (ART : Adaptive Resonance Theory) คือสามารถเรียนรู้แพทเทินอินพุทใหม่โดยไม่ลืมแพทเทินอินพุทที่ได้เรียนรู้ไปแล้ว (Incremental Learning System) โดยการจับคู่แพทเทินขนาด M ไปยังแพทเทินขนาด N นั่นคือเมื่อมีแพทเทินอินพุทใหม่ถูกป้อนเข้าสู่ระบบ ระบบก็จะตรวจสอบว่าแพทเทินอินพุทใหม่ที่ป้อนเข้ามีความคล้ายกับกลุ่มของแพทเทินอินพุทที่ระบบได้เรียนรู้ไปแล้วหรือไม่ ถ้ามีความคล้ายกันมากพอแพทเทินอินพุทใหม่ก็จะถูกจัดกลุ่มเข้ากับกลุ่มเดิมที่ระบบได้เคยเรียนรู้ไปแล้ว แต่ถ้าแพทเทินอินพุทที่ถูกป้อนเข้าไม่มีความคล้ายกับกลุ่มใดเลยระบบก็จะสร้างกลุ่มใหม่ขึ้นมา ซึ่งโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพได้รับการยอมรับว่ามีสมรรถนะในการรู้จักแพทเทินที่ดีกว่าเทคนิคอื่น

คำว่า fuzzy หมายถึง โครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพได้รวมเอาฟัซซี่ลอจิก (Fuzzy Logic) เข้ามาด้วย คำว่า ART หมายถึง Adaptive Resonance Theory ส่วนคำว่า

MAP หมายถึงการทำงานของระบบที่เรียนรู้โดยการจับคู่แพทเทินอินพุทกับคำตอบของแพทเทินอินพุทนั้น

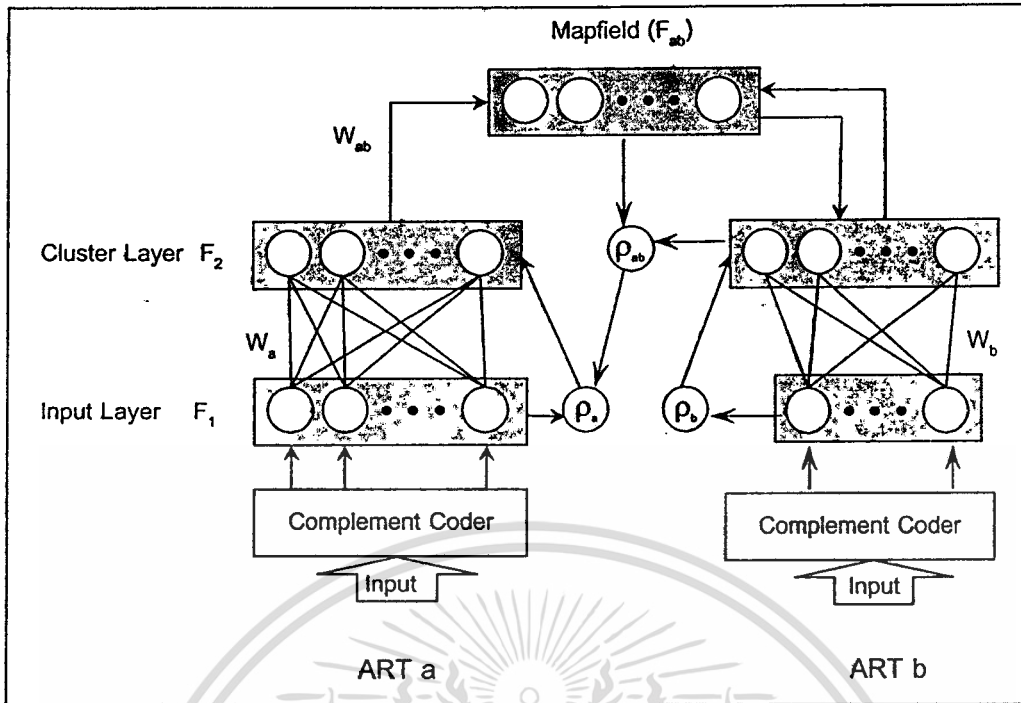
เพื่อให้เข้าใจการทำงานภายในของโครงข่ายประสาทเทียมแบบพีชชีอาทแมพจะอธิบายเกี่ยวกับพีชชีลอจิกพอเป็นพื้นฐาน นั่นคือในโลกคอมพิวเตอร์นั้นจะใช้ตัวเลข 2 ตัวคือ 0 และ 1 ซึ่งถ้าใช้ไบนารีลอจิกจะแทนได้แค่ 2 ประเภทนั่นคืออาจแทนอากาศร้อนด้วย 1 อากาศหนาวก็จะแทนได้ด้วย 0 เท่านั้น แต่ในพีชชีลอจิกจะใช้ค่าที่อยู่ระหว่าง 0 ถึง 1 ด้วย เช่น สามารถใช้ค่าที่อยู่ระหว่าง 0 ถึง 1 แทนอุณหภูมิหลายๆ แบบได้ เช่น อากาศหนาวมากแทนด้วย 0 อากาศเย็นแทนด้วย 0.2 อากาศอบอุ่นแทนด้วย 0.5 อากาศร้อนแทนด้วย 0.7 อากาศร้อนมากแทนด้วย 1 เป็นต้น จะเห็นว่าพีชชีลอจิกจะอธิบายอุณหภูมิได้ดีกว่าไบนารีลอจิกและใกล้เคียงกับโลกความเป็นจริงมากกว่า

โครงข่ายประสาทเทียมแบบพีชชีอาทแมพมีข้อดีที่น่าสนใจมาก นั่นคือมีค่าพารามิเตอร์เพียงค่าเดียวที่ต้องกำหนดคือค่าวิจิลแลนซ์ (Vigilance) ซึ่งเป็นค่าที่ใช้วัดความเหมือนของแพทเทินอินพุทในการจัดกลุ่ม และนอกจากนี้พีชชีอาทแมพจะใช้จำนวนรอบในการเรียนรู้้น้อยมากเมื่อเปรียบเทียบกับ การเรียนรู้ของโครงข่ายประสาทเทียมแบบอื่น จากคุณสมบัติเด่นดังกล่าว โครงข่ายประสาทเทียมแบบพีชชีอาทแมพจึงเป็นโครงข่ายประสาทเทียมประเภทหนึ่งที่ได้รับการสนใจพัฒนาไปประยุกต์ใช้อย่างแพร่หลาย

2.5.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบพีชชีอาทแมพ

โครงข่ายประสาทเทียมแบบพีชชีอาทแมพประกอบด้วยพีชชีอาท 2 โมดูลคือ ART_0 และ ART_1 ทั้งสองโมดูลนี้ถูกเชื่อมด้วยโมดูลกลางเรียกว่า map field (แทนด้วย F^{ab}) ทั้ง ART_0 และ ART_1 ประกอบด้วย 2 ชั้นคือชั้น F_1 และชั้น F_2 ทุกนิวรอน (Neuron) ในชั้น F_2 จะเชื่อมต่อกับทุกนิวรอนในชั้น F_1 ด้วยเวกเตอร์น้ำหนัก (Weight Vectors) แบบทั่วถึง (Fully Completed Connection) ดังรูปที่ 2.5 [8], [9]

พีชชีอาทแมพมีความสามารถในการรู้จำและจัดกลุ่มได้เป็นอย่างดี โดยสามารถเรียนรู้เพื่อจับคู่แพทเทินอินพุทกับแพทเทินเอาต์พุทที่สัมพันธ์กัน โดย ART_0 จะรับแพทเทินอินพุทที่ละแพทเทิน $\{a^{(n)}\}$ และ ART_1 รับแพทเทินเอาต์พุทที่ละแพทเทิน $\{b^{(n)}\}$ โดยที่แพทเทินเอาต์พุท $\{b^{(n)}\}$ คือเอาต์พุทเป้าหมายแพทเทินอินพุท $\{a^{(n)}\}$ และ n คือลำดับของแพทเทินอินพุทและแพทเทินเอาต์พุทที่ใช้ในการเรียนรู้ ทั้ง ART_0 และ ART_1 จะถูกเชื่อมต่อและควบคุมการเรียนรู้โดยโมดูลที่เรียกว่า map field โดย map field นี้จะทำหน้าที่สร้างความสัมพันธ์ระหว่างกลุ่มของแพทเทินอินพุทและกลุ่มของแพทเทินเอาต์พุทโดยผ่านกฎการเรียนรู้ของน้ำหนัก ถ้าแพทเทินอินพุทถูกจัดกลุ่มใน ART_0 แต่กลุ่มที่อยู่ไม่ได้สัมพันธ์กับคำตอบที่ต้องการใน ART_1 ส่วน map field ก็จะไปกระตุ้นโดยการเพิ่มค่าวิจิลแลนซ์ของ ART_0 (แทนด้วย ρ_0) เพื่อค้นหากลุ่มใหม่หรือสร้างกลุ่มใหม่



รูปที่ 2.5 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ

ให้แพทเทินอินพุทนั้นเพื่อให้ได้คำตอบที่ถูกต้องใน ART_b กระบวนการนี้เรียกว่า กฎการจับคู่ (Match Tracking Rule)

2.5.2 การเรียนรู้ของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพ

ก่อนที่จะอธิบายวิธีการเรียนรู้ของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพจะอธิบายพารามิเตอร์ที่สำคัญก่อน

2.5.2.1 แพทเทินอินพุท (Input Pattern)

แพทเทินอินพุทที่ป้อนเข้าสู่ ART_a และ ART_b จะเป็นจำนวนจริงใดๆ ที่มีค่าอยู่ระหว่าง 0 ถึง 1 แพทเทินอินพุทที่ป้อนเข้าสู่ ART_a และ ART_b จะต้องถูกทำ complement coding ดังนั้นขนาดของแพทเทินอินพุทที่ถูกป้อนเข้าจึงมีขนาดเป็น 2 เท่าของข้อมูลจริง ถ้า A คือแพทเทินอินพุทที่ป้อนเข้าสู่ ART_a

$$A = \{a_1, \dots, a_d\} \quad (2.6)$$

complement coding ของแพทเทินอินพุท A คือ

$$A^c = 1 - A \quad (2.7)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดังนั้น

$$I = (A, A^c) = \{a_1, a_2, \dots, a_d, 1-a_1, 1-a_2, \dots, 1-a_d\} \quad (2.8)$$

2.5.2.2 เวกเตอร์น้ำหนัก (Weight Vectors)

เวกเตอร์น้ำหนักแต่ละเวกเตอร์ใน ART_a และ ART_b ก็คือกลุ่ม (Category) 1 กลุ่มซึ่งแสดงด้วยโหนด (Node) 1 โหนด ส่วนเวกเตอร์น้ำหนักของ map field จะเป็นตัวเชื่อมระหว่างกลุ่มของ ART_a และกลุ่มของ ART_b เวกเตอร์น้ำหนักของ ART_a , ART_b และของ map field แทนด้วย W_a , W_b และ W_{ab} ตามลำดับ โดยเริ่มต้นค่าเวกเตอร์น้ำหนักของจะถูกกำหนดให้มีค่าเท่ากับ 1 ทั้งหมด

$$W_a = W_b = W_{ab} = 1 \quad (2.9)$$

ถ้าเวกเตอร์น้ำหนักมีค่าเท่ากับ 1 ทั้งหมดแสดงว่าเวกเตอร์น้ำหนักอยู่ในภาวะที่เรียกว่า uncommit คือยังไม่ได้เรียนรู้เพื่อที่จะเป็นตัวแทนของกลุ่มของแพทเทินอินพุทใดๆ แต่ถ้าค่าเวกเตอร์น้ำหนักเปลี่ยนไปแสดงว่าเวกเตอร์น้ำหนักนี้ได้เรียนรู้และถูกเข้ารหัสเพื่อเป็นตัวแทนของกลุ่มของแพทเทินอินพุทแล้วและจะเรียกว่าอยู่ในภาวะ commit ในระหว่างการเรียนรู้เวกเตอร์น้ำหนักจะถูกลดค่าลงเพียงอย่างเดียว

2.5.2.3 ค่าพารามิเตอร์ที่สำคัญ

1. ค่าวิจิเลนซ์ (Vigilance) คือค่าความเชื่อมั่นหรือความเหมือนในการจัดกลุ่มซึ่งจะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้าค่าวิจิเลนซ์เข้าใกล้ 1 แสดงว่าแพทเทินอินพุทที่จะถูกจัดกลุ่มเข้าด้วยกันมีความคล้ายกันมาก แต่ถ้าค่าวิจิเลนซ์มีค่าต่ำแสดงว่าแพทเทินอินพุทไม่จำเป็นต้องเหมือนกันมากก็สามารถจัดให้อยู่ในกลุ่มเดียวกันได้ ซึ่งจะทำให้ขนาดของกลุ่มมีขนาดใหญ่และมีจำนวนกลุ่มน้อย การกำหนดค่าวิจิเลนซ์ไว้สูงจะทำให้จำนวนของกลุ่มมีจำนวนมากและกลุ่มมีขนาดเล็กในพีชชี อาทแมพค่าวิจิเลนซ์ประกอบด้วยวิจิเลนซ์ของ ART_a , ART_b และ map field โดยแทนด้วย ρ_a , ρ_b และ ρ_{ab} ตามลำดับ กรณีที่ต้องการให้คำตอบตรงตามเอทพุทที่ต้องการให้กำหนด ρ_{ab} และ ρ_b ให้เท่ากับ 1 ส่วน ρ_a กำหนดตามต้องการ

2. อัตราการเรียนรู้ (Learning Rate) แทนด้วย β โดยใช้ค่าอยู่ระหว่าง 0 ถึง 1 ประกอบด้วยค่าการเรียนรู้ของ ART_a , ART_b และ map field แทนด้วย β_a , β_b และ β_{ab} ตามลำดับ สามารถกำหนดใช้อัตราการเรียนรู้ 2 ค่าได้โดยใช้อัตราการเรียนรู้เท่ากับ 1 เมื่อโหนดที่ชนะนั้นยังไม่ได้ commit มาก่อนและใช้อัตราการเรียนรู้ $0 < \beta < 1$ เมื่อโหนดที่ชนะนั้นได้ผ่านการ commit

มาแล้ว การใช้อัตราการเรียนรู้ 2 ค่านี้เรียกว่า fast-commit slow record option โดยทั่วไปแล้ว อัตราการเรียนรู้จะกำหนดเป็นค่าต่างๆ เช่น 0.05, 0.1 เป็นต้น

3. ค่าช้อยพารามิเตอร์ (Choice Parameter) แทนด้วย α ค่านี้จะกำหนดให้มากกว่า 0 เช่นอาจใช้ค่าที่มากกว่าหรือเท่ากับ 0.001 เพื่อให้มันมีผลต่อการคำนวณค่าช้อยฟังก์ชันในกรณีที่มี โหนดที่ให้ค่าช้อยฟังก์ชันเท่ากับ 1 หลายโหนด การใส่ค่าช้อยพารามิเตอร์จะทำให้ระบบจัดกลุ่มได้เหมาะสมยิ่งขึ้น

4. ค่าเอาต์พุตของ ART_a , ART_b และ map field แทนด้วย Y^a , Y^b และ X^{ab} ตามลำดับ

พีชคณิตแมทริกซ์การเรียนรู้แบบมีผู้สอน ซึ่งขั้นตอนการเรียนรู้มีดังนี้

1. เริ่มต้นกำหนดเวกเตอร์น้ำหนักของ W_a , W_b และ W_{ab} ของ ART_a , ART_b และ map field เท่ากับ 1 และขนาดของเวกเตอร์น้ำหนัก W_a เท่ากับขนาดของแพทเทินอินพุตที่ป้อนเข้าไปยัง ART_a ส่วนขนาดของเวกเตอร์น้ำหนัก W_b และ W_{ab} เท่ากับขนาดของเวกเตอร์อินพุตที่ป้อนเข้าไปยัง ART_b

2. กำหนดค่าวิจิลแลนซ์ของ ART_a เท่ากับ baseline vigilance (แทนด้วย \bar{p}) เนื่องจากค่าวิจิลแลนซ์ของ ART_a จะเพิ่มขึ้นระหว่างการเรียนรู้

3. เมื่อแพทเทินอินพุตถูกป้อนเข้าสู่ชั้น F_1 ก็จะถูกคำนวณค่าช้อยฟังก์ชัน (Choice Function) ในชั้น F_2 ของ ART_a และ ART_b ตามเงื่อนไข

$$T_j(I) = \frac{|I \wedge W_j|}{\alpha + |W_j|} \quad (2.10)$$

โดยที่ \wedge คือแอนดโอเปอร์เรเตอร์ของพีชคณิต นิยามโดย

$$(p \wedge q)_i \equiv \min(p_i, q_i) \quad (2.11)$$

$| \cdot |$ คือค่า norm ของเวกเตอร์ นิยามโดย

$$|P| \equiv \sum_{i=1}^M P_i \quad (2.12)$$

I คือแพทเทินอินพุตที่ต้องการให้ระบบเรียนรู้

W_j คือเวกเตอร์น้ำหนักของนิวรอนที่ j ในชั้น F_2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

α คือช้อยพารามิเตอร์

4. เมื่อคำนวณได้ค่าช้อยฟังก์ชันแล้วก็เลือกช้อยฟังก์ชันที่มีค่าสูงสุดตามเงื่อนไข

$$T_j = \max\{T_j(I)\} \quad (2.13)$$

โดยที่ $j = 1, 2, 3, \dots, N$

ถ้ามีนิรอนมากกว่า 1 นิรอนที่ให้ช้อยฟังก์ชันสูงสุดเท่ากันให้เลือกนิรอน j ทางซ้ายสุด ดังนั้นนิรอนในชั้น F_2 จะ commit ตามลำดับ $j = 1, 2, 3, \dots, N$ ในกรณีที่ยื่นพู่ใหม่ไม่สามารถจัดให้อยู่ในกลุ่มใดได้โครงข่ายจะสร้างกลุ่มใหม่ให้

5. ที่โมดูล ART_b เมื่อเลือกนิรอนที่ให้ค่าช้อยฟังก์ชันสูงสุดได้แล้วขั้นตอนต่อไปคือตรวจสอบค่าความเหมาะสมในการจัดกลุ่มโดยเปรียบเทียบกับค่าวิจิแลนซ์ของ ART_b (ρ_b) ตามเงื่อนไข

$$\frac{|I \wedge W_j|}{|I|} \geq \rho_b \quad (2.14)$$

ถ้าเงื่อนไขที่ (2.14) เป็นเท็จต้องกำหนดค่าช้อยฟังก์ชันของนิรอนที่ถูกเลือกขณะนั้นเป็น 0 หรือเป็นค่าต่ำๆ แล้ววนกลับไปทำที่ข้อ 4 และข้อ 5 จนกว่าเงื่อนไขที่ (2.14) จะเป็นจริงและจะได้เอาที่พู่ของ ART_b คือ Y^b

6. ที่โมดูล ART_a ก็เช่นกันตรวจสอบค่าความเหมาะสมในการจัดกลุ่มโดยเปรียบเทียบกับค่าวิจิแลนซ์ของ ART_a (ρ_a) ตามเงื่อนไขที่ (2.15)

$$\frac{|I \wedge W_j|}{|I|} \geq \rho_a \quad (2.15)$$

ถ้าเงื่อนไขที่ (2.15) เป็นเท็จต้องกำหนดค่าช้อยฟังก์ชันของนิรอนที่ถูกเลือกขณะนั้นเป็น 0 หรือเป็นค่าต่ำๆ แล้ววนกลับไปทำที่ข้อ 4 และข้อ 6 จนกว่าเงื่อนไขที่ (2.15) จะเป็นจริง จะได้เอาที่พู่ของ ART_a คือ Y^a และ J คือตำแหน่งของนิรอนที่ชนะในการจัดกลุ่ม

ถ้าเงื่อนไขที่ (2.14) และเงื่อนไขที่ (2.15) เป็นจริงเรียกว่าเกิด resonance แต่ถ้าเป็นเท็จก็จะเกิดการ reset การ reset ก็คือช้อยฟังก์ชันของนิรอน J ที่ถูกเลือกจะถูกกำหนดให้มีค่าต่ำๆ หรือกำหนดให้เป็น 0 เพื่อที่จะได้ไม่ถูกเลือกอีก

7. ตรวจสอบความถูกต้องการในการจัดกลุ่มของ ART_a และ ART_b ว่าอินพุทแพทเทินที่ ถูกจัดกลุ่มใน ART_a ตรงกับคำตอบใน ART_b หรือไม่ตามเงื่อนไข

$$\frac{|Y^b \wedge W_J^{ab}|}{|Y^b|} \geq \rho_{ab} \quad (2.16)$$

โดยที่ W_J^{ab} คือเวกเตอร์น้ำหนักของ map field ที่เชื่อมต่อกับนิรอน J ซึ่งเป็นนิรอนที่ชนะในการจัดกลุ่มใน ART_a

Y^b คือเอาท์พุทที่ได้จาก ART_b

ρ_{ab} คือค่าวิจิลแลนซ์ของ map field เป็นค่าความเชื่อมั่นที่ใช้วัดว่ากลุ่มของ ART_a ตรงกับกลุ่มของ ART_b ที่เป็นคำตอบที่ความเชื่อมั่นเท่าไร โดยทั่วไปมักจะกำหนดให้เท่ากับ 1

8. ถ้าเงื่อนไขที่ (2.16) เป็นเท็จจะกำหนดให้ค่าช้อยฟังก์ชันของโหนดที่ชนะใน ART_a เป็น 0 หรือค่าต่ำๆ เสร็จแล้วเพิ่มค่า ρ_a เพื่อให้เงื่อนไขที่ (2.17) เป็นจริงแล้วกลับไปทำข้อ 4 และข้อ 6 แต่ถ้าเงื่อนไขที่ (2.16) เป็นจริงให้ไปทำข้อ 9

$$\rho_a > \frac{|I \wedge W_J^a|}{|I|} \quad (2.17)$$

โดยที่ I คือแพทเทินอินพุทที่ป้อนเข้า ART_a

W_J^a คือเวกเตอร์ของ ART_a ที่เชื่อมระหว่างอินพุท I กับนิรอนที่ชนะ

9. คำนวณค่าเอาท์พุทของพีชชีอาทแมพตามเงื่อนไข

$$X^{ab} = \begin{cases} Y_b \wedge W_J^{ab} & \text{เมื่อ } ART_a \text{ และ } ART_b \text{ ทำงานทั้งคู่} \\ W_J^{ab} & \text{เมื่อ } ART_a \text{ ทำงานแต่ } ART_b \text{ ไม่ทำงาน} \\ Y^b & \text{เมื่อ } ART_a \text{ ไม่ทำงานแต่ } ART_b \text{ ทำงาน} \\ 0 & \text{เมื่อทั้ง } ART_a \text{ และ } ART_b \text{ ไม่ทำงาน} \end{cases} \quad (2.18)$$

10. ปรับค่าน้ำหนักของ ART_a และ ART_b ตามเงื่อนไขที่ (2.19) และปรับค่าน้ำหนักของ map field ตามเงื่อนไขที่ (2.20)

$$W_J^{new} = \beta(I \wedge W_J^{old}) + (1 - \beta)W_J^{old} \quad (2.19)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$W_{J(new)}^{ab} = \beta_{ab} X^{ab} + (1 - \beta_{ab}) W_{J(old)}^{ab} \quad (2.20)$$

11. เมื่อมีอินพุตใหม่เข้ามากลับไปทำข้อ 2

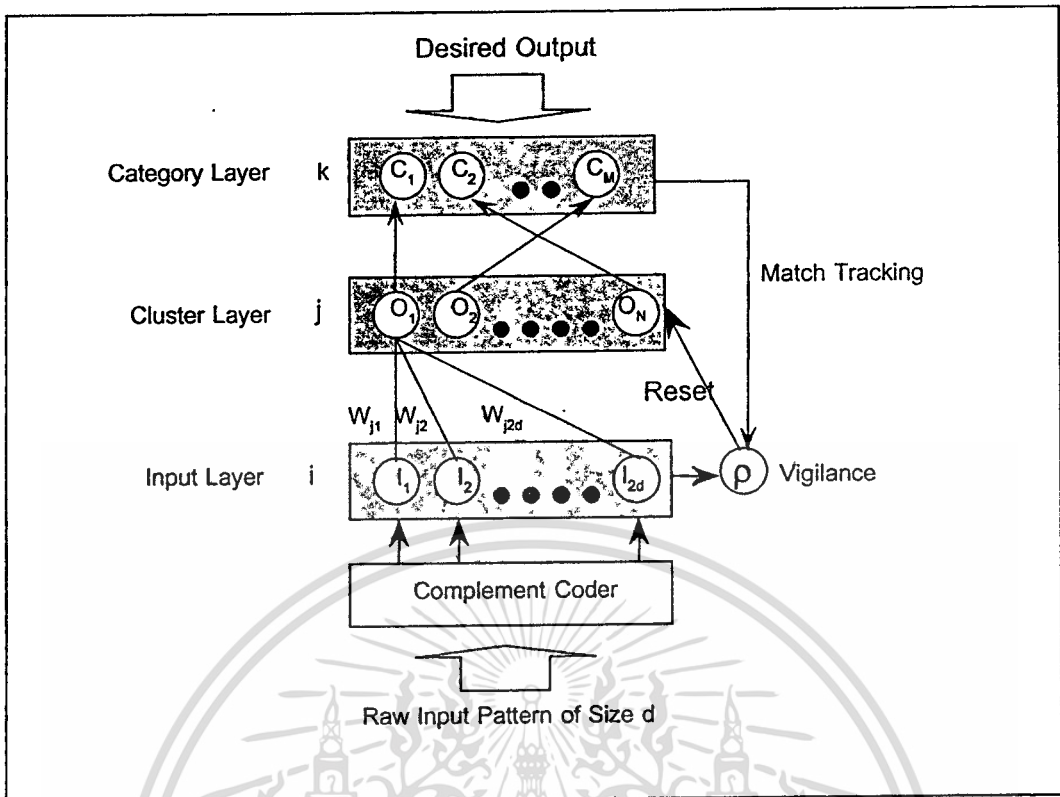
เป็นที่น่าสังเกตคือ ในระหว่างการเรียนรู้ค่าน้ำหนักของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพจะไม่เพิ่มขึ้นแต่จะลดค่าลงเพียงอย่างเดียว

2.6 โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ (Simplified Fuzzy ARTMAP)

โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพได้รับการพัฒนามาจากโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพโดย Tom Kasuba ในปี พ.ศ.2536 เพื่อลดการคำนวณที่ไม่จำเป็นลงและตัดสถาปัตยกรรมที่ซ้ำซ้อนของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพออก นั่นคือสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพจะคล้ายโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพที่ถูกตัด ART₀ ออกเหลือเพียง ART₁ และส่วน map field เท่านั้น ซึ่งทำให้โครงข่ายเรียนรู้ได้เร็วขึ้น แต่โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพยังคงรักษาข้อเด่นของโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพไว้ดังเดิมโดยเฉพาะความสามารถในเรียนรู้เพิ่มโดยไม่ลืมแพทเทินอินพุตที่ได้เรียนรู้ไปแล้ว ข้อดีอีกประการหนึ่งก็คือผู้ใช้กำหนดค่าพารามิเตอร์เพียงค่าเดียวคือค่าวิจิแลนซ์ซึ่งเป็นค่าที่ใช้วัดความเหมือนของแพทเทินอินพุต [10]

2.6.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพ

โครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซี่อาทแมพประกอบด้วย 3 ชั้นคือชั้นอินพุต (Input Layer) ชั้นคลัสเตอร์ (Cluster Layer) และชั้นแคททิกอรี (Category Layer) ชั้นอินพุตเป็นชั้นที่รับแพทเทินอินพุตที่เป็นจำนวนจริงใดๆ ที่มีค่าอยู่ระหว่าง 0 ถึง 1 เหมือนโครงข่ายประสาทเทียมแบบฟัซซี่อาทแมพและแพทเทินอินพุตนั้นต้องผ่านการทำ complement coding ซึ่งจะทำให้ขนาดของแพทเทินอินพุตเป็น 2 เท่าของขนาดแพทเทินจริง ส่วนแพทเทินเอาท์พุตไม่ต้องทำ complement coding ในชั้นคลัสเตอร์ (Cluster Layer) ประกอบไปด้วยกลุ่มหรือคลัสเตอร์ของแพทเทินอินพุตที่ถูกจัดเข้าไว้ด้วยกัน ซึ่งจำนวนโหนดในชั้นคลัสเตอร์นี้จะมีจำนวนไม่จำกัด และจะเพิ่มจำนวนขึ้นเรื่อยๆ ระหว่างการเรียนรู้ ทุกโหนดในชั้นคลัสเตอร์และทุกโหนดในชั้นอินพุตจะเชื่อมต่อกันหมดด้วยเวกเตอร์น้ำหนัก W_{jk} ซึ่งเวกเตอร์น้ำหนัก 1 เวกเตอร์ก็คือตัวแทนของคลัสเตอร์ 1 คลัสเตอร์หรือ 1 กลุ่มนั่นเอง ดังรูปที่ 2.6 โหนดในชั้นคลัสเตอร์ 1 โหนดจะชี้ไปยังโหนด 1 โหนดในชั้นแคททิกอรี (Category Layer) ซึ่งโหนดในชั้นแคททิกอรีก็คือคำตอบของแพทเทินอินพุตที่ระบบต้องเรียนรู้นั่นเอง



รูปที่ 2.6 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัซซีที่อาทแมพ

นอกจากนี้ซิมพลิไฟด์ฟัซซีที่อาทแมพยังประกอบด้วยค่าวิจิแลนซ์ (แทนด้วย ρ) ซึ่งทำหน้าที่ควบคุมกระบวนการจัดกลุ่มในชั้นคลัสเตอร์คล้ายค่าวิจิแลนซ์ของ ART₁ และ ART₂ ในฟัซซีที่อาทแมพ ค่าวิจิแลนซ์จะมีค่าอยู่ระหว่าง 0 ถึง 1 ถ้ากำหนดค่าวิจิแลนซ์ไว้สูงแสดงว่าแพทเทินอินพุตจะต้องมีความคล้ายกันมากจึงจะถูกจัดไว้ในคลัสเตอร์เดียวกันได้ ซึ่งทำให้คลัสเตอร์มีขนาดเล็กและเจาะจงและทำให้จำนวนคลัสเตอร์มีมาก ในทางตรงกันข้ามถ้ากำหนดค่าวิจิแลนซ์ไว้ต่ำแพทเทินอินพุตไม่จำเป็นต้องมีความคล้ายกันมากก็สามารถจัดไว้ในคลัสเตอร์เดียวกันได้ทำให้คลัสเตอร์มีขนาดใหญ่และจำนวนคลัสเตอร์มีน้อย เมื่อมีแพทเทินอินพุตใหม่ถูกส่งเข้าระบบค่าวิจิแลนซ์จะถูกกำหนดให้เท่ากับ base-line vigilance ($\bar{\rho}$) เนื่องจากค่าวิจิแลนซ์จะเพิ่มขึ้นระหว่างการเรียนรู้

ส่วน match tracking ทำหน้าที่ควบคุมความถูกต้องในการเรียนรู้ว่าแพทเทินอินพุตถูกจัดไว้ในคลัสเตอร์และคลัสเตอร์นั้นชี้ไปยังคำตอบในชั้นแคททิกอรีที่ตรงกับคำตอบของแพทเทินอินพุตที่กำลังเรียนรู้หรือไม่ ถ้าปรากฏว่าคลัสเตอร์นั้นชี้ไปยังคำตอบในชั้นแคททิกอรีแต่คำตอบนั้นไม่ตรงกับคำตอบของแพทเทินอินพุตแสดงว่าแพทเทินอินพุตนั้นไม่เหมาะสมที่ถูกจัดไว้ในคลัสเตอร์นั้น ถ้าเหตุการณ์นี้เกิดขึ้นระบบกำหนดให้ค่าฟังก์ชันกระตุ้นของโหนดคลัสเตอร์ที่ชนะในขณะนั้นให้เป็น 0 หรือเป็นค่าต่ำๆ และเพิ่มค่าวิจิแลนซ์ขึ้นเพื่อไม่ให้ระบบกลับมาเลือกคลัส

เดอรินี้ก็ แต่ถ้าโหนดคลาสเดอรินั้นชี้ไปยังคำตอบที่ตรงกับคำตอบของแพทเทินอินพุทระบบก็จะไปปรับเวกเดอริน้ำหนัก (W_j) นั้นแสดงว่าระบบได้เรียนรู้แพทเทินอินพุทนั้นแล้ว

2.6.2 การเรียนรู้ของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพ

แพทเทินอินพุทที่ป้อนเข้าโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพจะต้องผ่านการนอรัลไลซ์ (Normalization) โดยการทำให้ complement coding ประโยชน์ของการทำให้ complement coding คือเพื่อป้องกันไม่ให้ระบบเลือก weights เพียงอย่างเดียวเพราะหากระบบเลือก weights เพียงอย่างเดียวโหนดที่มีค่า weights ต่ำจะเป็นโหนดที่ชนะอยู่เสมอซึ่งจะทำให้การจัดกลุ่มไม่เหมาะสมได้ ผลรวมของแพทเทินอินพุทที่ผ่านการทำให้ complement coding จะเท่ากับขนาดของแพทเทินอินพุทที่ยังไม่ได้ทำ complement coding จากเงื่อนไขที่ (2.6) เงื่อนไขที่ (2.7) และเงื่อนไขที่ (2.8) จะได้เงื่อนไขที่ (2.21) ดังนี้

$$|I| = |(a, a^c)| = \sum_{i=1}^d a_i + \left(d - \sum_{i=1}^d a_i \right) = d \quad (2.21)$$

โครงข่ายประสาทเทียมซิมพลิไฟด์พีชชีอาทแมพจะเรียนรู้เพื่อจัดกลุ่มแพทเทินอินพุทตามกลุ่มคำตอบที่สัมพันธ์กับแพทเทินอินพุทนั้น อัลกอริทึมการเรียนรู้ของโครงข่ายประสาทเทียมซิมพลิไฟด์พีชชีอาทแมพมีดังนี้

1. รับแพทเทินอินพุทใหม่ที่ผ่านการทำให้ complement coding แล้ว
2. กำหนดให้ค่าวิจิลแลนซ์ (ρ) เท่ากับค่า baseline vigilance (แทนด้วย $\bar{\rho}$)
3. คำนวณค่าฟังก์ชันกระตุ้น (T_j) หรือค่าช้อยฟังก์ชันในพีชชีอาทแมพตามเงื่อนไขที่

(2.22)

$$T_j(I) = \frac{|I \wedge W_j|}{\alpha + |W_j|} \quad (2.22)$$

ค่า α จะใช้ค่าต่างๆ เข้าใกล้ 0 เช่น 0.0001 เป็นต้น

4. เลือกโหนดหรือนิวรอนที่ให้ค่าช้อยฟังก์ชันสูงสุดตามเงื่อนไข

$$winner = \max\{T_j(I)\} \quad \text{โดยที่ } j = 1, 2, 3, \dots, N \quad (2.23)$$

ถ้ามีนิรอนมากกว่า 1 นิรอนที่ให้ค่าฟังก์ชันกระตุ้นสูงสุดเท่ากันให้เลือกนิรอน j ทางซ้ายสุด ดังนั้นนิรอนในชั้นคัลล์เตอร์จะ commit ตามลำดับ $j = 1, 2, 3, \dots, N$

5. คำนวณหาค่าแมทฟังก์ชันและตรวจสอบความเหมาะสมในการจัดกลุ่มโดยเปรียบเทียบกับค่าวิจิลแลนซ์ตามเงื่อนไข

$$\frac{|I \wedge W_j|}{d} \geq \rho \quad (2.24)$$

โดยที่ J คือตำแหน่งของนิรอนที่ชนะในการจัดกลุ่ม

ถ้าเงื่อนไขที่ (2.24) เป็นจริงแสดงว่าคัลล์เตอร์ J ที่ถูกเลือกเหมาะสมที่จะรวมแพทเทินอินพุท (I) เข้าไว้ในกลุ่ม แต่ถ้าเงื่อนไขที่ (2.24) เป็นเท็จต้องกำหนด (Reset) ให้ค่าฟังก์ชันกระตุ้นของโหนดคัลล์เตอร์ที่ถูกเลือกตอนนั้นเป็นค่าต่ำๆ หรือเป็น 0 เพื่อรอบต่อไปจะได้ไม่ถูกเลือกอีกและกลับไปเลือกโหนดคัลล์เตอร์ใหม่ตามเงื่อนไขที่ (2.23) จนทำจนกว่าเงื่อนไขที่ (2.24) จะเป็นจริงแล้วจึงไปทำข้อ 6 ต่อไป

6. เมื่อเลือกโหนดคัลล์เตอร์ที่ชนะได้แล้วขั้นต่อไปต้องไปตรวจสอบว่าโหนดคัลล์เตอร์ J ที่ชนะนั้นเข้าไปถึงคำตอบหรือโหนดในชั้นแคททิกริที่ตรงกับคำตอบ (O) ของแพทเทินอินพุทที่กำลังเรียนรู้ในปัจจุบันหรือไม่ ตามเงื่อนไขที่ (2.25)

$$Category_{(J)} = O \quad (2.25)$$

โดยที่ $Category_{(J)}$ คือแคททิกริที่โหนดคัลล์เตอร์ที่ชนะชื่ออยู่

O คือคำตอบของแพทเทินอินพุทที่ระบบกำลังเรียนรู้อยู่

ถ้าเงื่อนไขที่ (2.25) เป็นเท็จนั่นคือคำตอบของโหนดคัลล์เตอร์ชนะหรือที่ถูกเลือกนั้นไม่ตรงกับคำตอบของแพทเทินอินพุทที่กำลังเรียนรู้ก็กำหนดให้ค่าฟังก์ชันกระตุ้นที่โหนดคัลล์เตอร์ที่ถูกเลือกตอนนั้นเป็นค่าต่ำๆ หรือเป็น 0 และเพิ่มค่าวิจิลแลนซ์ขึ้นอีกเล็กน้อยพอที่จะทำให้เงื่อนไขที่ (2.26) เป็นจริง

$$\rho > \frac{|I \wedge W_j|}{|I|} \quad (2.26)$$

จากนั้นจึงวนกลับไปเลือกโหนดคลัสเตอร์ใหม่ตามเงื่อนไขที่ (2.23) และตรวจสอบเงื่อนไขที่ (2.24) และเงื่อนไขที่ (2.25) จนทำจนกว่าเงื่อนไขที่ (2.24) และเงื่อนไขที่ (2.25) จะเป็นจริงหรือโหนดคลัสเตอร์ที่ถูกเลือกนั้นชี้ไปยังโหนดในชั้นแคตทิทอรีที่ตรงกับคำตอบของแพทเทินอินพุทที่กำลังเรียนรู้อยู่แล้วจึงไปทำข้อ 7

7. ปรับค่านำหนักของโหนดคลัสเตอร์ที่ชนะตามเงื่อนไขที่ (2.27)

$$W_j^{new} = \beta(I \wedge W_j^{old}) + (1 - \beta)W_j^{old} \quad (2.27)$$

เมื่อเรียนรู้แพทเทินอินพุทนี้เสร็จแล้วให้วนกลับไปกำหนดค่าวิจแลนซ์ให้เท่ากับค่า base-line vigilance ก่อนรับแพทเทินอินพุทใหม่มาเรียนรู้ต่อไป

2.6.3 การทำนายของซิมพลิไฟด์ฟัชชีอาทแมพ (Simplified Fuzzy ARTMAP Classification)

การทำนายของโครงข่ายประสาทเทียมแบบฟัชชีอาทแมพและโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัชชีอาทแมพจะคล้ายกัน กล่าวคือการทำนายของโครงข่ายประสาทเทียมแบบฟัชชีอาทแมพจะใช้แค่ ART₀ และส่วน mapfiled เท่านั้น ส่วน ART₀ ไม่ได้ถูกใช้งาน นั่นคือจะใช้เวกเตอร์นำหนักของ ART₀ (W_a) และเวกเตอร์นำหนักของ map filed (W_{ab}) ที่ได้จากขั้นตอนการเรียนรู้ แพทเทินอินพุทที่ใช้ในการทดสอบจะถูกป้อนเข้า ART₀ และคำนวณค่าช้อยฟังก์ชันจากนั้นก็เลือกโหนดที่ให้ค่าช้อยฟังก์ชันสูงสุด ถ้ามีโหนดที่ให้ค่าช้อยฟังก์ชันสูงสุดเท่ากันมากกว่า 1 โหนด จะเลือกโหนดที่อยู่ทางซ้ายมือสุดเหมือนช่วงการเรียนรู้ ซึ่งจะได้โหนด J ที่ชนะซึ่งจะมีเพียงโหนดเดียวเท่านั้น จากนั้นจึงดูว่าโหนดที่ชนะนั้นเชื่อมกับเวกเตอร์นำหนัก W_{ab} อะไรซึ่งก็จะให้คำตอบออกมา

ส่วนการทำนายในโครงข่ายประสาทเทียมแบบซิมพลิไฟด์ฟัชชีอาทแมพเมื่อแพทเทินอินพุทถูกส่งเข้ามาก็จะถูกคำนวณเพื่อหาค่าฟังก์ชันกระตุ้น จากนั้นก็เลือกโหนดคลัสเตอร์ที่ให้ค่าฟังก์ชันกระตุ้นสูงสุด โดยถ้ามีโหนดคลัสเตอร์ที่ให้ค่าฟังก์ชันกระตุ้นสูงสุดเท่ากันก็จะเลือกโหนดคลัสเตอร์อยู่ทางซ้ายมือสุดเหมือนช่วงการเรียนรู้ เมื่อได้โหนดคลัสเตอร์ที่ชนะก็จะดูว่าโหนดคลัสเตอร์ที่ชนะนั้นชี้ไปยังแคตทิทอรีโหนดใดซึ่งโหนดนั้นก็คือคำตอบนั่นเอง

2.7 โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นเป็นโครงข่ายประสาทเทียมที่ถูกนำไปใช้ในประโยชน์ในงานหลายด้านเช่น การจัดกลุ่มข้อมูล การรู้จำแพทเทิน การทำนายข้อมูลอนุกรมเวลา เป็นต้น โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นเป็นโครงข่ายประสาทเทียมแบบ

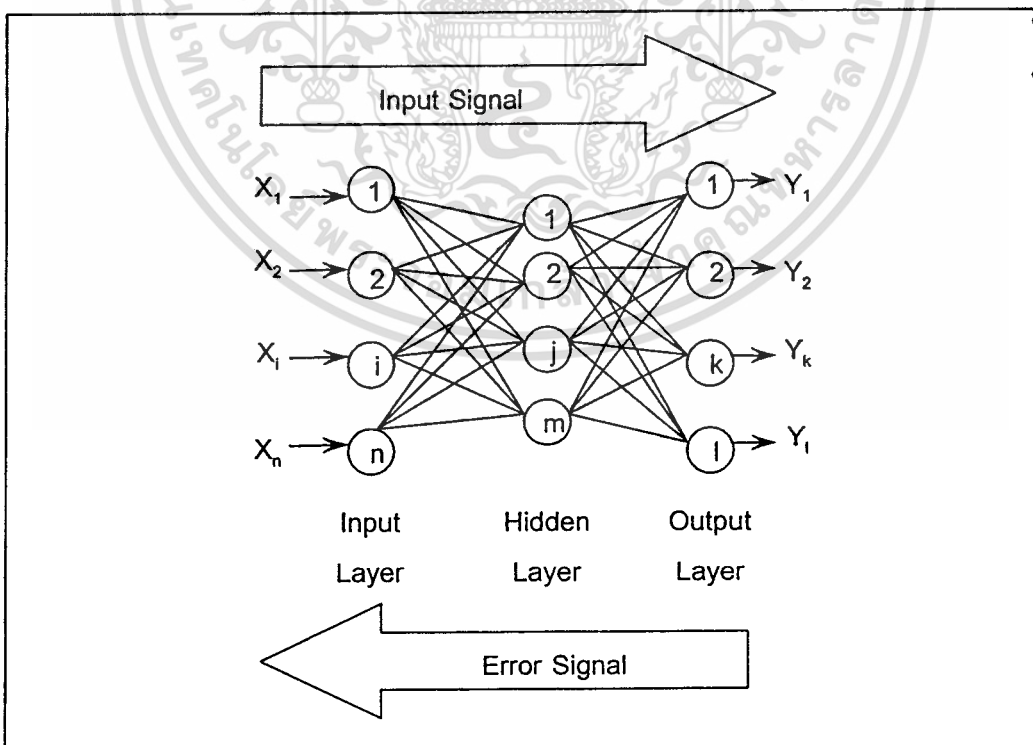
feedforward ที่มีการเรียนรู้แบบมีผู้สอน feedforward คือทิศทางการไหลของข้อมูลจะไหลจากชั้นอินพุตผ่านชั้นฮิดเดนไปยังชั้นเอาต์พุต ส่วนการเรียนรู้แบบมีผู้สอนเป็นกระบวนการเปรียบเทียบกับคำตอบที่ได้จากโครงข่ายกับคำตอบที่ถูกต้องแล้วปรับค่าน้ำหนักบนพื้นฐานของค่าความผิดพลาด (Error) ที่เกิดขึ้น เพื่อให้โครงข่ายให้คำตอบที่ใกล้เคียงกับคำตอบจริงมากที่สุด

โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันได้รับการพัฒนาเพื่อแก้ปัญหาการเรียนรู้แบบ perceptron โดยเพิ่มชั้นฮิดเดนเข้าไปในโครงข่ายเพื่อให้สามารถแก้ปัญหาที่ไม่เป็นเชิงเส้นได้ การเพิ่มชั้นฮิดเดนเข้าไปแล้วทำให้สามารถแก้ปัญหาของ perceptron ได้เนื่องจากนิวรอนในชั้นฮิดเดนจะทำหน้าที่เหมือนเป็นตัวจับลักษณะเด่น (Features) ของแพทเทินอินพุตโดยเข้ารหัสไว้ในค่าน้ำหนักซึ่งลักษณะเด่นจะถูกใช้โดยนิวรอนในชั้นเอาต์พุตเพื่อให้คำตอบออกมา

โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน เป็นการเรียนรู้ที่พยายามปรับค่าความผิดพลาดให้เหลือน้อยที่สุดโดยการเปรียบเทียบคำตอบที่ได้จากโครงข่ายกับคำตอบที่ต้องการแล้วทำการปรับค่าน้ำหนักเพื่อให้โครงข่ายให้คำตอบที่ใกล้เคียงกับคำตอบที่ต้องการมากที่สุด

2.7.1 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันประกอบด้วยนิวรอนหรือโหนด (Node) แทนด้วยวงกลม นิวรอนจะรวมกันเป็นชั้น (Layer) โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันประกอบด้วย 3 ชั้น คือชั้นอินพุต (Input Layer) ชั้นฮิดเดน (Hidden Layer) และชั้นเอาต์พุต



รูปที่ 2.7 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(Output Layer) นิวรอนในทุกชั้นจะเชื่อมต่อแบบถึงกันหมดโดยเส้นที่เชื่อมต่อกันจะมีน้ำหนัก (Weights) กำกับอยู่ด้วยเพื่อบอกถึงความสำคัญของเส้นที่เชื่อมต่อนั้น ดังรูปที่ 2.7 โครงข่ายประสาทเทียมแบบแบบคพรอพพาเกชันเป็น mapping network นั่นคือโครงข่ายสามารถคำนวณฟังก์ชันในบางลักษณะที่กำเนิดความสัมพันธ์ระหว่างอินพุตและเอาต์พุตโดยอาศัยการเรียนรู้แบบแพร่ย้อนกลับ ลักษณะการเรียนรู้ของโครงข่ายประสาทเทียมแบบแบบคพรอพพาเกชันจะกระทำ 2 ขั้นตอนคือขั้นแรกแพทเทินอินพุตถูกส่งเข้าไปยังชั้นอินพุตซึ่งในขั้นนี้จะไม่มีการคำนวณใดๆ ชั้นฮิดเดนเป็นชั้นที่รับข้อมูลจากชั้นอินพุตมาคำนวณจากนั้นก็ส่งค่าที่ได้จากการคำนวณไปยังชั้นเอาต์พุตจนได้คำตอบออกมา ขั้นตอนที่ 2 เมื่อได้คำตอบออกมาก็จะเปรียบเทียบกับคำตอบที่ได้กับคำตอบจริง ถ้าคำตอบไม่ถูกต้องก็จะคำนวณค่าความผิดพลาดและแพร่ย้อนกลับค่าความผิดพลาดที่เกิดขึ้นไปยังชั้นเอาต์พุตและชั้นฮิดเดนโดยการปรับค่าน้ำหนักในชั้นฮิดเดนและชั้นเอาต์พุตเพื่อให้โครงข่ายให้คำตอบที่ใกล้เคียงกับคำตอบจริงมากที่สุด

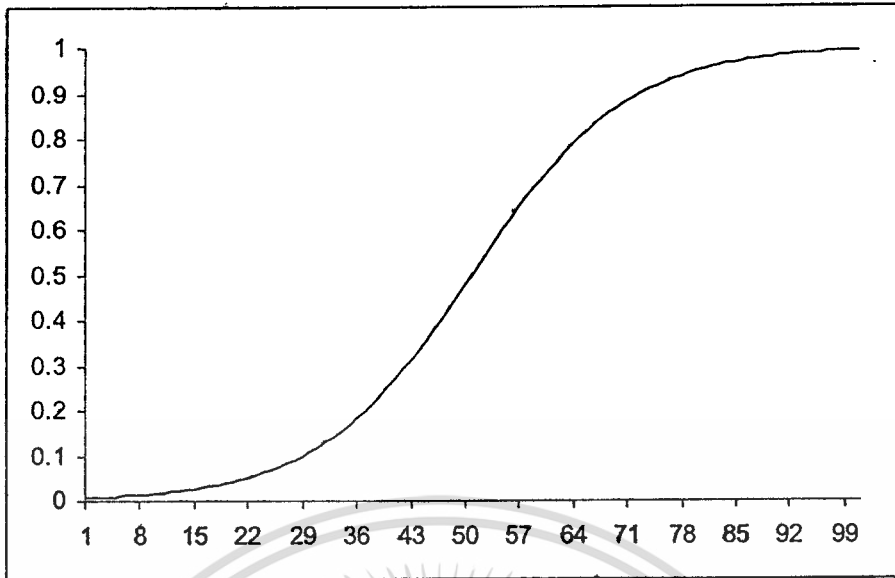
ฟังก์ชันกระตุ้น (Transfer Function) ของนิวรอนในชั้นฮิดเดนและชั้นเอาต์พุตของโครงข่ายประสาทเทียมแบบแบบคพรอพพาเกชันใช้ฟังก์ชัน sigmoid ดังเงื่อนไขที่ (2.26) ซึ่งมีรูปแบบเป็นรูปตัว s ดังรูปที่ 2.8 โดยค่าของเอาต์พุตที่ได้จากฟังก์ชัน sigmoid จะอยู่ระหว่าง 0 ถึง 1

$$Y = \frac{1}{1 + e^{-x}} \quad (2.26)$$

โดยที่ Y คือค่าที่ผ่านฟังก์ชัน sigmoid
 x คือข้อมูลอินพุต

เมื่อเริ่มต้นการเรียนรู้ค่าน้ำหนักของโครงข่ายได้จากการสุ่มค่าซึ่งมักจะเป็นค่าต่ำๆ เช่น อาจสุ่มค่าที่อยู่ระหว่าง -1 ถึง 1 หรืออาจสุ่มค่าที่อยู่ระหว่าง 0 ถึง 1 เป็นต้น แต่หลังจากโครงข่ายได้เรียนรู้ไปแล้วค่าน้ำหนักอาจจะเป็นค่าใดๆ ก็ได้ซึ่งอาจมากกว่าหรือน้อยกว่าค่าที่ได้จากการสุ่มครั้งแรก จำนวนชั้นฮิดเดนอาจใช้ 1 ชั้นหรือมากกว่า 1 ชั้นก็ได้ ถ้าใช้จำนวนชั้นฮิดเดนมากการคำนวณก็จะเพิ่มมากตามไปด้วย

ข้อมูลจะต้องถูกนอร์มัลไลซ์ให้อยู่ในช่วง 0 ถึง 1 ก่อนจะถูกป้อนเข้าสู่ชั้นอินพุต จำนวนของนิวรอนในชั้นอินพุตจะเท่ากับจำนวนค่าที่อยู่ในเวกเตอร์อินพุต เช่น ถ้าแพทเทินอินพุต 1 แพทเทินประกอบด้วย $0.5 \ 0.6 \ 0.4 \ 0.5$ จำนวนนิวรอนในชั้นอินพุตก็จะเท่ากับ 4 นิวรอน เป็นต้น จำนวนนิวรอนในชั้นเอาต์พุตเท่ากับจำนวนคำตอบที่ต้องการ ส่วนการกำหนดจำนวนนิวรอนในชั้นฮิดเดนไม่มีหลักเกณฑ์ที่แน่นอนต้องลองผิดลองถูก



รูปที่ 2.8 ฟังก์ชัน sigmoid

2.7.2 การเรียนรู้ของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

พิจารณาสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน ตามรูปที่ 2.7 ประกอบด้วย 3 ชั้นคือชั้นอินพุท ชั้นฮิดเดนและชั้นเอาต์พุทโดยที่ i คือนิวรอนในชั้นอินพุท j คือนิวรอนในชั้นฮิดเดนและ k คือนิวรอนในชั้นเอาต์พุท โดยมีอินพุท $X_1, X_2, \dots, X_i, X_n$ แพร่ผ่านเครือข่ายจากซ้ายไปขวาและค่าความผิดพลาด e_1, e_2, \dots, e_i แพร่จากขวาไปซ้าย W_{ji} แทนค่าน้ำหนักจากนิวรอน i ในชั้นอินพุทไปนิวรอน j ในชั้นฮิดเดนและ W_{kj} แทนค่าน้ำหนักจากนิวรอน j ในชั้นฮิดเดนไปยังนิวรอน k ในชั้นเอาต์พุท

โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันมีขั้นตอนการเรียนรู้ ดังนี้ [11]

1. กำหนดค่าน้ำหนัก (Weight Vector) โดยการสุ่มค่าที่อยู่ในช่วงแคบๆ เช่น -1 ถึง 1
2. กำหนดแพทเทินอินพุทเช่น $X_1(p), X_2(p), \dots, X_n(p)$ และเอาต์พุทที่สัมพันธ์กับแพทเทินอินพุทคือ $Y_{d,1}(p), Y_{d,2}(p), \dots, Y_{d,n}(p)$ เมื่อ p คือรอบการเรียนรู้
3. คำนวณผลลัพธ์ในชั้นต่างๆ ดังนี้

3.1 คำนวณผลลัพธ์ในชั้นฮิดเดนตามเงื่อนไขที่ (2.27)

$$Y_j(p) = \text{sigmoid} \left[\sum_{i=1}^n X_i(p) \times W_{ji}(p) \right] \quad (2.27)$$

โดยที่ $Y_j(p)$ คือค่าเอาต์พุทที่ออกจากนิวรอน j ของรอบที่ p

sigmoid คือฟังก์ชันกระตุ้น sigmoid

X_i คืออินพุต

W_{ji} คือค่าน้ำหนักจากนิวรอน i ไปยังนิวรอน j

n คือจำนวนอินพุตทั้งหมดที่ส่งเข้านิวรอน j ในชั้นฮิดเดน

3.2 คำนวณผลลัพท์ในชั้นเอาต์พุตตามเงื่อนไขที่ (2.28)

$$Y_k(p) = \text{sigmoid} \left[\sum_{i=1}^m X_j(p) \times W_{kj}(p) \right] \quad (2.28)$$

โดยที่ $Y_k(p)$ คือค่าเอาต์พุตที่ออกจากนิวรอน k ของรอบที่ p

sigmoid คือฟังก์ชันกระตุ้น sigmoid

X_j คืออินพุต

W_{kj} คือค่าน้ำหนักจากนิวรอน j ไปยังนิวรอน k

m คือจำนวนอินพุตทั้งหมดที่ส่งเข้านิวรอน k ในชั้นเอาต์พุต

4. ปรับค่าน้ำหนักโดยการแพร่ย้อนกลับค่าความผิดพลาด

4.1 คำนวณค่า error gradient สำหรับนิวรอนในชั้นเอาต์พุตตามเงื่อนไขที่

(2.29)

$$\delta_k(p) = Y_k(p) \times [1 - Y_k(p)] \times e_k(p) \quad (2.29)$$

โดยที่ $Y_k(p)$ คือค่าเอาต์พุตจากนิวรอน k ในชั้นเอาต์พุต

$e_k(p)$ คือค่าความผิดพลาดที่ได้จากการนำค่าเอาต์พุตที่ต้องการลบ
คำตอบที่ได้จากระบบคำนวณได้ตามเงื่อนไขที่ (2.30)

$$e_k(p) = Y_{d,k}(p) - Y_k(p) \quad (2.30)$$

โดยที่ $Y_{d,k}(p)$ คือคำตอบหรือเอาต์พุตที่ต้องการที่รอบ p

$Y_k(p)$ คือคำตอบหรือเอาต์พุตที่ออกจากนิวรอน k ที่รอบ p

4.2 คำนวณการเปลี่ยนแปลงค่าน้ำหนักตามเงื่อนไขที่ (2.31)

$$\Delta W_{kj}(p) = \alpha \times Y_j(p) \times \delta_k(p) \quad (2.31)$$

โดยที่ α คืออัตราการเรียนรู้ (Learning Rate)

$Y_j(p)$ คือเอาต์พุตที่ออกจากนิวรอน j ในชั้นฮิดเดนที่รอบ p

$\delta_k(p)$ คือค่า error gradient ของนิวรอน k ในชั้นเอาต์พุตที่รอบ p

4.3 ปรับค่าน้ำหนักของนิวรอนในชั้นเอาต์พุตตามเงื่อนไข (2.32)

$$W_{kj}(p+1) = W_{kj}(p) + \Delta W_{kj}(p) \quad (2.32)$$

4.4 คำนวณค่า error gradient สำหรับนิวรอนในชั้นฮิดเดนตามเงื่อนไข

(2.33)

$$\delta_j(p) = Y_j(p) \times [1 - Y_j(p)] \times \sum_{k=1}^l \delta_k(p) \times W_{kj}(p) \quad (2.33)$$

โดยที่ $Y_j(p)$ คือค่าเอาต์พุตที่ออกจากนิวรอน j รอบที่ p

$\delta_k(p)$ คือค่า error gradient ของนิวรอน k ในชั้นเอาต์พุตรอบที่ p

W_{kj} คือค่าน้ำหนักจากนิวรอน j ไปนิวรอน k รอบที่ p

4.5 คำนวณการเปลี่ยนแปลงค่าน้ำหนักตามเงื่อนไข (2.34)

$$\Delta W_{ji}(p) = \alpha \times X_i(p) \times \delta_j(p) \quad (2.34)$$

โดยที่ $X_i(p)$ คืออินพุต X_i ในชั้นอินพุตของรอบที่ p

$\delta_j(p)$ คือค่า error gradient ของนิวรอน j ในชั้นฮิดเดนรอบที่ p

4.6 ปรับค่าน้ำหนักของนิวรอนในชั้นฮิดเดนตามเงื่อนไข (2.35)

$$W_{ji}(p+1) = W_{ji}(p) + \Delta W_{ji}(p) \quad (2.35)$$

5. เพิ่มจำนวนรอบหรือเพิ่มค่า p อีก 1 และวนกลับไปทำขั้นตอนที่ 2

อัตราการเรียนรู้ (Learning Rate) สำหรับโครงข่ายประสาทเทียมแบบแบคพรอพพา เกชันใช้ค่าระหว่าง 0 ถึง 1 ซึ่งอัตราการเรียนรู้จะเป็นตัวบอกอัลกอริทึมว่าจะต้องปรับค่าน้ำหนัก มากน้อยเท่าไร ถ้ากำหนดอัตราการเรียนรู้ต่ำจะทำให้ค่าน้ำหนักค่อยๆ เปลี่ยนแปลงและใช้เวลา เรียนรู้นาน ถ้ากำหนดอัตราการเรียนรู้สูงจะทำให้ค่าน้ำหนักมีการเปลี่ยนแปลงเร็วและทำให้การ

เรียนรู้เร็วขึ้นแต่ถ้ากำหนดอัตราการเรียนรู้สูงมากเกินไปอัลกอริทึมอาจข้ามค่าน้ำหนักที่เหมาะสมไป โดยทั่วไปอัตราการเรียนรู้จะใช้ค่าต่ำๆ เช่น 0.05-0.1 เป็นต้น

2.8 Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) เป็นวิธีการทำนายทางสถิติที่พัฒนาโดย George Box และ Gwilym Jenkins และเผยแพร่ในช่วงต้นของคริสต์ทศวรรษ 1970 ในหนังสือชื่อ Time Series Analysis : Forecasting and Control ARIMA เป็นวิธีที่ค่อนข้างยุ่งยากในการวิเคราะห์และใช้เวลาในการคำนวณมาก อย่างไรก็ตามในปัจจุบันได้มีการพัฒนาโปรแกรมสำเร็จรูปขึ้นเพื่อช่วยในการคำนวณ เช่น Minitab SPSS เป็นต้น ทำให้ใช้งานได้ง่ายขึ้น

การทำนายกระทำได้โดยการกำหนดรูปแบบ (Model) ให้กับอนุกรมเวลาในรูปแบบของ ARIMA(p,d,q) (Autoregressive Integrated Moving Average order p and q Model) ซึ่งเป็นการรวมส่วนของรูปแบบ AR(p) และรูปแบบ IMA(q) เข้าด้วยกัน รูปแบบ AR(p) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t ขึ้นอยู่กับค่าสังเกต Y_{t-1}, \dots, Y_{t-p} หรือค่าสังเกตที่เกิดขึ้นก่อนหน้า p ค่า ส่วนรูปแบบ IMA(q) หมายถึงรูปแบบที่แสดงว่าค่าสังเกต Y_t ขึ้นอยู่กับค่าความคลาดเคลื่อน $\varepsilon_{t-1}, \dots, \varepsilon_{t-q}$ หรือค่าความคลาดเคลื่อนที่อยู่ก่อนหน้า q ค่า ส่วน d คือจำนวนครั้งในการทำ first differencing เพื่อให้ข้อมูลอยู่ในรูปสเตชันนารี (Stationary)

ในกรณีที่มีอิทธิพลฤดูกาลในอนุกรมเวลาจะกำหนดรูปแบบ SARIMA(P,D,Q) (Seasonal Autoregressive Integrated Moving Average order P,D,Q) ให้กับอนุกรมเวลาด้วย โดยที่ P คืออันดับของ SARI และ Q คืออันดับของ SIMA และ D คือจำนวนครั้งในการหาผลต่างฤดูกาลเพื่อทำให้ข้อมูลสเตชันนารี การกำหนดรูปแบบจะกำหนดให้กับข้อมูลที่เป็นสเตชันนารี เท่านั้น

2.8.1 เครื่องหมาย B (Backshift Notation)

เพื่อให้การเขียนรูปแบบง่ายขึ้นจะเขียนในรูปแบบเครื่องหมาย B (Backshift Notation) โดย B คือการเลื่อนย้อนกลับ 1 ตำแหน่ง โดยที่

$$BY_t = Y_{t-1} \quad (2.36)$$

ถ้าย้อนกลับ 2 ตำแหน่งเขียนได้ดังนี้

$$B(BY_t) = B^2Y_t = Y_{t-2} \quad (2.37)$$

ถ้าเป็นข้อมูลรายเดือนและต้องการย้อนกลับ 12 เดือนเขียนได้ดังนี้ คือ $B^{12}Y_t = Y_{t-12}$ เครื่องหมาย B จะถูกนำไปใช้ในการอธิบายการทำ first differencing ซึ่งสามารถเขียนได้ดังนี้

$$Y'_t = Y_t - Y_{t-1} = Y_t - BY_t = (1-B)Y_t \quad (2.37)$$

ดังนั้น first differencing จึงสามารถแทนด้วย $(1-B)$ ถ้าต้องการเขียน second order differencing หรือ first differencing of first differencing สามารถเขียนได้ดังนี้

$$\begin{aligned} Y''_t &= (Y'_t - Y'_{t-1}) \\ &= (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \\ &= (1 - 2B + B^2)Y_t \\ &= (1-B)^2 Y_t \end{aligned} \quad (2.38)$$

ดังนั้น second order differencing จึงแทนด้วย $(1-B)^2$ แต่ second differencing แทนด้วย $1-B^2$ ดังนั้น d-order differencing เขียนได้ว่า $(1-B)^d Y_t$ และ ถ้าข้อมูลผ่านการทำ first differencing และ seasonal differencing สามารถเขียนได้ว่า

$$(1-B)(1-B^s)Y_t \quad (2.39)$$

จากสมการที่ (2.39) สามารถเขียนในรูปของสมการเพื่อนำไปใช้ในการทำนายได้ ดังนี้

$$(1-B)(1-B^s)Y_t = (1-B-B^s+B^{s+1})Y_t = Y_t - Y_{t-1} - Y_{t-s} + Y_{t-s-1} \quad (2.40)$$

2.8.2 ขั้นตอนการทำนายโดยวิธี ARIMA

1. กำหนดรูปแบบ (Identification) ประกอบด้วย 2 ขั้นตอนคือ

1.1 แปลงอนุกรมเวลาให้เป็นสเตชันนารี

1.2 เลือกโมเดลที่เหมาะสมให้กับอนุกรมเวลา โดยการพิจารณากราฟค่า

สัมประสิทธิ์สหสัมพันธ์แบบออโต (Autocorrelation Function : ACF) และค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบออโต (Partial Autocorrelation Function : PACF) ของอนุกรมเวลา

2. ประมาณและทดสอบ (Estimation and Testing)

2.1 ประมาณพารามิเตอร์ในรูปแบบ (Estimation) ซึ่งทำได้ 2 วิธีคือการหาค่าประมาณแบบง่ายและค่าประมาณที่ได้จากการวิเคราะห์ตัวเลข (Numerical Analysis) สำหรับค่าประมาณแบบง่ายจะทำโดยการสร้างสมการที่มาจากความสัมพันธ์ระหว่างค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและพารามิเตอร์ โดยสมการที่สร้างขึ้นจะเท่ากับจำนวนพารามิเตอร์ที่ต้องการประมาณ ส่วนค่าประมาณที่ได้จากการวิเคราะห์ตัวเลขจะได้รับการแก้สมการที่สร้างขึ้นด้วยวิธีกำลังสองน้อยที่สุด ขั้นตอนของการวิเคราะห์ตัวเลขจะต้องกำหนดค่าประมาณเริ่มต้นก่อน เช่น โปรแกรม Minitab จะใช้ค่าประมาณเริ่มต้นที่ 0.1 จากนั้นจึงใช้เทคนิคการค้นหาลายๆ รอบจนได้ค่าพารามิเตอร์สุดท้ายที่ให้ค่ากำลังสองน้อยที่สุด เมื่อสิ้นสุดการวิเคราะห์จะได้ค่าประมาณสุดท้ายที่จะนำไปสร้างสมการเพื่อทำนายต่อไป

2.2 ตรวจสอบความเหมาะสมของรูปแบบ (Diagnostic Checking) เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ได้แล้วจึงตรวจสอบความเหมาะสมของรูปแบบ เช่น พิจารณากราฟค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติของค่าความผิดพลาด (Residuals) ที่ได้จากรูปแบบ หรือการทดสอบค่าความเหมาะสมของรูปแบบโดยใช้การทดสอบของ Box และ Pierce หรือใช้การทดสอบของ Box และ Ljung การทดสอบค่าความเหมาะสมของพารามิเตอร์ด้วยแบบทดสอบแบบ t หากพบว่ารูปแบบที่ได้รับการทดสอบเหมาะสมก็ใช้รูปแบบในการทำนายต่อไป แต่ถ้ารูปแบบไม่เหมาะสมต้องย้อนกลับไปทำหัวข้อที่ 1.2 จนกว่าจะได้รูปแบบที่เหมาะสม

3. การใช้โมเดลเพื่อทำนาย (Forecasting) ซึ่งมี 2 แบบคือการทำนายแบบจุด (Point Forecast) และการทำนายแบบช่วง (Interval Forecast)

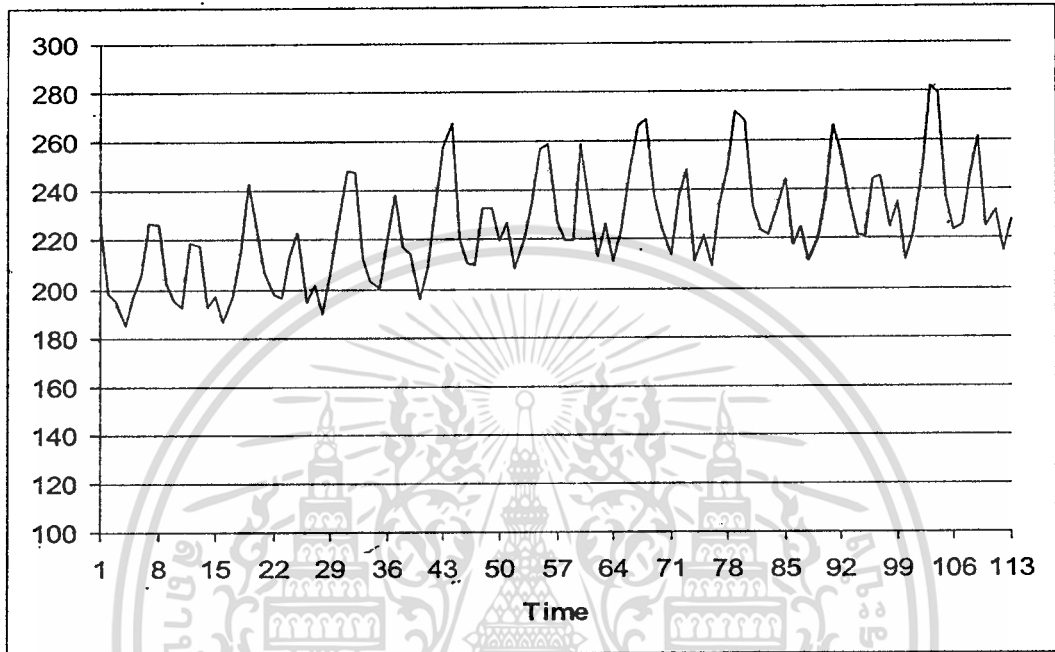
2.8.3 อนุกรมเวลาที่เป็นสเตชันนารี (Stationary Time Series)

การกำหนดรูปแบบ ARIMA(p,d,q) ให้กับอนุกรมเวลาจะกระทำกับข้อมูลที่เป็นสเตชันนารีเท่านั้น ถ้าให้ $\{Y_t\}$ เป็นอนุกรมเวลา ลักษณะของข้อมูลที่เป็นสเตชันนารีคืออนุกรมเวลาที่มีค่าเฉลี่ย $E(Y_t)$ และค่าความแปรปรวน $V(Y_t)$ คงที่สำหรับแต่ละช่วงเวลา t ซึ่งอนุกรมเวลาที่มีแนวโน้มและหรือฤดูกาลจะมี $E(Y_t)$ ไม่คงที่ การพิจารณาว่าข้อมูลเป็นสเตชันนารีหรือไม่พิจารณาจาก

1. $E(Y_t)$ คงที่สำหรับทุกค่าของ t หรือไม่ การพิจารณาทำได้โดยการแบ่งอนุกรมเวลาออกเป็นส่วนๆ แล้วหาค่าเฉลี่ยของอนุกรมเวลาของแต่ละส่วน ถ้าค่าเฉลี่ยของแต่ละส่วนย่อยไม่ต่างกันมากจะสรุปได้ว่า $E(Y_t)$ คงที่สำหรับแต่ละช่วงเวลา t
2. $V(Y_t)$ คงที่สำหรับทุกค่าของ t หรือไม่ การพิจารณาจะกระทำในทำนองเดียวกับการพิจารณาว่า $E(Y_t)$ คงที่หรือไม่โดยการหาส่วนเบี่ยงเบนมาตรฐานของอนุกรมเวลาแต่ละส่วน ถ้า

ส่วนเบี่ยงเบนมาตรฐานในแต่ละส่วนย่อยไม่ต่างกันมากจะสรุปได้ว่า $V(Y_t)$ คงที่สำหรับแต่ละช่วงเวลา t

3. พิจารณาแนวโน้มหรือฤดูกาลจากกราฟอนุกรมเวลา วิธีนี้เป็นวิธีที่ง่ายที่สุดในกรณีที่มีแนวโน้มและหรือฤดูกาลมักจะเห็นจากรูปชัดเจน

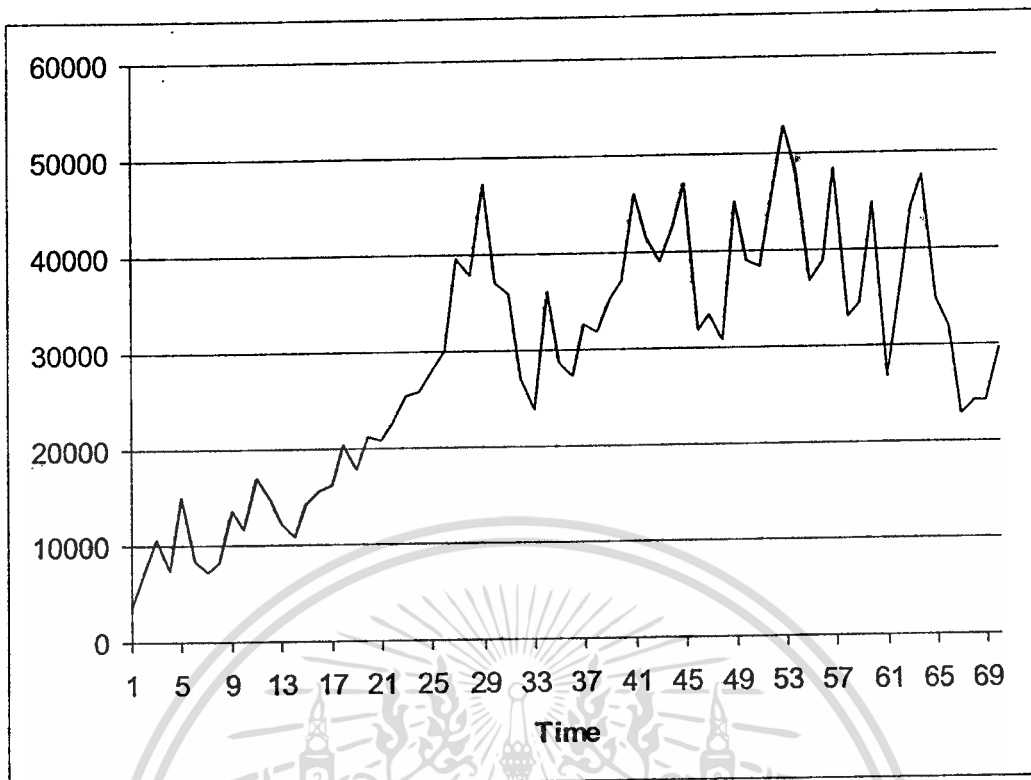


รูปที่ 2.9 แสดงลักษณะอนุกรมเวลาที่ไม่เป็นสเตชันนารีเนื่องจากแนวโน้มและอิทธิพลฤดูกาล

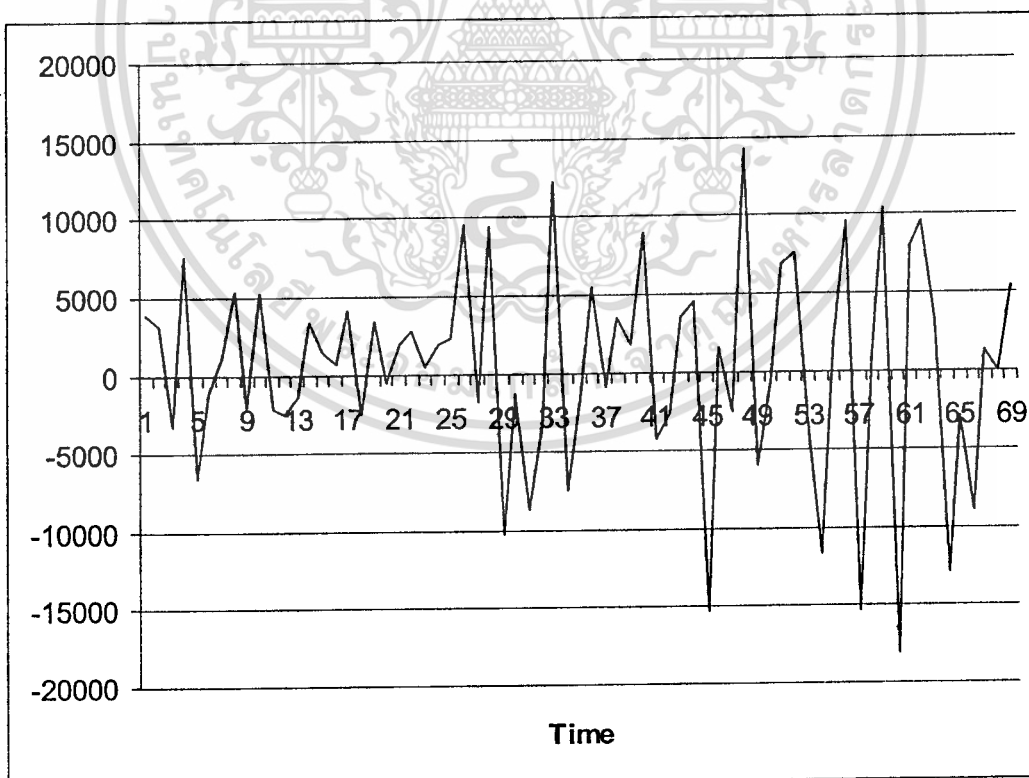
จากรูปที่ 2.9 เป็นลักษณะของอนุกรมเวลาที่ไม่เป็นสเตชันนารีเพราะจะเห็นว่าอนุกรมเวลามีฤดูกาลและมีแนวโน้มรวมอยู่ด้วยกัน อนุกรมเวลาที่ไม่เป็นสเตชันนารีจะต้องนอร์มัลไลซ์ให้เป็นสเตชันนารีก่อน

วิธีการนอร์มัลไลซ์ให้อนุกรมเวลาเป็นสเตชันนารี

1. การหาผลต่าง (Regular Differencing) ของอนุกรมเวลาเพื่อลบแนวโน้มออกจากอนุกรมเวลา ถ้าอนุกรมเวลา $\{Y_t\}$ เป็นอนุกรมเวลาที่มีแนวโน้มจะแปลงเป็นอนุกรมเวลาที่ไม่แนวโน้ม $\{Z_t\}$ โดย $Z_t = \nabla^d Y_t$ เมื่อ d เป็นลำดับของการหาผลต่าง ตัวอย่างเช่น ถ้า $d = 1$, $Z_t = \nabla Y_t = Y_t - Y_{t-1}$ ถ้า $d = 2$, $Z_t = \nabla^2 Y_t = \nabla Y_t - \nabla Y_{t-1} = Y_t - Y_{t-1} - Y_{t-1} + Y_{t-2} = Y_t - 2Y_{t-1} + Y_{t-2}$ เป็นต้น จำนวนครั้งของการหาผลต่างขึ้นอยู่กับว่าเมื่อหาผลต่างแล้วอนุกรมเวลาใหม่เป็นสเตชันนารีหรือไม่ ถ้ายังไม่เป็นต้องหาผลต่างต่อไป โดยทั่วไปถ้าอนุกรมเวลาที่มีแนวโน้มเป็นเส้นตรง (Linear) จะใช้ $d = 1$ แต่ถ้าอนุกรมเวลาที่แนวโน้มแบบกำลังสอง (Quadratic) จะใช้ $d = 2$



รูปที่ 2.10 แสดงลักษณะอนุกรมเวลาที่ไม่เป็นสเดชันนารีเนื่องจากแนวโน้ม

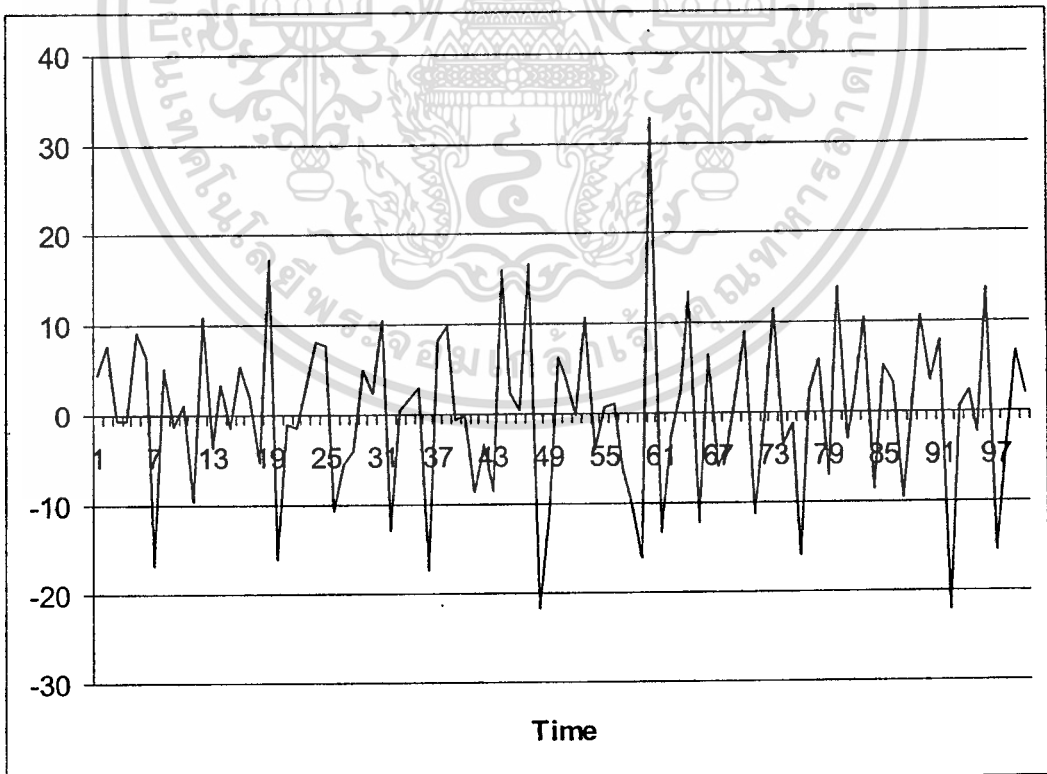


รูปที่ 2.11 แสดงลักษณะอนุกรมเวลาที่เป็นสเดชันนารีหลังจากลบแนวโน้ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การหาผลต่างฤดูกาล (Seasonal Differencing) ของอนุกรมเวลา ถ้าอนุกรมเวลาไม่มีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง จะแปลงอนุกรมเวลาเดิม $\{Y_t\}$ ให้เป็นอนุกรมเวลาใหม่ที่ไม่มีอิทธิพลฤดูกาล $\{Z_t\}$ โดย $Z_t = \nabla_L^D Y_t$ โดย D เป็นลำดับของการหาผลต่างและ L เป็นจำนวนฤดูกาลต่อปี เช่น อนุกรมเวลารายเดือน (L = 12) ถ้า D = 1, $Z_t = \nabla_{12} Y_t = Y_t - Y_{t-12}$ ถ้า D = 2, $Z_t = \nabla_{12}^2 Y_t = \nabla_{12}(Y_t - Y_{t-12}) = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-24} = Y_t - 2Y_{t-12} + Y_{t-24}$ เป็นต้น ผลต่างจะหนักี่ครั้งขึ้นอยู่กับว่าเมื่อหาผลต่างแล้วอนุกรมเวลาเป็นสเตชันนารีหรือไม่ ถ้ายังก็ต่อต้องหาผลต่างต่อไป

3. การหาผลต่างและการหาผลต่างฤดูกาล (Regular Differencing and Seasonal Differencing) กรณีที่อนุกรมเวลาที่มีทั้งแนวโน้มและฤดูกาล การปรับอนุกรมเวลาให้เป็นสเตชันนารี จะทำโดยการหาผลต่างและผลต่างฤดูกาล โดยที่อันดับการหาผลต่าง d และ D ขึ้นอยู่กับว่าเมื่อทำแล้วอนุกรมเวลาเป็นสเตชันนารีหรือยัง เช่น อนุกรมเวลารายเดือนที่มีทั้งแนวโน้มและฤดูกาล เมื่อ d = 1 และ D = 1 จะแปลงอนุกรมเวลาเดิม $\{Y_t\}$ ให้เป็นอนุกรมเวลาใหม่ $\{Z_t\}$ ได้โดย $Z_t = \nabla \nabla_{12} Y_t = \nabla(Y_t - Y_{t-12}) = \nabla Y_t - \nabla Y_{t-12} = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$ เป็นต้น จากรูปที่ 2.9 หลังจากหาผลต่างและผลต่างฤดูกาลแล้วปรากฏว่าอนุกรมเวลามีการเปลี่ยนแปลงรอบๆค่า 0 ดังนั้นอนุกรมเวลาจึงเป็นสเตชันนารี



รูปที่ 2.12 แสดงลักษณะอนุกรมเวลาที่เป็นสเตชันนารีหลังจากลบแนวโน้มและฤดูกาล

4. การหาลอการิทึมของค่าสังเกตในอนุกรมเวลา นั่นคือการแปลงอนุกรมเวลา $\{Y_t\}$ ให้เป็นอนุกรมเวลาใหม่ $\{Z_t\}$ โดยที่ $Z_t = \log(Y_t)$ การแปลงนี้จะกระทำเมื่อความผันแปรของอนุกรมเวลาไม่คงที่ นั่นคือ $V(Y_t)$ ไม่คงที่สำหรับแต่ละช่วงเวลา t

2.8.4 การกำหนดรูปแบบ (Model Identification)

การกำหนดรูปแบบจะทำได้โดยการพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติ ซึ่งคำนวณจากสมการดังนี้
ค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติ

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (2.36)$$

ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติ

$$r_{kk} = \begin{cases} r_1 & \text{เมื่อ } k=1 \\ \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \text{เมื่อ } k=2,3,\dots \end{cases} \quad (2.37)$$

โดยที่

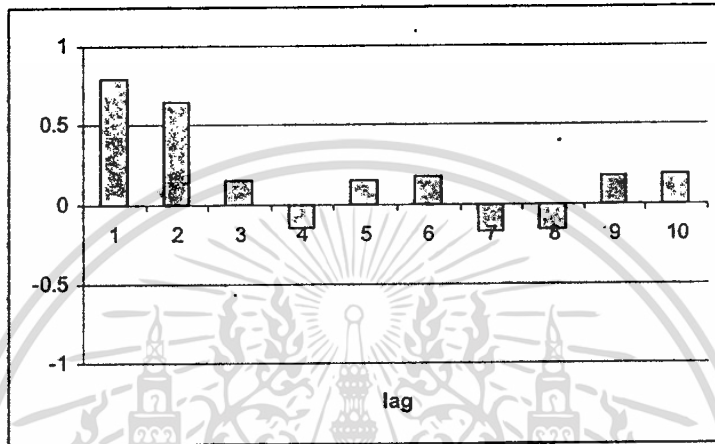
$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1} r_{k-j} \quad \text{เมื่อ } j=1,2,\dots,k-1 \quad (2.38)$$

ค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติ หรือ r_k มีความหมายทำนองเดียวกับค่าสัมประสิทธิ์สหสัมพันธ์แต่เป็นการวัดความสัมพันธ์ของอนุกรมเวลากับตัวมันเอง นั่นคือ เป็นค่าวัดสหสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่อยู่ห่างกัน k ช่วงเวลาหรือค่าวัดสหสัมพันธ์ระหว่าง Y_t และ Y_{t-k} ค่า r_k จะอยู่ระหว่าง -1 ถึง 1 ถ้า $|r_k|$ มีค่าเข้าใกล้ 1 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลา มีสหสัมพันธ์กันสูง ถ้า $|r_k|$ มีค่าเข้าใกล้ 0 แสดงว่าค่าสังเกตที่อยู่ห่างกัน k ช่วงเวลา มีสหสัมพันธ์กันต่ำ ถ้า r_k มีค่าน้อยกว่า 0 แสดงว่ามีสหสัมพันธ์กันทางตรงกันข้ามและถ้า r_k มีค่ามากกว่า 0 แสดงว่ามีสหสัมพันธ์ในทางตามกัน

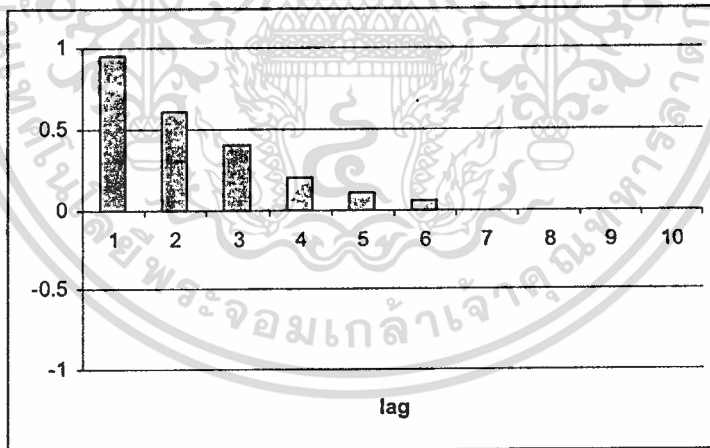
การกำหนดรูปแบบให้อนุกรมเวลาจะพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติที่อยู่ในและนอกช่วงวิกฤติ นั่นคือพิจารณาว่า

$$|r_k| > \frac{2}{\sqrt{n}} \text{ และพิจารณาว่า } |r_{kk}| > \frac{2}{\sqrt{n}} \text{ เมื่อ } k \text{ คือจำนวนค่าสังเกตในอนุกรมเวลา}$$

ต่อไปนี้เป็นลักษณะกราฟของค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติ

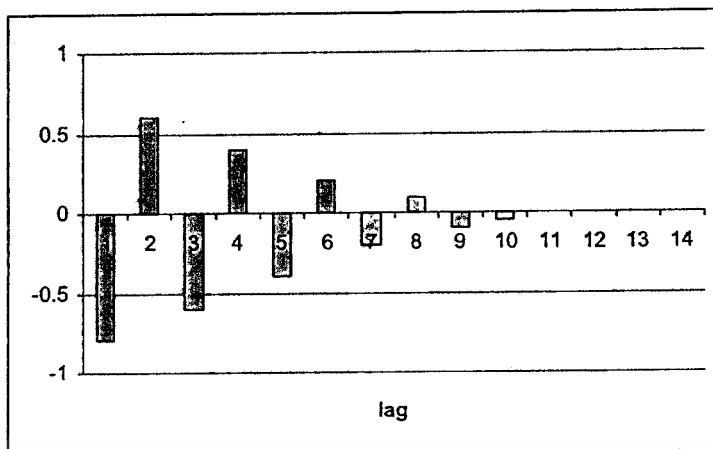


รูปที่ 2.13 แสดงลักษณะของ ACF หรือ PACF ที่หายไป(Cut Off)

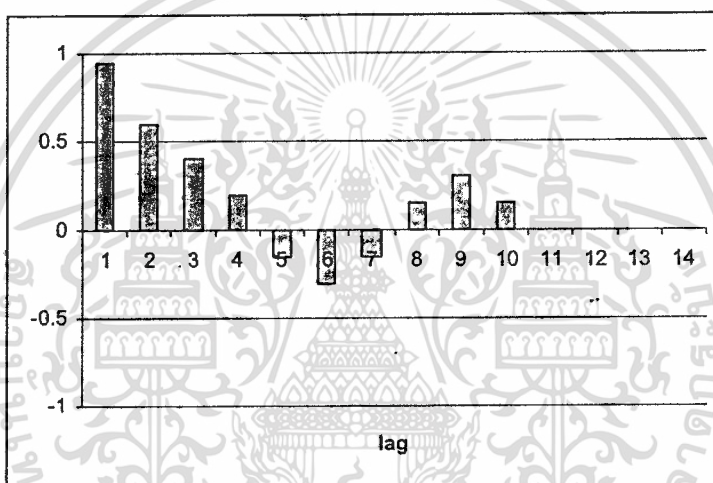


รูปที่ 2.14 แสดงลักษณะของ ACF หรือ PACF ที่ลดลงแบบ exponential แต่ไม่แกว่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.15 แสดงลักษณะของ ACF หรือ PACF ที่ลดลงแบบ exponential แต่แกว่ง



รูปที่ 2.16 แสดงลักษณะของ ACF หรือ PACF ที่มีรูปแบบ sine wave

จากรูปที่ 2.13 กราฟ spike ที่ lag 1 และ 2 และหัดสั้นเข้าหรือหายไป (Cut off) หลัง lag ที่ 2 รูปที่ 2.14 แสดงลักษณะการลดลงแบบ exponential แต่ไม่แกว่ง (Oscillation) รูปที่ 2.15 แสดงลักษณะการลดลงแบบ exponential แต่แกว่งและรูปที่ 2.16 แสดงลักษณะการลดลงแบบ sine wave

จากลักษณะกราฟของค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติจะถูกนำมาใช้เป็นแนวทางในการกำหนดรูปแบบของอนุกรมเวลา แนวทางในการพิจารณา $|r_x|$ และ $|r_{kk}|$ เพื่อกำหนดรูปแบบสำหรับอนุกรมเวลาที่ไม่ใช่ฤดูกาล อธิบายตามตารางที่ (2.1) [12]

ตารางที่ 2.1 แนวทางการกำหนดรูปแบบสำหรับอนุกรมเวลาที่ไม่มียุคฤดูกาล

แนวทางที่	ACF	PACF	รูปแบบที่ใช้
1	ACF spike ที่ lag 1,2,...,q และหดหรือตัดหายไป (Cut Off) หลัง lag q	PACF ค่อยๆ ลดลง (Die Down)	ใช้ IMA order q
	<p>ตัวอย่างเช่น</p> <ol style="list-style-type: none"> ACF spike ที่ lag 1 และตัดหายไปหลัง lag 1 และ PACF ค่อยๆ ลดลงในรูปแบบที่เป็น sine wave หรือลดลงเร็วแบบ exponential รูปแบบที่ใช้คือ IMA(1) ACF spike ที่ lag 1 และ 2 และตัดหายไปหลัง lag 2 และ PACF ค่อยๆ ลดลงในรูปแบบที่เป็น sine wave หรือลดลงเร็วแบบ exponential รูปแบบที่ใช้คือ IMA(2) 		
2	ACF ค่อยๆ ลดลง	PACF spike ที่ lag 1,2,...,p และหดหรือตัดหายไปหลัง lag p	ใช้ ARI order p
	<p>ตัวอย่างเช่น</p> <ol style="list-style-type: none"> PACF spike ที่ lag 1 และตัดหายไปหลัง lag 1 และ ACF ค่อยๆ ลดลงในรูปแบบที่เป็น sine wave หรือลดลงเร็วแบบ exponential รูปแบบที่ใช้คือ ARI(1) PACF spike ที่ lag 1 และ 2 และตัดหายไปหลัง lag 2 และ ACF ค่อยๆ ลดลงในรูปแบบที่เป็น sine wave หรือลดลงเร็วแบบ exponential รูปแบบที่ใช้คือ ARI(2) 		
3	ACF spike ที่ lag 1,2,...,q และตัดหายไปหลัง lag q	PACF spike ที่ lag 1,2,...,p และตัดหายไปหลัง lag p	ใช้ ARI order p หรือ IMA order q
	<p>ตัวอย่างเช่น</p> <ol style="list-style-type: none"> ถ้า ACF ตัดหายไปเร็วกว่า PACF ใช้รูปแบบ IMA แต่ถ้า PACF ตัดหายไปเร็วกว่า ACF ใช้รูปแบบ ARI ถ้า ACF ตัดหายไปเร็วเท่าๆ กับ PACF ใช้รูปแบบ ARI order p หรือ IMA order q อย่างใดอย่างหนึ่งเท่านั้นโดยเลือกรูปแบบที่ให้ค่าความถูกต้องมากที่สุด โดยทั่วไป IMA order q มักจะเป็นรูปแบบที่ดี ดังนั้นให้พิจารณารูปแบบ IMA order q ก่อน 		
4	ไม่มี ACF spike ที่ lag ไต ๆ	ไม่มี PACF spike ที่ lag ไต ๆ	ไม่ต้องกำหนดรูปแบบ
5	ACF มีรูปแบบที่ลดลง	PACF มีรูปแบบที่ลดลง	กำหนดทั้ง ARI และ IMA ในรูปแบบ
	ถ้าทั้ง ACF และ PACF มีลักษณะที่ลดลงทั้งสอง ให้กำหนดรูปแบบอย่างง่ายก็เพียงพอ เช่น ARIMA(1,d,1) เป็นต้น		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำหรับแนวทางในการพิจารณา $|r_k|$ และ $|r_{kk}|$ เพื่อกำหนดรูปแบบสำหรับอนุกรมเวลาที่มีฤดูกาลปรากฏตามตารางที่ (2.2) โดยที่ L คือฤดูกาล [12]

ตารางที่ 2.2 แนวทางการกำหนดรูปแบบสำหรับอนุกรมเวลาที่มีฤดูกาล

แนวทางที่	ACF	PACF	รูปแบบฤดูกาลที่ใช้
6	ACF spike ที่ lag $L, 2L, \dots, QL$ และหายไปหรือหัดสั้น (Cut Off) เข้าหลัง lag QL	PACF ค่อยๆ หดหาย (Die Down)	ใช้ IMA order Q
	<p>ตัวอย่างเช่น</p> <p>ถ้า ACF spike ที่ lag L และตัดหายไปหลัง lag L รูปแบบที่ใช้คือ $IMA(1)_L$ เช่น ถ้าเป็นข้อมูลรายเดือนและ 1 ฤดูกาลมี 12 เดือน ดังนั้น $L = 12$ ถ้า ACF spike ที่ lag 12 และตัดหายไปหลัง lag 12 รูปแบบที่ใช้คือ $IMA(1)_{12}$ ถ้า ACF spike ที่ lag 12 และ lag 24 และตัดหายไปหลัง lag 24 รูปแบบที่ใช้คือ $IMA(2)_{12}$ เป็นต้น</p>		
7	ACF ค่อยๆ หดหาย (Die Down)	PACF spike ที่ lag $L, 2L, \dots, PL$ และหายไปหรือหัดสั้น (Cut Off) เข้าหลัง lag PL	ใช้ ARI order P
	<p>ตัวอย่างเช่น</p> <p>ถ้า PACF spike ที่ lag L และตัดหายไปหลัง lag L รูปแบบที่ใช้คือ $ARI(1)_L$ เช่น ถ้าเป็นข้อมูลรายเดือนและ 1 ฤดูกาลมี 12 เดือน ดังนั้น $L = 12$ ถ้า PACF spike ที่ lag 12 และตัดหายไปหลัง lag 12 รูปแบบที่ใช้คือ $ARI(1)_{12}$ ถ้า PACF spike ที่ lag 12 และ lag 24 และตัดหายไปหลัง lag 24 รูปแบบที่ใช้คือ $ARI(2)_{12}$ เป็นต้น</p>		
8	ACF spike ที่ lag $L, 2L, \dots, QL$ และตัดหายไปหลัง lag QL	PACF spike ที่ lag $1L, 2L, \dots, PL$ และตัดหายไปหลัง lag PL	ใช้ ARI order P หรือ IMA order Q
	<p>ตัวอย่างเช่น</p> <ol style="list-style-type: none"> 1. ถ้า ACF ตัดหายไปเร็วกว่า PACF ใช้รูปแบบ IMA แต่ถ้า PACF ตัดหายไปเร็วกว่า ACF ใช้รูปแบบ ARI 2. ถ้า ACF ตัดหายไปเร็วเท่าๆ กับ PACF ใช้รูปแบบ ARI order Q หรือ IMA order P อย่างไรก็ตามหนึ่งเท่านั้นโดยเลือกรูปแบบที่ให้ค่าความถูกต้องมากที่สุด โดยทั่วไป IMA order Q มักจะเป็นรูปแบบที่ดี ดังนั้นให้พิจารณารูปแบบ IMA order Q ก่อน 		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 (ต่อ)

แนวทาง ที่	ACF	PACF	รูปแบบฤดูกาลที่ใช้
4	ไม่มี ACF spike ที่ seasonal lag ใด ๆ	ไม่มี PACF spike ที่ seasonal lag ใด ๆ	ไม่ต้องกำหนดรูปแบบ ฤดูกาล
5	ACF มีรูปแบบที่ลดลงอย่าง รวดเร็วที่ seasonal lag	PACF มีรูปแบบที่ลดลงอย่าง รวดเร็วที่ seasonal lag	กำหนดทั้ง ARI และ IMA ในรูปแบบ
	ถ้าทั้ง ACF และ PACF มีลักษณะที่ลดลงทั้งสอง ให้กำหนดรูปแบบอย่างง่ายก็เพียงพอ เช่น SARIMA(1,D,1) เป็นต้น		

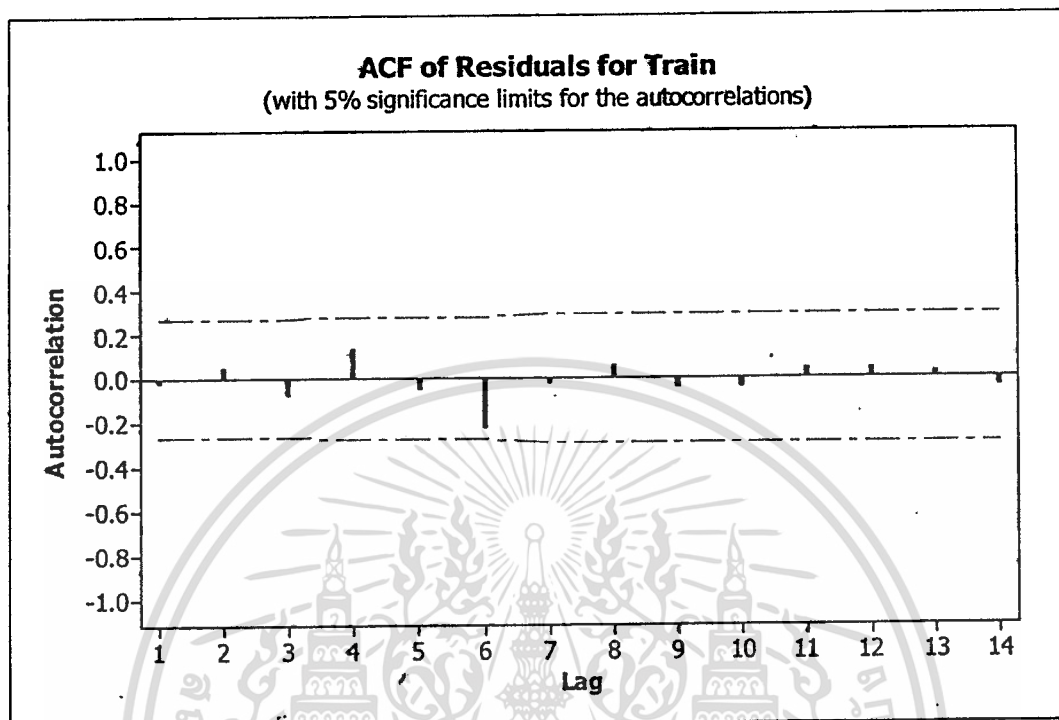
จากตารางที่ 2.1 และ ตารางที่ 2.2 เป็นเพียงแนวทางเบื้องต้นในการกำหนดรูปแบบเท่านั้นเพื่อให้การกำหนดรูปแบบง่ายขึ้น แต่จริงๆ แล้วรูปแบบที่เหมาะสมอาจไม่เป็นไปตามแนวทางที่แนะนำไว้ก็ได้ ซึ่งต้องลองถูกลองผิดและอาศัยประสบการณ์จึงจะสามารถกำหนดรูปแบบที่เหมาะสมเร็วยิ่งขึ้น สำหรับการพิจารณาใส่ค่าคงที่ในรูปแบบนั้นจะพิจารณาที่ค่าเฉลี่ยของอนุกรมเวลานั้นคือ ถ้าค่าเฉลี่ยมีค่าเข้าใกล้ 0 มากหรือค่าเฉลี่ยมีค่าสูงหรือต่ำมากเมื่อเทียบกับค่าเบี่ยงเบนมาตรฐานของอนุกรมเวลานั้นจะไม่ใส่ค่าคงที่ในรูปแบบ

2.8.5 การทดสอบความเหมาะสมของรูปแบบ

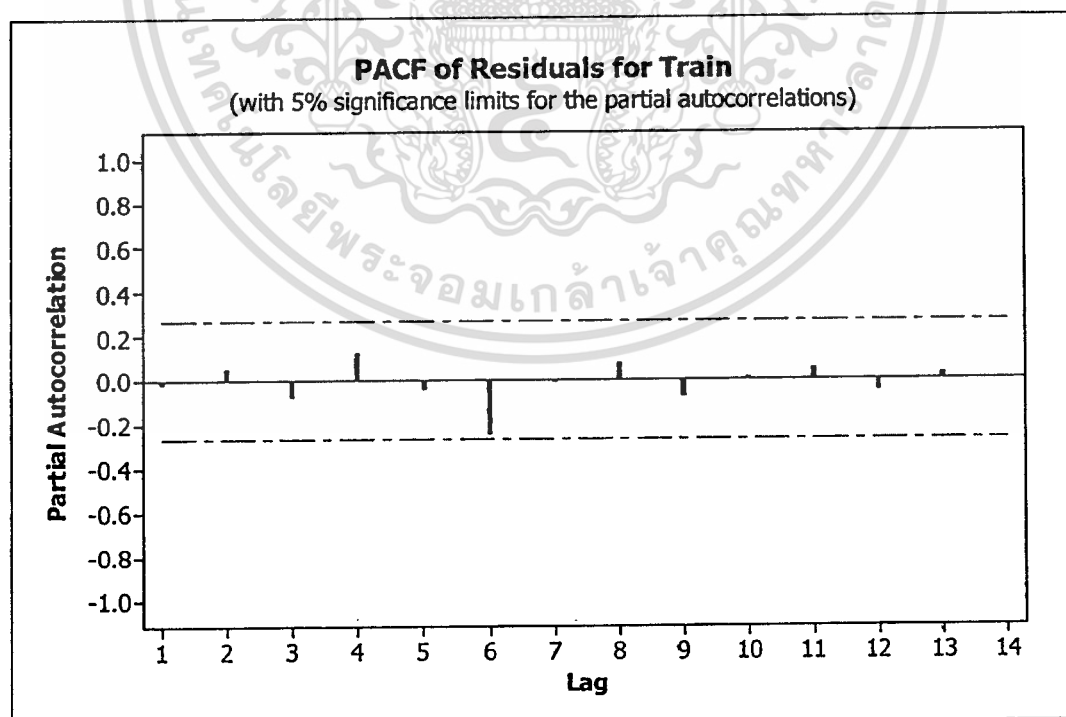
หลังจากกำหนดรูปแบบให้กับอนุกรมเวลาแล้วต้องมีการทดสอบความเหมาะสมของรูปแบบ วิธีการตรวจสอบรูปแบบนั้นมีหลายวิธีซึ่งการทดสอบนั้นจะใช้หลายวิธีพร้อมกัน

1. การพิจารณาอนุกรมเวลาของค่าความคลาดเคลื่อนหรือค่าความผิดพลาด (Residual หรือ Error) ของรูปแบบ $\{e_t\}$ ว่ามีลักษณะการเคลื่อนไหวที่เป็นอิสระกันหรือไม่ ซึ่งค่าความคลาดเคลื่อนของการทำนาย e_t คำนวณได้จากผลต่างระหว่างค่าจริงและค่าทำนาย การทดสอบนี้เป็นการตรวจสอบข้อจำกัดของรูปแบบที่กำหนดว่า e_t ต้องมีการแจกแจงที่เป็นอิสระกัน หากพบว่า $\{e_t\}$ มีลักษณะการเคลื่อนไหวที่เป็นอิสระกันจริงสรุปได้ว่ารูปแบบนั้นเหมาะสมกับอนุกรมเวลาแล้วซึ่งทำได้โดยการทดสอบสมมติฐาน $H_0 : r_k(e_t) = 0$ และ $H_1 : r_k(e_t) \neq 0$ สำหรับ $k = 1, 2, \dots$ โดยจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.05 ถ้า $|r_k(e_t)| \geq 2/\sqrt{n}$ เมื่อ n คือขนาดของอนุกรมเวลา $\{e_t\}$ ตัวอย่างเช่น ในโปรแกรม Minitab จะพิจารณากราฟค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติของ $\{e_t\}$ ถ้าค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติและค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนแบบอัตโนมัติอยู่ในช่วงค่าวิกฤติ (เส้นประ) ก็ถือว่ารูปแบบนั้นเหมาะสมกับอนุกรมเวลานั้นหมายความว่าค่าความคลาดเคลื่อนหรือ

ค่าความผิดพลาดไม่มีความสัมพันธ์กันหรือมีการเคลื่อนไหวที่เป็นอิสระกัน รูปที่ 2.17 และรูปที่ 2.18 ประกอบ



รูปที่ 2.17 กราฟ ACF ของ residuals ที่เป็นอิสระกัน



รูปที่ 2.18 กราฟ PACF ของ residuals ที่เป็นอิสระกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. การพิจารณาว่า $r_k(e_t) = 0$ สำหรับ $k = 1, \dots, m$ หรือไม่ นั่นคือพิจารณาว่าค่าความคลาดเคลื่อนของการทำนายที่อยู่ห่างกัน $1, \dots, m$ ช่วงเวลาว่าเป็นอิสระกันหรือไม่ โดยทดสอบสมมติฐาน $H_0 : r_1(e_t) = \dots = r_m(e_t) = 0$ กับ $H_1 : r_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 สำหรับ $k = 1, \dots, m$ การทดสอบนี้จะใช้การทดสอบของ Box และ Pierce หรือการทดสอบของ Box และ Ljung ซึ่งวิธีการทดสอบทั้งสองกำหนดสมมติฐานหลักและสมมติฐานทางเลือกเหมือนกัน มีช่วงวิกฤติเหมือนกันแต่ใช้ตัวทดสอบสถิติต่างกัน วิธีการมีดังนี้คือ

2.1 สมมติฐานที่ใช้คือ

$$H_0 : r_1(e_t) = \dots = r_m(e_t) = 0$$

$$H_1 : r_k(e_t) \text{ อย่างน้อยหนึ่งค่าไม่เท่ากับ } 0 \text{ สำหรับ } k = 1, \dots, m$$

ความหมายคือ ถ้า r_k เท่ากับ 0 เมื่อ $k = 1, \dots, m$ แสดงว่าค่าความคลาดเคลื่อนไม่มีความสัมพันธ์กัน ในทางตรงกันข้ามถ้า มี r_k อย่างน้อยหนึ่งค่ามีค่าไม่เท่ากับ 0 เมื่อ $k = 1, \dots, m$ แสดงว่าค่าความคลาดเคลื่อนมีความสัมพันธ์กัน

2.2 ตัวทดสอบสถิติ

ตัวทดสอบสถิติที่ใช้กันทั่วไปคือตัวทดสอบ Box และ Pierce และตัวทดสอบ Box และ Ljung ซึ่งตัวทดสอบตัวหลังนี้ปรับปรุงมาจากตัวทดสอบ Box และ Pierce ตัวทดสอบสถิติทั้งสองตัวนี้ใช้ทดสอบ r_k เมื่อ $k = 1, 2, \dots, m$ พร้อมกันในครั้งเดียวว่ามีค่าต่างจาก 0 หรือไม่ ถ้า r_k มีค่าเข้าใกล้ 0 ค่า Q และค่า Q_m จะมีค่าต่ำ แต่ถ้า $|r_k|$ มีค่ามากกว่า Q และค่า Q_m จะมีค่าสูง คำนวณได้ดังสมการ

$$\text{Box-Pierce} \quad Q = n \sum_{k=1}^m r_k^2(e_t) \quad (2.39)$$

$$\text{Box-Ljung} \quad Q_m = n(n+2) \sum_{k=1}^m (r_k^2(e_t)) / (n-k) \quad (2.40)$$

เมื่อ n คือขนาดอนุกรมเวลา $\{e_t\}$

m คือจำนวน lag สูงสุดของ e_t ในอนุกรมเวลา $\{e_t\}$ ที่กำลังพิจารณา

k คือ lag ที่กำลังพิจารณา

2.3 ช่วงวิกฤติ

หลังจากคำนวณได้ค่า Q และค่า Q_m ก็นำไปเปรียบเทียบกับช่วงวิกฤติ ซึ่งช่วงวิกฤติคำนวณได้โดย

$$Q \text{ หรือ } Q_m \geq \chi^2_{\alpha, m-a} \quad (2.41)$$

เมื่อ α คือระดับความเชื่อมั่นที่ใช้ ซึ่งโดยส่วนมากใช้ระดับความเชื่อมั่นที่ 0.05

m คือจำนวน lag สูงสุดของ e_t ในอนุกรมเวลา $\{e_t\}$ ที่กำลังพิจารณา ถ้ายังไม่ได้กำหนดรูปแบบใดๆ ให้กับอนุกรมเวลาจะใช้ $m = 0$

a คือจำนวนพารามิเตอร์ทั้งหมดในรูปแบบรวมทั้งค่าคงที่ด้วย

2.4 สรุปผลการทดสอบ

ถ้าค่า Q หรือ Q_m น้อยกว่าช่วงวิกฤติจะปฏิเสธ H_1 ยอมรับ H_0 นั่นคือ $r_k(e_t)$ เท่ากับ 0 เมื่อ $k = 1, \dots, m$ นั่นคือรูปแบบที่กำหนดให้อนุกรมเวลานั้นเหมาะสมแล้ว แต่ถ้า Q หรือ Q_m มากกว่าช่วงวิกฤติปฏิเสธ H_0 นั่นคือมี $r_k(e_t)$ อย่างน้อยหนึ่งค่าไม่เท่ากับ 0 เมื่อ $k = 1, \dots, m$ นั่นคือรูปแบบที่กำหนดให้อนุกรมเวลานั้นยังไม่เหมาะสม

ในโปรแกรม Minitab จะใช้ตัวทดสอบของ Box และ Ljung ซึ่งโปรแกรมจะคำนวณให้ โดยจะคำนวณที่ lag ต่างๆ ดังตัวอย่าง

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	5.7	16.6	23.4	30.9
DF	9	21	33	45
P-Value	0.773	0.734	0.893	0.946

จากตัวอย่างโปรแกรม Minitab จะคำนวณหาค่า chi-square ของตัวทดสอบ Box และ Ljung ที่ lag 12, 24, 36, 48 เพื่อทดสอบสมมติฐานว่า residual มีความสัมพันธ์กันหรือไม่โดยพิจารณาค่าสัมประสิทธิ์สหสัมพันธ์แบบอัตโนมัติที่ lag ต่างๆ พร้อมกัน

จากการคำนวณช่วงวิกฤติของ Chi-Square ที่ lag ต่างๆ เมื่อ n คือ lag ที่กำลังพิจารณา a คือจำนวนพารามิเตอร์ในรูปแบบและระดับความเชื่อมั่น $\alpha = 0.05$ เป็นดังนี้คือ

$$\text{ที่ lag } 12 = \chi^2_{0.05, 12-3} = 16.919$$

$$\text{ที่ lag } 24 = \chi^2_{0.05, 24-3} = 32.67$$

$$\text{ที่ lag } 36 = \chi^2_{0.05, 36-3} = 51.5$$

$$\text{ที่ lag } 48 = \chi^2_{0.05, 48-3} = 66$$

จากตัวอย่างวิเคราะห์ได้ว่าค่า chi-square ไม่ออกนอกช่วงวิกฤติ ดังนั้น ยอมรับ H_0 นั่นคือ r_k ของค่าความคลาดเคลื่อนที่ lag ต่างๆ มีค่าเท่ากับ 0 ตามสมมติฐานหรือค่าความคลาดเคลื่อนไม่มีความสัมพันธ์กันนั่นเอง

จากตัวอย่างจะเห็นว่าโปรแกรม Minitab จะคำนวณค่า P-Value (Probability Value) ของค่า chi-square ให้ด้วย หลักการพิจารณาค่า P-Value จะพิจารณาเปรียบเทียบกับค่า α -level หรือค่าความเชื่อมั่นที่กำหนดไว้ก่อนหน้า ซึ่งโดยทั่วไปจะกำหนดที่ 0.05 หลักการพิจารณาคือ

ถ้า P-Value น้อยกว่าหรือเท่ากับ α -level จะปฏิเสธ H_0

ถ้า P-Value มากกว่า α -level จะปฏิเสธ H_1

จากการวิเคราะห์โดยโปรแกรม Minitab พบว่าค่า P-Value ที่ lag ต่างๆ มากกว่า 0.05 ทุกค่า ดังนั้นจึงยอมรับ H_0 ซึ่งสอดคล้องกับวิเคราะห์ค่า chi-square ของ Box' และ Ljung จากตัวอย่างจะเห็นว่าในโปรแกรม Minitab การพิจารณาค่า P-Value จะสะดวกกว่าการพิจารณาค่า chi-square เนื่องจากสามารถเปรียบเทียบค่า P-Value กับค่า α -level ได้เลย

3. การพิจารณาว่าค่าสัมประสิทธิ์ในรูปแบบมีค่าเป็น 0 หรือไม่ โดยการทดสอบสมมติฐานคือ

H_0 : ค่าสัมประสิทธิ์ในรูปแบบ = 0

H_1 : ค่าสัมประสิทธิ์ในรูปแบบ ไม่เท่ากับ 0

ตัวอย่างการคำนวณของโปรแกรม Minitab

Type	Coef	SE Coef	T	P
AR 1	0.5216	0.1335	3.91	0.000
MA 1	0.9635	0.0672	14.34	0.000
Constant	337.10	46.64	7.23	0.000
Mean	704.66	97.50		

สิ่งที่ต้องพิจารณา

1. พิจารณาค่า P-Value หลักการพิจารณาคือ

ถ้า P-value น้อยกว่าหรือเท่ากับ 0.05 จะยอมรับ H_1

ถ้า P-value มากกว่า 0.05 จะยอมรับ H_0

ความหมายของ P-Value คือถ้า P-Value = 0.00% คือมีโอกาส 0.00 % ที่ค่าสัมประสิทธิ์ในรูปแบบ จะมีค่าเป็น 0

2. พิจารณาค่า T-value นั่นคือ ค่า absolute T-value ต้องมากกว่า 2 ซึ่งหมายความว่าค่าสัมประสิทธิ์ในรูปแบบมีค่าต่างจาก 0 อย่างมีนัยสำคัญ

3. SE Coef เป็นค่าใช้ตรวจสอบความแม่นยำในการประมาณค่า ถ้าค่าต่ำแสดงว่ามีความแม่นยำในการประมาณสูง

จากตัวอย่างจะเห็นว่าทั้งค่า P-Value และค่า T-Value ผ่านเงื่อนไข ดังนั้นจะยอมรับ H_1 ดังนั้นค่าสัมประสิทธิ์ที่กำหนดในรูปแบบจึงเหมาะสมแล้ว

ในกรณีที่การทดสอบค่าสัมประสิทธิ์ไม่ผ่านซึ่งอาจเกิดเนื่องจาก multicollinearity ของอนุกรมเวลานั้นคือค่าพารามิเตอร์มีความสัมพันธ์กัน ในกรณีนี้ให้พิจารณาข้อที่ 1 และข้อที่ 2 เป็นสำคัญถ้าผ่านก็ถือว่ารูปแบบที่กำหนดนั้นเหมาะสมแล้ว

4. ถ้ากำหนดรูปแบบให้อนุกรมเวลามากกว่า 1 รูปแบบให้พิจารณาค่า Mean Square Error (MSE) ของรูปแบบโดยเลือกรูปแบบที่ให้ค่า MSE น้อยที่สุด ตัวอย่างเช่น

Number of observations: 55

Residuals: SS = 1406051233 (backforecasts excluded)

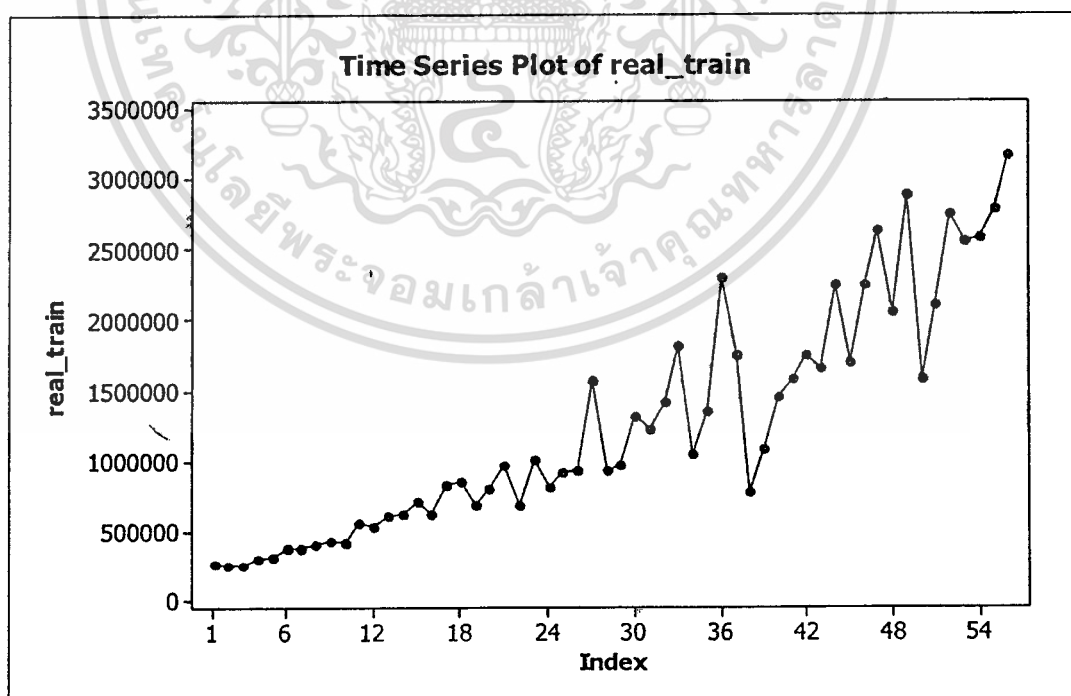
MS = 27039447 DF = 52

จากตัวอย่างเป็นการคำนวณของโปรแกรม Minitab ซึ่งจะคำนวณค่า MSE ให้ในที่นี้ค่า MS ก็คือค่า MSE นั้นเอง

2.8.6 ตัวอย่างการทำนายโดยวิธี ARIMA

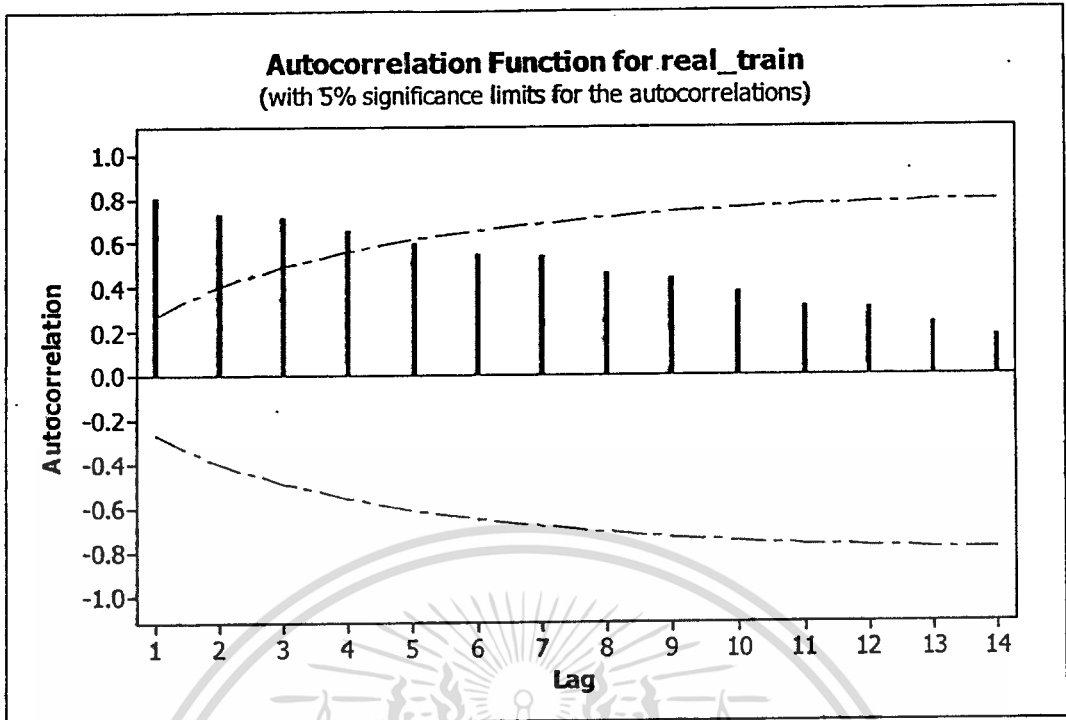
1. แบ่งข้อมูลอนุกรมเวลาออกเป็น 2 ส่วนคือส่วนสำหรับกำหนดรูปแบบและส่วนสำหรับทดสอบในอัตราส่วน 80 เปอร์เซ็นต์และ 20 เปอร์เซ็นต์ตามลำดับ

2. พิจารณาว่าข้อมูลเป็น stationary หรือไม่ โดยการเขียนกราฟดูลักษณะของข้อมูล

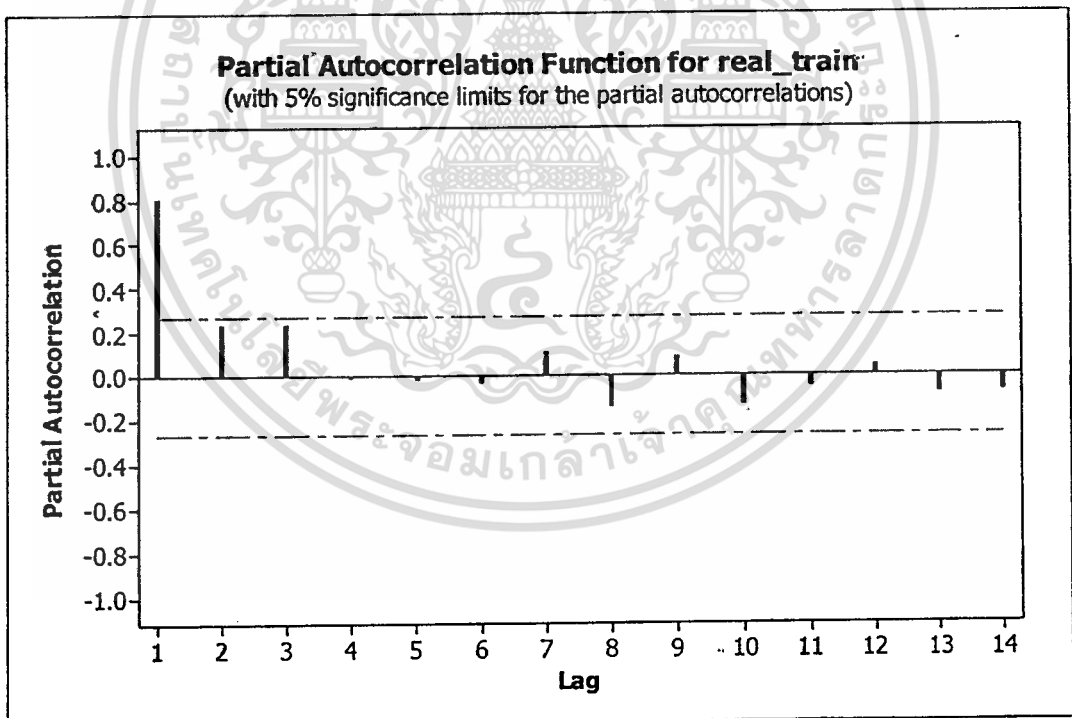


รูปที่ 2.19 แสดงลักษณะของข้อมูลจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.20 แสดง ACF ของข้อมูลจริง

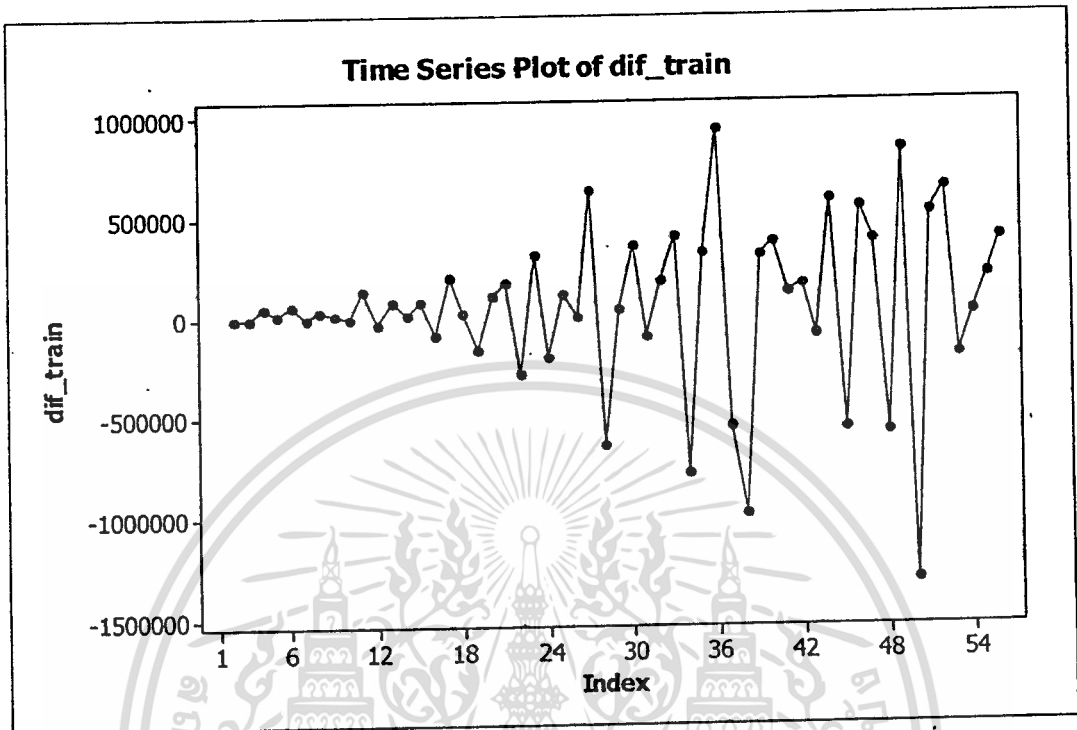


รูปที่ 2.21 แสดง PACF ของข้อมูลจริง

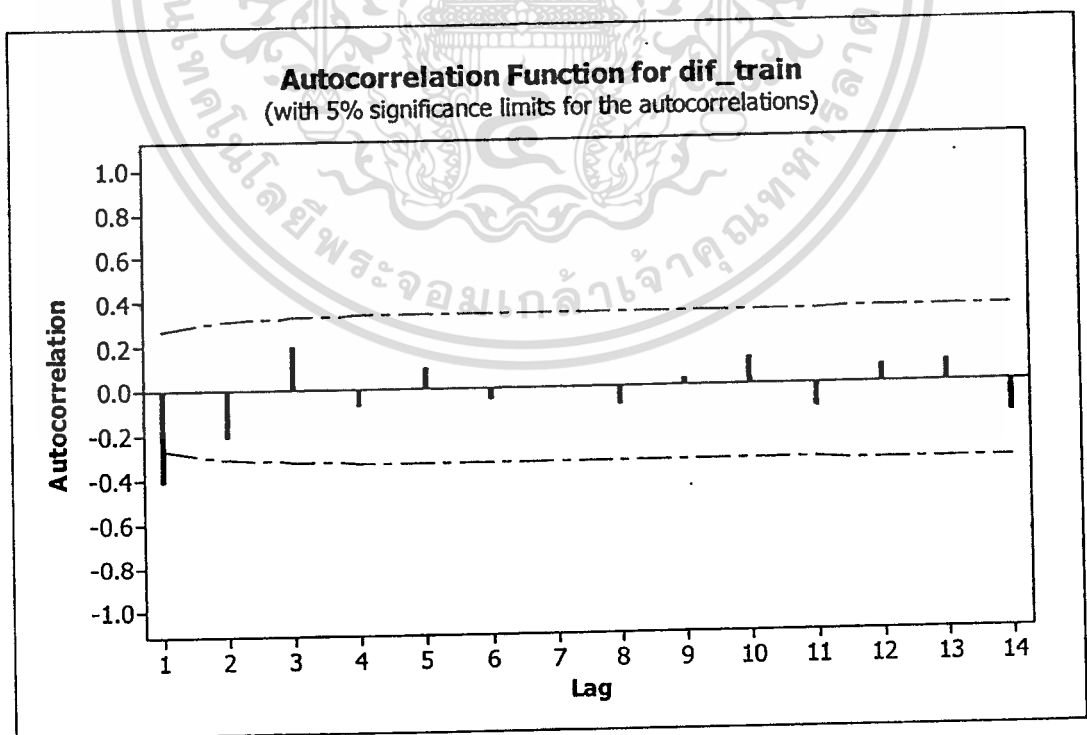
พิจารณารูปที่ 2.19 จะเห็นว่าข้อมูลมีแนวโน้มในรูปแบบชัดเจนและเมื่อพิจารณารูปที่ 2.20 จะเห็นว่าค่า ACF ค่อยๆ ลดค่าลงแต่ไม่ลดถึง 0 แสดงว่าข้อมูลมีแนวโน้มในรูปแบบ ดังนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลจึงยังไม่เป็น stationary เพื่อทำให้ข้อมูลเป็น stationary ให้ทำ first differencing เพื่อลบแนวโน้ม

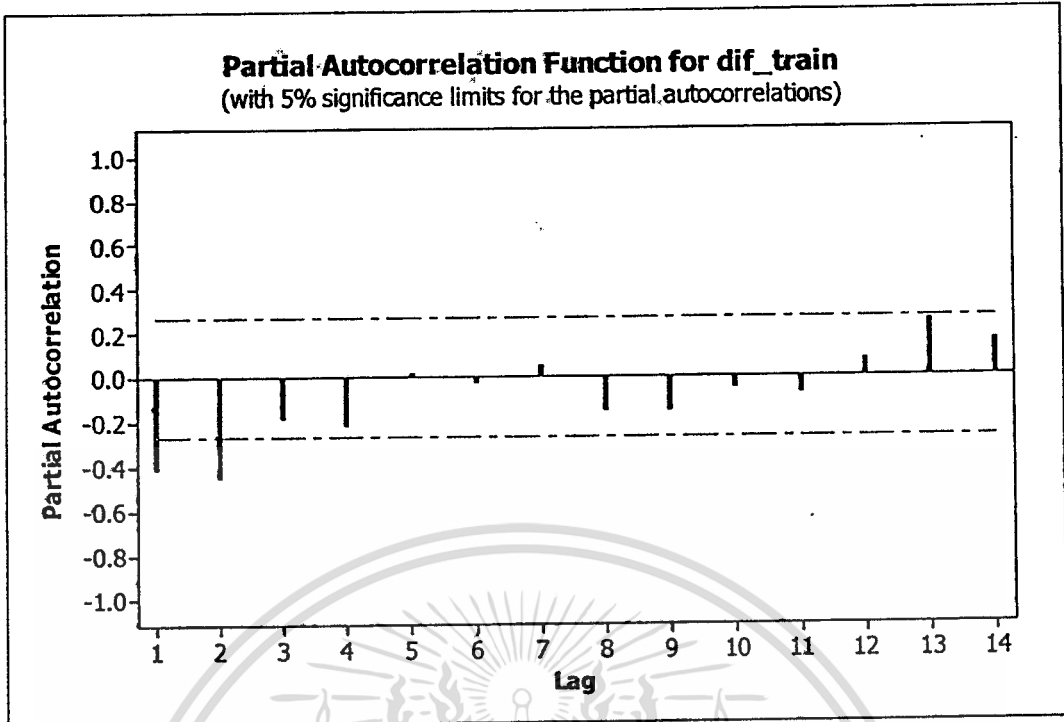


รูปที่ 2.22 แสดงลักษณะของข้อมูลหลังทำ first differencing



รูปที่ 2.23 ACF ของข้อมูลหลังทำ first differencing

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.24 PACF ของข้อมูลหลังทำ first differencing

รูปที่ 2.22 เป็นกราฟข้อมูลที่ผ่านการทำ first differencing จะเห็นว่าข้อมูลมีการเปลี่ยนแปลงรอบๆ ค่า 0 แสดงว่าข้อมูลเป็น stationary แล้ว จากนั้นให้ เขียนกราฟ ACF และ PACF เพื่อกำหนดรูปแบบ

3. กำหนดรูปแบบโดยพิจารณากราฟ ACF และ PACF ดูรูปที่ 2.23 และรูปที่ 2.24 ประกอบ จากการพิจารณากำหนดรูปแบบ ARIMA(2,1,0) และเมื่อใช้โปรแกรม Minitab คำณวนปรากฏผลดังนี้

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	-0.5660	0.1266	-4.47	0.000
AR 2	-0.4249	0.1269	-3.35	0.001

Number of observations: 55

Residuals: SS = 6609363350278 (backforecasts excluded)

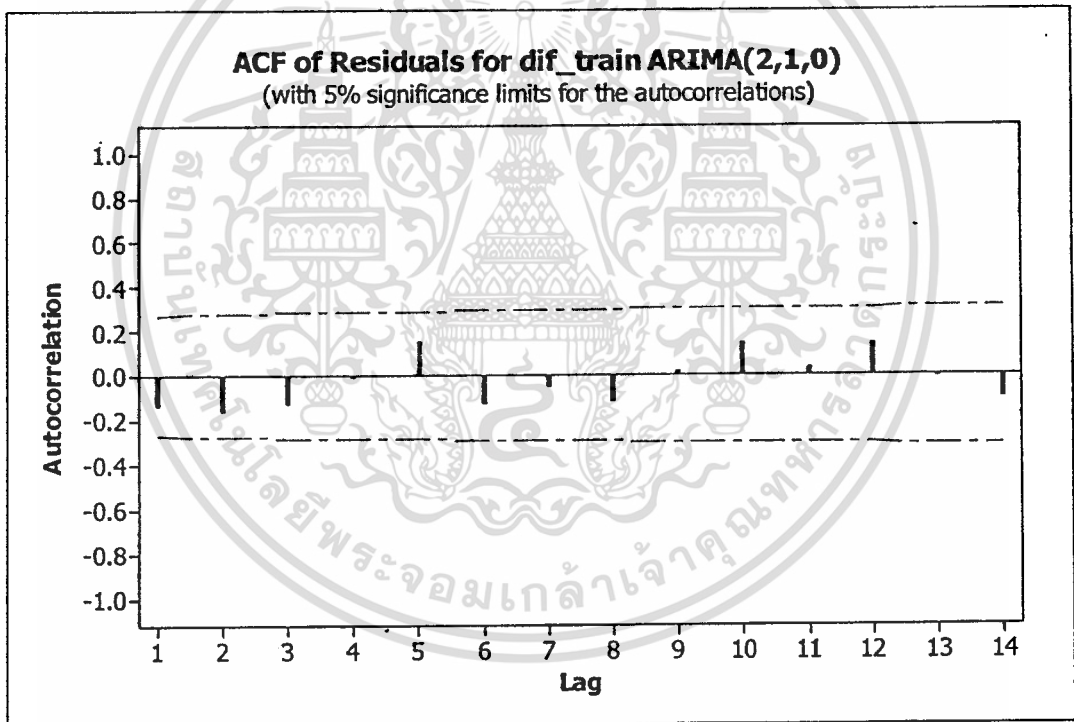
MS = 124704968873 DF = 53

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

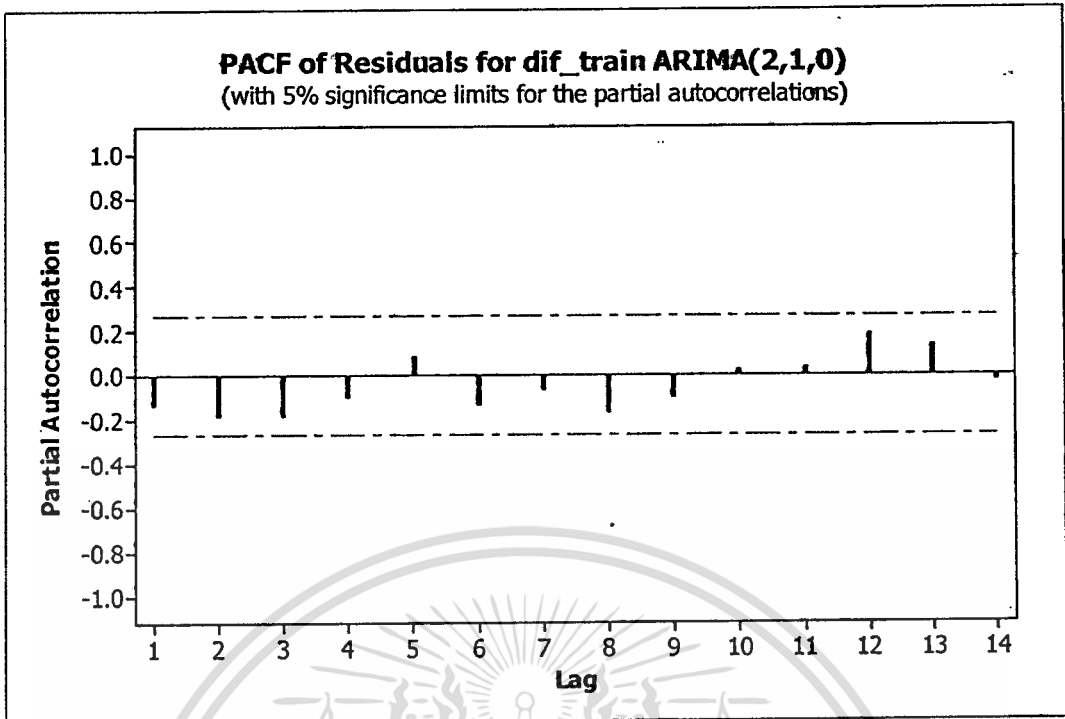
Lag	12	24	36	48
Chi-Square	9.9	17.9	20.3	21.0
DF	10	22	34	46
P-Value	0.450	0.712	0.969	0.999

3. พิจารณาความเหมาะสมของรูปแบบ

3.1 พิจารณากราฟ ACF และ PACF ของ residual จากรูปที่ 2.25 และรูปที่ 2.26 ประกอบปรากฏว่าไม่มี ACF และ PACF ที่ spike ออกนอกช่วงวิกฤติ (เส้นประ) ซึ่งหมายความว่า residuals ที่ lag ต่างๆ ไม่มีความสัมพันธ์กันจึงมีการเคลื่อนไหวที่เป็นอิสระกันหรือกล่าวได้ว่า residuals มีลักษณะที่เป็น random ซึ่งแสดงว่ารูปแบบที่กำหนดเหมาะสมแล้ว



รูปที่ 2.25 ACF ของ residuals



รูปที่ 2.26 PACF ของ residuals

3.2 พิจารณาค่า residuals ที่ lag ใดๆ ว่ามีความสัมพันธ์กันหรือไม่โดยการพิจารณาว่า r_k ของ residuals ที่ lag ใดๆ มีค่าเป็น 0 หรือไม่ ด้วยการพิจารณาดัชนีทดสอบสถิติของ Box และ Ljung โดยจะพิจารณาค่า P-Value ของค่า Chi-Square ปรากฏว่าค่า P-Value ที่ lag ใดๆ มีค่ามากกว่า α -level ที่ 0.05 ทุกค่า ดังนั้นแสดงว่า r_k ที่ lag ใดๆ มีค่าเท่ากับ 0 ซึ่งหมายความว่า residuals ไม่มีความสัมพันธ์กันที่ lag ใดๆ หรือกล่าวอีกนัยได้ว่า residuals ไม่มีความสัมพันธ์กันที่ lag ใดๆ

3.3 พิจารณาค่าพารามิเตอร์ว่าเป็น 0 หรือไม่ โดยการพิจารณาค่า T-Value และค่า P-Value ของค่าพารามิเตอร์แต่ละตัว จะเห็นว่าค่า P-Value ของทุกพารามิเตอร์มีค่าต่ำกว่า α -level ที่ 0.05 ทุกค่า ดังนั้นแสดงว่าค่าพารามิเตอร์มีค่าห่างจาก 0 อย่างมีนัยสำคัญ

3.4 ค่า MSE ของรูปแบบไม่ได้พิจารณา เนื่องจากกำหนดรูปแบบเดียวไม่มีรูปแบบอื่นมาเปรียบเทียบ

จากการพิจารณาความเหมาะสมของรูปแบบจะเห็นว่า ARIMA(2,1,0) เหมาะสมกับอนุกรมเวลาแล้ว

4. นำรูปแบบที่ได้ไปทำนายข้อมูลทดสอบซึ่งสามารถเขียนรูปแบบได้ดังนี้คือ

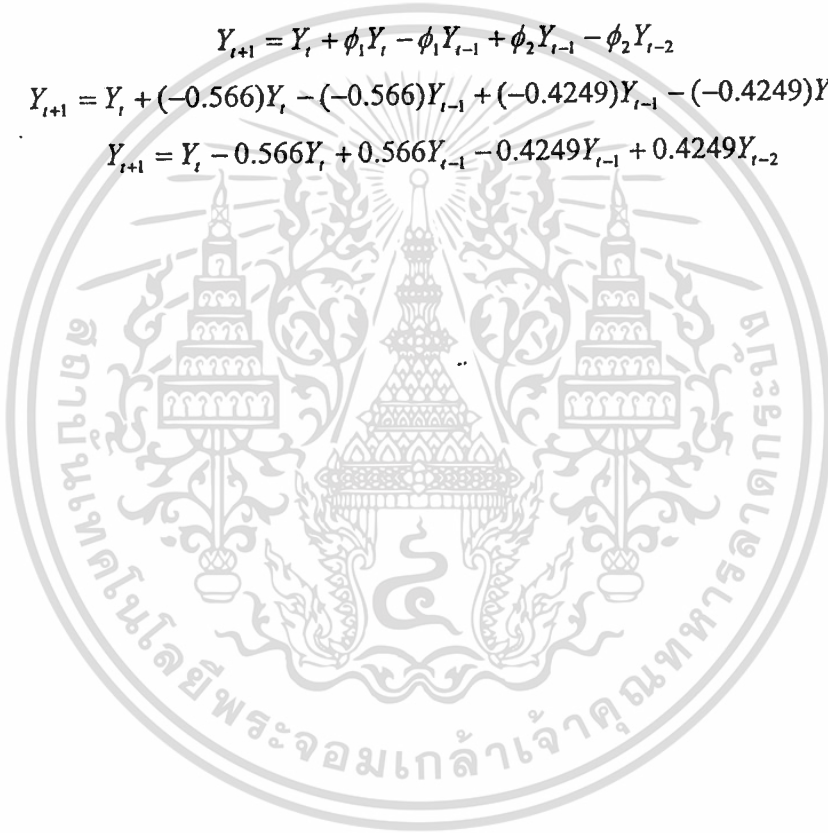
$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)Y_t = e_t \quad (2.41)$$

เนื่องจาก error ที่เวลา t ไม่ทราบค่า ดังนั้นจึงกำหนดให้เป็น 0 ดังนั้นจึงสามารถกระจายและเขียนสมการใหม่ได้เป็น

$$\begin{aligned}(1 - B - \phi_1 B + \phi_1 B^2 - \phi_2 B^2 + \phi_2 B^3)Y_t &= 0 \\ Y_t - Y_{t-1} - \phi_1 Y_{t-1} + \phi_1 Y_{t-2} - \phi_2 Y_{t-2} + \phi_2 Y_{t-3} &= 0 \\ Y_t &= Y_{t-1} + \phi_1 Y_{t-1} - \phi_1 Y_{t-2} + \phi_2 Y_{t-2} - \phi_2 Y_{t-3}\end{aligned}$$

ดังนั้นสมการที่ใช้ทำนายคือ

$$\begin{aligned}Y_{t+1} &= Y_t + \phi_1 Y_t - \phi_1 Y_{t-1} + \phi_2 Y_{t-1} - \phi_2 Y_{t-2} \\ Y_{t+1} &= Y_t + (-0.566)Y_t - (-0.566)Y_{t-1} + (-0.4249)Y_{t-1} - (-0.4249)Y_{t-2} \\ Y_{t+1} &= Y_t - 0.566Y_t + 0.566Y_{t-1} - 0.4249Y_{t-1} + 0.4249Y_{t-2}\end{aligned}$$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การทำนายข้อมูลอนุกรมเวลาโดยโมเดลแบบผสม

เนื่องจากข้อมูลอนุกรมเวลามีลักษณะที่มีการเปลี่ยนแปลงขึ้นลง ถ้าหากสามารถทำนายทิศทางการขึ้นหรือลงของอนุกรมเวลาได้ถูกต้องก็น่าจะทำนายค่าได้ถูกต้องมากขึ้นด้วย จากแนวคิดดังกล่าวจึงได้เสนอโมเดลแบบผสมซึ่งได้จากการนำโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นมาประกอบกัน โดยโมเดลที่นำเสนอนี้แบ่งการทำงานออกเป็น 2 ส่วน คือส่วนทำนายทิศทางของอนุกรมเวลาใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพและส่วนทำนายค่าในอนาคตใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น ก่อนจะอธิบายถึงสถาปัตยกรรมและการเรียนรู้ของโมเดลที่นำเสนอจะอธิบายการแสดงรูปแบบข้อมูลอนุกรมเวลาและการนอร์มัลไลซ์ข้อมูลก่อน

3.1 การแสดงรูปแบบข้อมูลอนุกรมเวลา

สมมติข้อมูลอนุกรมเวลาถูกแสดงในรูปของเวกเตอร์เวลา $Y = \{Y_1, Y_2, Y_3, \dots, Y_{t-1}, Y_t\}$ ข้อมูลอนุกรมเวลาจะถูกแปลงให้อยู่ในรูปของ segment ก่อนที่จะถูกส่งเข้าไปยังโครงข่ายประสาทเทียม โดยที่ segment คือขนาดการเปลี่ยนแปลงของข้อมูลจากเวลาที่ $t-1$ ไปยังเวลาที่ t เช่น $Y_t - Y_{t-1}$, $Y_{t-1} - Y_{t-2}$ หรืออาจจะแสดงในรูปของเวกเตอร์ที่ได้จากลบกันของข้อมูลอนุกรมเวลาที่อยู่ติดกัน $\delta = (\delta_1, \delta_2, \delta_3, \dots, \delta_{n-1})$ โดยที่ $\delta_i = Y_{i+1} - Y_i$ และ $1 \leq i \leq t-1$

3.1.1 การนอร์มัลไลซ์ข้อมูลอนุกรมเวลา

การนอร์มัลไลซ์ข้อมูลอนุกรมเวลา (Time Series Normalization) คือ การนำข้อมูลมาปรับค่าให้มีค่าอยู่ระหว่าง 0 ถึง 1 เพื่อให้สามารถใช้งานกับอัลกอริทึมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นได้ ก่อนการนอร์มัลไลเซชันจะปรับค่าสูงสุดของชุดข้อมูลเพิ่มอีก 5% ของค่าสูงสุดของข้อมูลอนุกรมเวลา และปรับค่าต่ำสุดของชุดข้อมูลลด 5% ของค่าต่ำสุดของข้อมูลอนุกรมเวลา สาเหตุที่เพิ่มและลด 5% ของข้อมูลกระทำบนสมมติฐานที่ว่าข้อมูลในอนาคตอาจจะมีการเปลี่ยนแปลงในทิศทางที่เพิ่มขึ้นหรือลดลงแตกต่างจากข้อมูลอนุกรมเวลาในปัจจุบัน จากนั้นข้อมูลอนุกรมเวลาที่เป็นข้อมูลจริงจะถูกนำมาคำนวณหาค่า segment ก่อน แล้วจึงนำข้อมูลที่อยู่ในรูป segment มานอร์มัลไลเซชันตามเงื่อนไขที่ 3.1

$$A_i = \frac{\delta_i - \min}{\max - \min} \quad (3.1)$$

โดยที่ δ คือ เวกเตอร์ segment ของข้อมูลอนุกรมเวลาที่ต้องการทำนอร์มัลไลเซชัน
 max คือ ค่าสูงสุดของเวกเตอร์ δ
 min คือ ค่าต่ำสุดของเวกเตอร์ δ

ตัวอย่างการนอร์มัลไลซ์ข้อมูล

ตารางที่ 3.1 การนอร์มัลไลซ์ข้อมูลอนุกรมเวลา

ลำดับ	Data	Segment	Normalized Data	Direction
1	3619.8	3846.9	0.6735	up
2	7466.7	3112.3	0.6517	up
3	10579	-3209.9	0.4646	down
4	7369.1	7618.9	0.7851	up
5	14988	-6589.5	0.3646	down
6	8398.5	-1144.2	0.5258	down
7	7254.3	1038.3	0.5904	up
8	8292.6	5365.4	0.7184	up
9	13658	-1978	0.5011	down
10	11680	0	0.5904	NC
11	11680			

จากตาราง อธิบายได้ดังนี้

Data คืออนุกรมเวลา

Segment คือค่าที่ได้จากการลบกับของค่าในอนุกรมเวลาที่อยู่ลำดับติดกัน โดยที่ segment ลำดับที่ 1 ได้มาจากนำ data ลำดับที่ 2 ลบ data ลำดับที่ 1 เป็นต้น

Normalized Data คือการแปลงค่า segment ที่ได้ให้อยู่ในช่วงระหว่าง 0 ถึง 1 โดยใช้
 สูตรที่ 3.1

Direction ของคำตอบพิจารณาจากค่า segment นั่นคือถ้าค่า segment มีค่ามากกว่า 0 ถือว่าคำตอบมีทิศทางขึ้น (Up) ถ้าค่า segment มีค่าเท่ากับ 0 ถือว่าคำตอบมีทิศทางไม่เปลี่ยนแปลง (NC) และถ้าค่า segment มีค่าน้อยกว่า 0 ถือว่าคำตอบมีทิศทางลง (Down)

3.1.2 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพ

จากตารางที่ 3.1 หลังจากนอร์มัลไลซ์ข้อมูลเสร็จแล้วก็นำมาจัดเป็นแพทเทิน จากตัวอย่าง ใช้ขนาดแพทเทินอินพุทเท่ากับ 4 และทำนาย 1 step ข้างหน้าโดย output pattern คือ ทิศทาง ซึ่งจะแทนทิศทางขึ้นด้วย 1 0 0 ทิศทางลงแทนด้วย 0 0 1 และทิศทางไม่เปลี่ยนแปลงแทนด้วย 0 1 0 ดังนั้น แพทเทินอินพุทและแพทเทินเอาต์พุทจะเป็นดังนี้

ตารางที่ 3.2 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ SFAM ก่อนทำ complement coding

ลำดับ	Input Pattern	Output Pattern
1	0.6735 0.6517 0.4646 0.7851	0 0 1
2	0.6517 0.4646 0.7851 0.3646	0 0 1
3	0.4646 0.7851 0.3646 0.5258	1 0 0
4	0.7851 0.3646 0.5258 0.5904	1 0 0
5	0.3646 0.5258 0.5904 0.7184	0 0 1
6	0.5258 0.5904 0.7184 0.5011	0 1 0

แพทเทินอินพุทสำหรับโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพนี้จะต้องผ่านการทำ complement coding ก่อน ดังนั้น ขนาดของแพทเทินอินพุทจะมีขนาดเป็น 2 เท่าของขนาดแพทเทินอินพุทเดิม ดังตารางที่ 3.3

ตารางที่ 3.3 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ SFAM หลังทำ complement coding

ลำดับ	Input Pattern	Output Pattern
1	0.6735 0.6517 0.4646 0.7851 0.3265 0.3483 0.5354 0.2149	0 0 1
2	0.6517 0.4646 0.7851 0.3646 0.3483 0.5354 0.2149 0.6354	0 0 1
3	0.4646 0.7851 0.3646 0.5258 0.5354 0.2149 0.6354 0.4742	1 0 0
4	0.7851 0.3646 0.5258 0.5904 0.2149 0.6354 0.4742 0.4096	1 0 0
5	0.3646 0.5258 0.5904 0.7184 0.6354 0.4742 0.4096 0.2816	0 0 1
6	0.5258 0.5904 0.7184 0.5011 0.4742 0.4096 0.2816 0.4989	0 1 0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.3 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

แพทเทินอินพุทและแพทเทินเอาต์พุทที่ถูกส่งเข้าโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นจะถูกแบ่งตามทิศทางของคำตอบ เพื่อส่งเข้าโครงข่าย 2 โครงข่ายคือโครงข่ายหนึ่งสำหรับเรียนรู้และทำนายค่าแพทเทินอินพุทที่คำตอบมีทิศทางขึ้น และอีกโครงข่ายหนึ่งสำหรับเรียนรู้และทำนายค่าแพทเทินอินพุทที่คำตอบมีทิศทางลง แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นเป็นดังนี้ คือ

ตารางที่ 3.4 แพทเทินอินพุทและแพทเทินเอาต์พุทสำหรับ BPNN

ลำดับ	Input Pattern	Output Pattern	ทิศทางของคำตอบ
1	0.6735 0.6517 0.4646 0.7851	0.3646	down
2	0.6517 0.4646 0.7851 0.3646	0.5258	down
3	0.4646 0.7851 0.3646 0.5258	0.5904	up
4	0.7851 0.3646 0.5258 0.5904	0.7184	up
5	0.3646 0.5258 0.5904 0.7184	0.5011	down
6	0.5258 0.5904 0.7184 0.5011	0.5904	nc

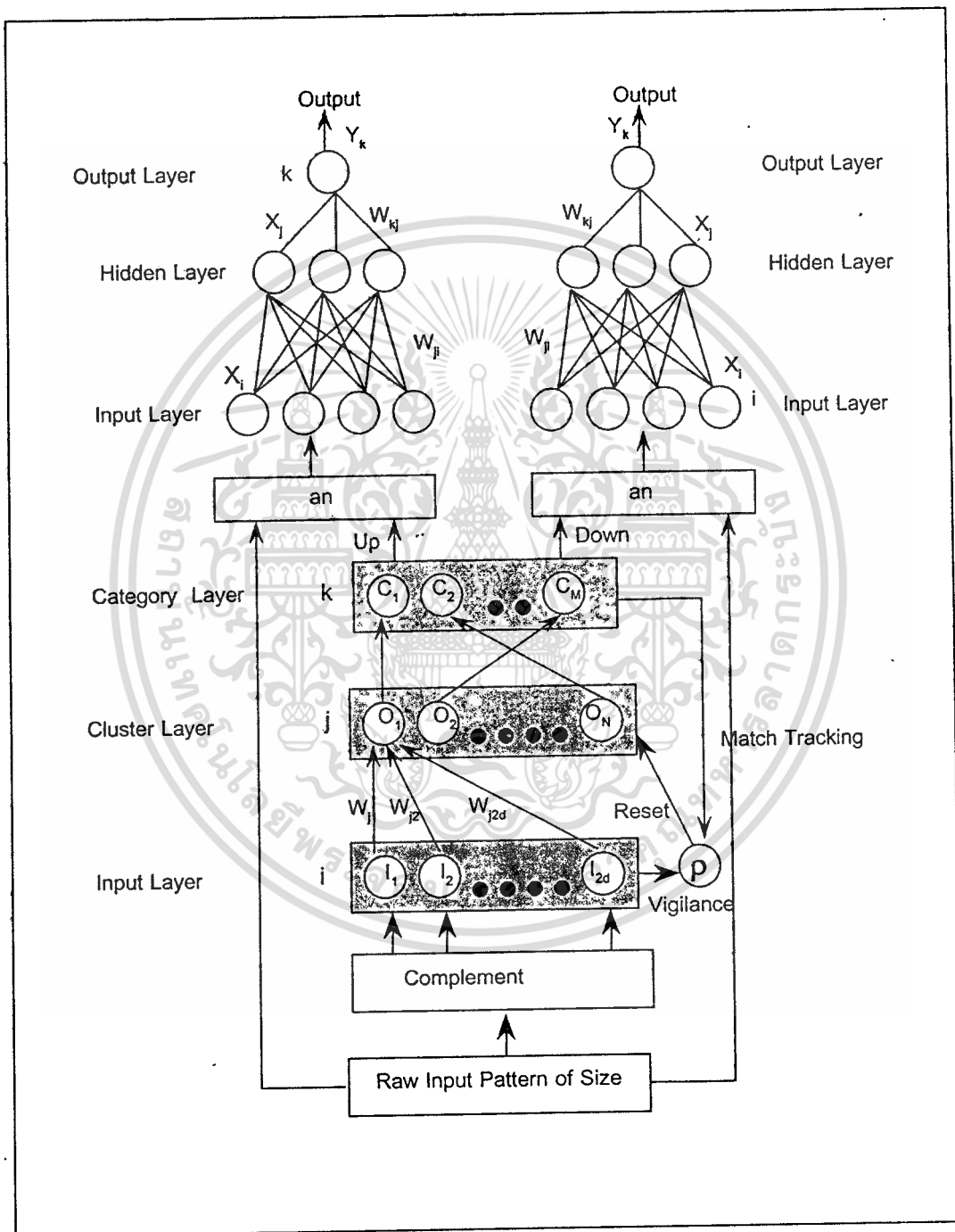
จากตารางที่ 3.4 แพทเทินที่ถูกส่งเข้าโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นสำหรับเรียนรู้และทำนายค่าแพทเทินอินพุทที่คำตอบมีทิศทางขึ้นคือ แพทเทินลำดับที่ 3 และ 4 ส่วนแพทเทินที่ถูกส่งเข้าโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นสำหรับเรียนรู้และทำนายค่าแพทเทินอินพุทที่คำตอบมีทิศทางลงคือ แพทเทินลำดับที่ 1, 2 และ 5

3.2 สถาปัตยกรรมของโมเดลแบบผสม

สถาปัตยกรรมของโมเดลแบบผสมสำหรับทำนายข้อมูลอนุกรมเวลานี้ประกอบด้วย 2 ส่วนคือส่วนทำนายทิศทางของอนุกรมเวลาใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพ ส่วนที่สองที่ใช้ทำนายค่าของการเปลี่ยนแปลงใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น ดังรูปที่ 3.1

ส่วนทำนายทิศทางของแพทเทินอินพุทที่ใช้สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพประกอบด้วย 3 ชั้นคือชั้นอินพุท (Input layer) ชั้นคลัสเตอร์ (Cluster Layer) และชั้นแคททิกอรี (Category Layer) ชั้นอินพุทจะรับแพทเทินอินพุทที่ผ่านการทำ complement coding ส่วนชั้น คลัสเตอร์ ก็คือกลุ่มของแพทเทินอินพุทที่ถูกจัดกลุ่มเข้าไว้ด้วยกัน

ทั้งสองชั้นนี้จะเชื่อมต่อแบบถึงกันหมดด้วยเวกเตอร์น้ำหนัก (W_{ji}) โหนด 1 โหนดในชั้นคลัสเตอร์ จะชี้ไปยังโหนดในชั้นแคททิทอรีที่เป็นคำตอบที่สัมพันธ์กับแพทเทินอินพุต โครงข่ายประสาทเทียม แบบซิมพลิไฟด์พีชชีอาทแมพจะจัดกลุ่มแพทเทินอินพุตตามคำตอบของมันซึ่งคำตอบแยกได้เป็น 3 กลุ่มคือคำตอบที่มีทิศทางขึ้น คำตอบที่มีทิศทางลงและคำตอบที่ทิศทางไม่เปลี่ยนแปลง



รูปที่ 3.1 สถาปัตยกรรมของโมเดลแบบผสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนทำนายค่าซึ่งเป็นส่วนที่ต่อจากส่วนทำนายทิศทางใช้สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันซึ่งประกอบด้วย 3 ชั้นคือ ชั้นอินพุท ชั้นฮิดเดนและชั้นเอาต์พุท โหนดทุกโหนดในทุกชั้นจะเชื่อมต่อแบบถึงกันหมดโดยเวกเตอร์น้ำหนัก ชั้นอินพุทเป็นชั้นที่รับแพทเทินอินพุทและส่งผ่านข้อมูลไปจนผลลัพธ์ออกไปที่ชั้นเอาต์พุท แพทเทินอินพุทที่ป้อนเข้าโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันไม่ต้องทำ complement coding ในการทำนายค่านี้จะใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน 2 โครงข่ายโดยโครงข่ายหนึ่งใช้ทำนายค่าของแพทเทินอินพุทที่คำตอบถูกทำนายว่ามีทิศทางขึ้นและอีกโครงข่ายหนึ่งใช้ทำนายค่าของแพทเทินอินพุทที่คำตอบถูกทำนายว่ามีทิศทางลง ส่วนแพทเทินอินพุทที่ถูกทำนายว่าคำตอบจะไม่มี การเปลี่ยนแปลงทิศทางจะใช้ค่าก่อนจุดที่ทำนายในข้อมูลอนุกรมเวลาเป็นคำตอบ

3.3 การเรียนรู้ของโมเดลแบบผสม

ในขั้นตอนการเรียนรู้ของโมเดลใหม่ที่น่าเสนอนี้จะแบ่งการทำงานออกเป็น 2 ส่วนคือ ส่วนแรกใช้เรียนรู้เพื่อจัดกลุ่มของแพทเทินอินพุทตามกลุ่มคำตอบที่สัมพันธ์กับแพทเทินอินพุทนั้น ซึ่งกลุ่มคำตอบที่ต้องการให้ระบบเรียนรู้มี 3 กลุ่มคำตอบคือกลุ่มที่มีทิศทางขึ้น ทิศทางลง และทิศทางไม่เปลี่ยนแปลง ส่วนการเรียนรู้ของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันที่ใช้ทำนายค่านั้นมี 2 โครงข่ายคือ โครงข่ายหนึ่งใช้เรียนรู้แพทเทินอินพุทที่คำตอบของมันมีทิศทางขึ้นเพียงอย่างเดียวและอีกโครงข่ายให้เรียนรู้แพทเทินอินพุทที่คำตอบของมันมีทิศทางลงเพียงอย่างเดียว ส่วนแพทเทินอินพุทที่คำตอบมีทิศทางไม่เปลี่ยนแปลงไม่ต้องทำอะไรเพราะจะใช้ค่าก่อนจุดที่จะทำนายในอนุกรมเวลาเป็นคำตอบเลย

3.3.1 การเรียนรู้ส่วนทำนายทิศทาง

กระบวนการเรียนรู้ในส่วนทำนายทิศทางนี้ใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชีอาทแมพในบทที่ 2 หัวข้อที่ 2.6

3.3.2 การเรียนรู้ส่วนทำนายค่า

แพทเทินอินพุทของชุดข้อมูลที่ใช้ในการเรียนรู้จะถูกแบ่งตามทิศทางของคำตอบที่สัมพันธ์กับแพทเทินอินพุทนั้นซึ่งทิศทางคำตอบของแพทเทินอินพุทมี 3 ทิศทางคือคำตอบที่มีทิศทางขึ้น ทิศทางลงและทิศทางไม่เปลี่ยนแปลง ในส่วนนี้จะใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันโดยโครงข่ายหนึ่งเรียนรู้แพทเทินอินพุทที่คำตอบมีทิศทางขึ้นและอีกโครงข่ายหนึ่งเรียนรู้ แพทเทินอินพุทที่คำตอบมีทิศทางลง แพทเทินอินพุทที่ส่งเข้าโครงข่ายนี้ไม่ต้องทำ complement coding เหมือนแพทเทินอินพุทที่ส่งเข้าส่วนเรียนรู้และทำนายทิศทาง อัลกอริทึมการเรียนรู้ในส่วนนี้จะใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันดังอธิบายในบทที่ 2 หัวข้อที่

2.7 จำนวนนิรอนในชั้นอินพุตเท่ากับจำนวนค่าในเวกเตอร์แพทเทินอินพุต จำนวนนิรอนในชั้นเอาต์พุตเท่ากับ 1 นิรอนและจำนวนนิรอนในชั้นฮิดเดนได้จากการทดลอง โครงข่ายจะหยุดเรียนรู้เมื่อครบจำนวนรอบที่กำหนดหรือเมื่อค่าความผิดพลาดเพิ่มขึ้นหรือลดลงน้อยกว่าหรือเท่ากับค่าที่กำหนด

3.4 การทำนายโดยโมเดลแบบผสม

ในขั้นตอนการทำนายของโมเดลแบบผสมนี้จะแบ่งการทำงานออกเป็น 2 ส่วนคือส่วนแรกใช้ทำนายทิศทางของแพทเทินที่นำมาทดสอบและส่วนทำนายค่า เมื่อมีแพทเทินอินพุตป้อนเข้าระบบก็จะถูกทำนายทิศทางของคำตอบว่าคำตอบมีทิศทางขึ้น ทิศทางลงหรือทิศทางไม่เปลี่ยนแปลง ถ้าโมเดลตอบว่าแพทเทินนั้นคำตอบมีทิศทางขึ้นก็จะส่งไปทำนายค่าในโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นที่ใช้ทำนายแพทเทินที่คำตอบมีทิศทางขึ้น ถ้าโมเดลตอบว่าแพทเทินนั้นคำตอบมีทิศทางลงก็จะส่งไปทำนายค่าในโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นที่ใช้ทำนายแพทเทินที่คำตอบมีทิศทางลง ถ้าโมเดลตอบว่าแพทเทินนั้นคำตอบมีทิศทางไม่เปลี่ยนแปลงจะใช้ค่าก่อนจุดที่จะทำนายในอนุกรมเวลาเป็นคำตอบ

บทที่ 4

การทดลองและผลการทดลอง

4.1 ข้อมูลที่ใช้ในการทดลอง

ในการทดลองนี้ใช้ข้อมูลจำนวน 2 กลุ่มดังนี้

1. ข้อมูลกลุ่มที่ 1 เป็นข้อมูลที่สร้างจากสมการประกอบด้วย 2 ชุดข้อมูล

1.1 ข้อมูลชุดที่ 1 เป็นข้อมูลที่ได้สร้างมาจากสมการที่ 4.1 จำนวนข้อมูลมีทั้งหมด 321 จุด ลักษณะของข้อมูลแสดงดังรูปที่ 4.1

$$y = \sin(x) , \quad 0 \leq x \leq 32 \quad (4.1)$$

1.2 ข้อมูลชุดที่ 2 เป็นข้อมูลที่สร้างมาจากสมการ Mackey-Glass Chaotic Time Series ดังสมการที่ 4.2 [13] เป็นอนุกรมเวลาที่มีลักษณะไม่เป็นเชิงเส้นที่นิยมใช้เป็นตัววัดประสิทธิภาพของระบบ จำนวนข้อมูลมีทั้งหมด 1500 จุด ลักษณะของข้อมูลแสดงดังรูปที่ 4.2

$$x(t+1) = x(t) + \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \quad (4.2)$$

โดยที่ $\tau = 17$

2. ข้อมูลกลุ่มที่ 2 เป็นข้อมูลจริงมีทั้งหมดจำนวน 8 ชุดข้อมูล แต่ละชุดข้อมูลมีลักษณะดังนี้

2.1 ข้อมูลชุดที่ 3 เป็นข้อมูลรายเดือนของอุตสาหกรรมรถยนต์ แสดงจำนวนรถยนต์ รุ่น Panter ที่ผลิตตั้งแต่เดือนมกราคม ค.ศ.1982 ถึงเดือนตุลาคม ค.ศ.1987 จำนวนข้อมูลทั้งหมด 70 จุด ช่วงของข้อมูลอยู่ระหว่าง 3,619.8 – 52,727.2 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.3

2.2 ข้อมูลชุดที่ 4 เป็นข้อมูลราย 3 เดือนแสดงผลผลิตภัณฑ์มวลรวมแห่งชาติ ตั้งแต่ไตรมาสที่ 1 ของ ค.ศ. 1947 ถึงไตรมาสที่ 3 ของ ค.ศ. 1987 จำนวนข้อมูลทั้งหมด 163 จุด ช่วงของข้อมูลอยู่ระหว่าง 1,056.5 – 3,769.4 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.4

2.3 ข้อมูลชุดที่ 5 เป็นข้อมูลรายเดือนแสดงยอดขายสินค้า Honeywell ตั้งแต่เดือนมกราคม ค.ศ. 1981 ถึงเดือนตุลาคม ค.ศ. 1987 จำนวนข้อมูลทั้งหมด 82 จุด ช่วงของข้อมูลอยู่ระหว่าง 24,611 – 222,062 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.5

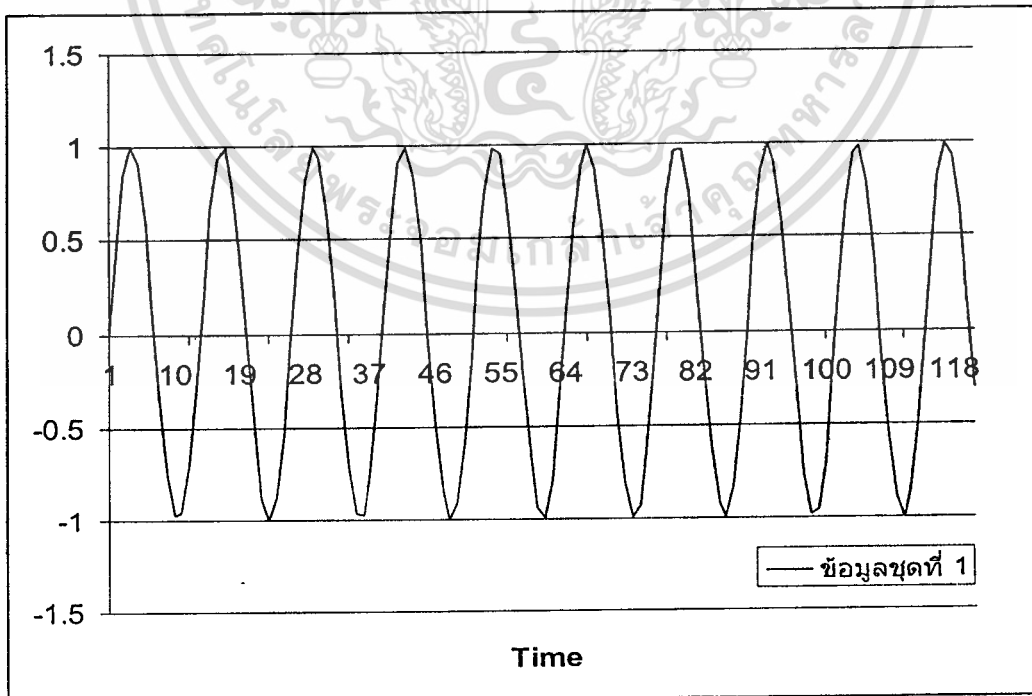
2.4 ข้อมูลชุดที่ 6 เป็นข้อมูลรายเดือนแสดงยอดขายสินค้า Squibb ตั้งแต่เดือนมกราคม ค.ศ. 1982 ถึงเดือนตุลาคม ค.ศ.1987 จำนวนข้อมูลทั้งหมด 70 จุด ช่วงของข้อมูลอยู่ระหว่าง 252,360 – 4,685,076 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.6

2.5 ข้อมูลชุดที่ 7 เป็นข้อมูลรายเดือนแสดงยอดสั่งซื้อกระดาษปกของยุโรป ตั้งแต่เดือนมกราคม ค.ศ. 1981 ถึงเดือนตุลาคม ค.ศ. 1987 จำนวนข้อมูล 82 จุด ช่วงของข้อมูลอยู่ระหว่าง 142,872 – 397,034 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.7

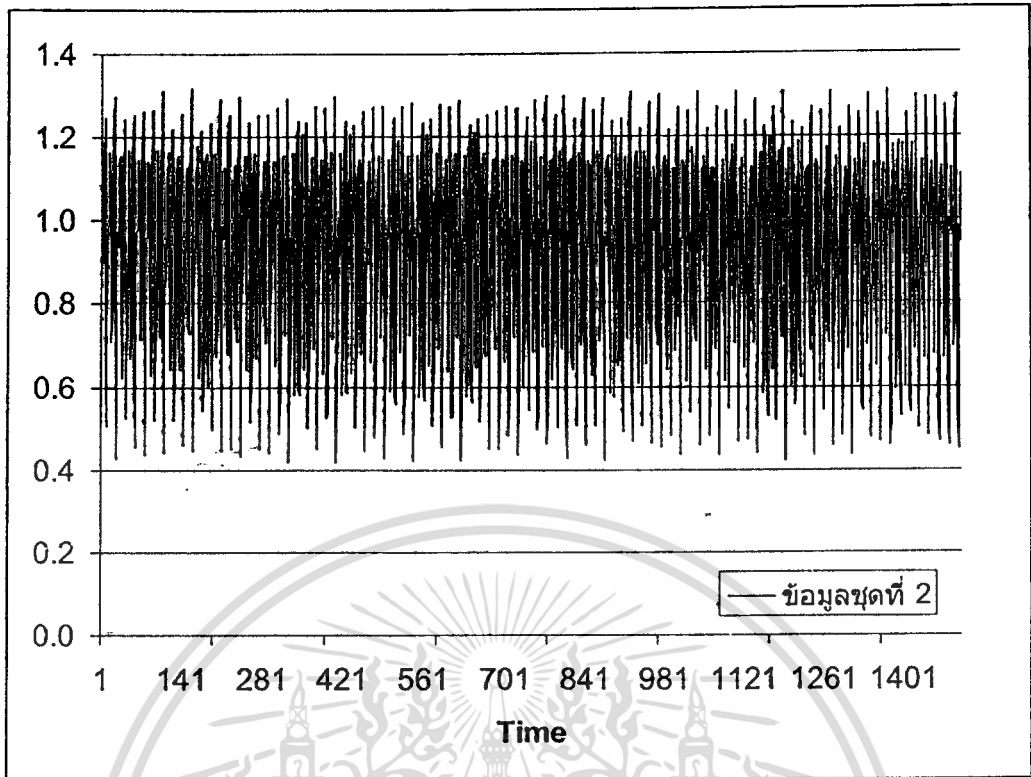
2.6 ข้อมูลชุดที่ 8 เป็นข้อมูลรายเดือนแสดงยอดสั่งซื้อกระดาษของประเทศสหรัฐอเมริกาตั้งแต่เดือนมกราคม ค.ศ. 1969 ถึงเดือนกันยายน ค.ศ. 1987 จำนวนข้อมูลทั้งหมด 225 จุด ช่วงของข้อมูลอยู่ระหว่าง 222,000 – 651,000 [14] ลักษณะของข้อมูลแสดงดังรูปที่ 4.8

2.7 ข้อมูลชุดที่ 9 เป็นข้อมูลรายเดือนแสดงการผลิตกระแสไฟฟ้าทั้งหมดของอุตสาหกรรมไฟฟ้าของประเทศสหรัฐอเมริกา ตั้งแต่เดือนมกราคม ค.ศ. 1985 ถึงเดือนตุลาคม ค.ศ. 1996 จำนวนข้อมูลทั้งหมด 142 จุด ช่วงของข้อมูลอยู่ระหว่าง 184.88 – 304.71 [6] ลักษณะของข้อมูลแสดงดังรูปที่ 4.9

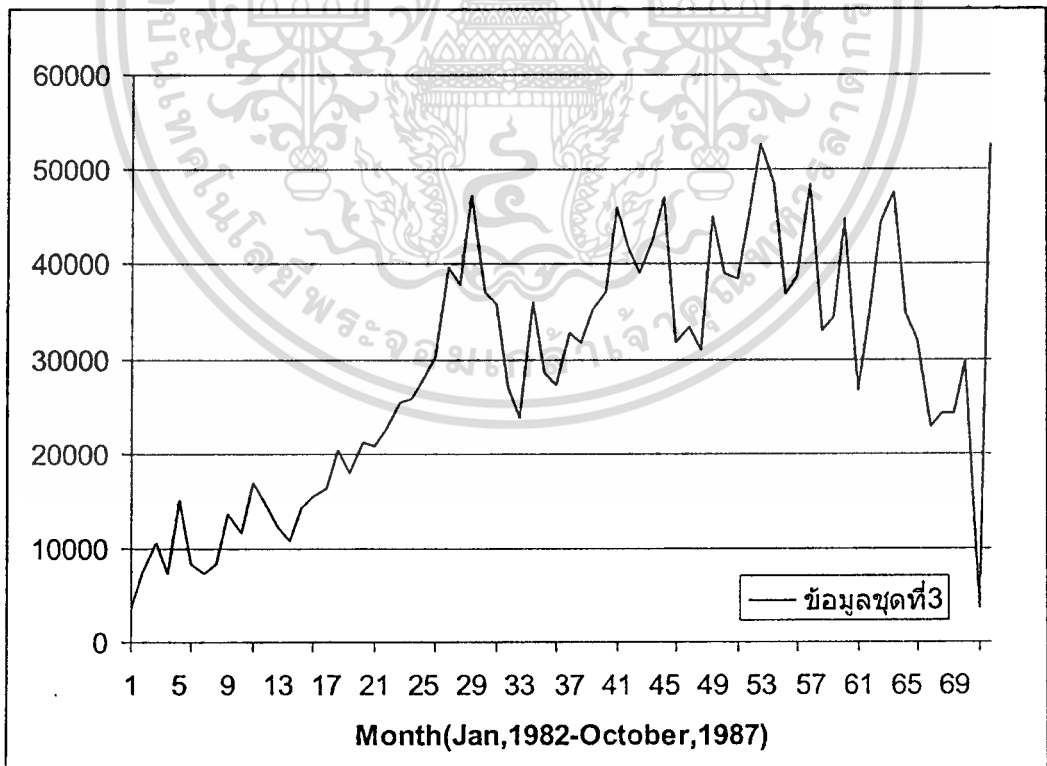
2.8 ข้อมูลชุดที่ 10 เป็นข้อมูลรายเดือนแสดงปริมาณการส่งออกปลาหูนาของประเทศไทย ตั้งแต่เดือนมกราคม พ.ศ. 2534 ถึงเดือนเมษายน พ.ศ. 2546 จำนวนข้อมูลทั้งหมด 148 จุด ช่วงของข้อมูลอยู่ระหว่าง 2,258,735 – 14,626,923 [15] ลักษณะของข้อมูลแสดงดังรูปที่ 4.10



รูปที่ 4.1 ลักษณะของข้อมูลชุดที่ 1

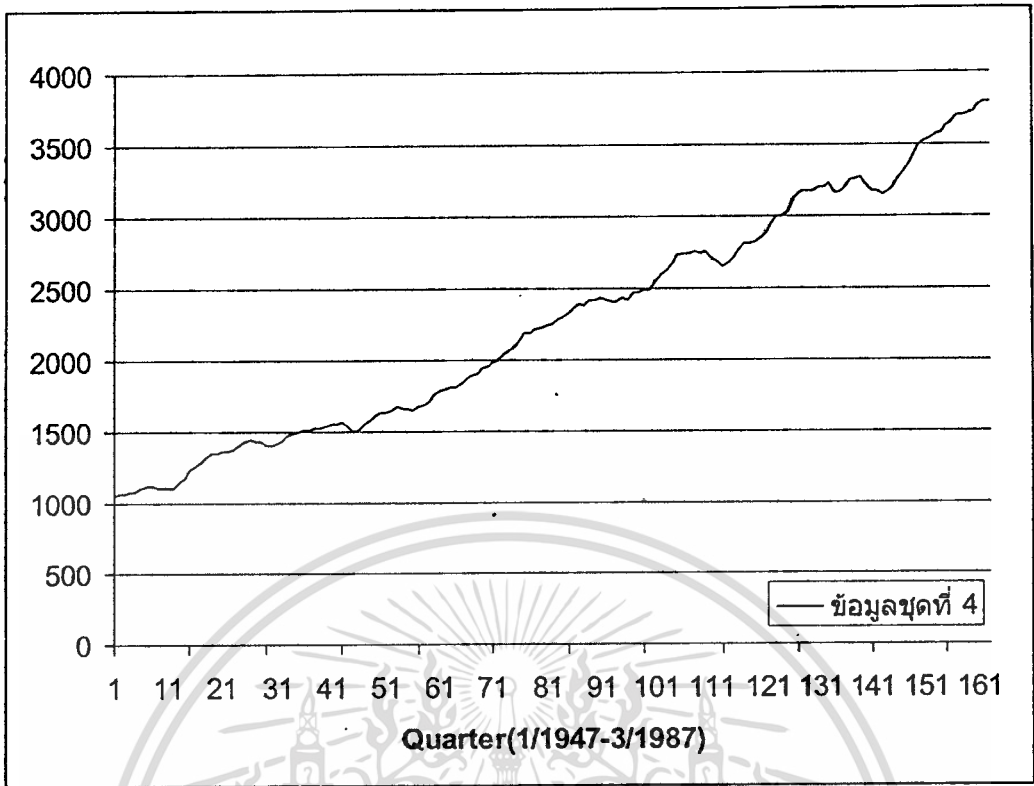


รูปที่ 4.2 ลักษณะของข้อมูลชุดที่ 2

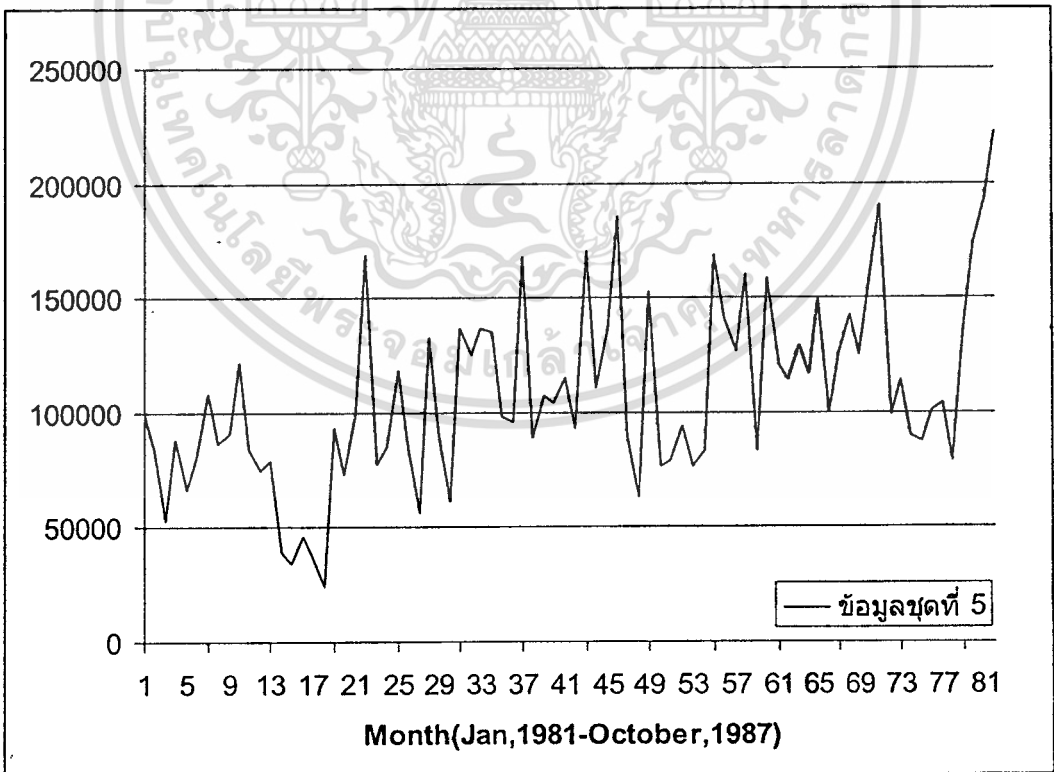


รูปที่ 4.3 ลักษณะของข้อมูลชุดที่ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

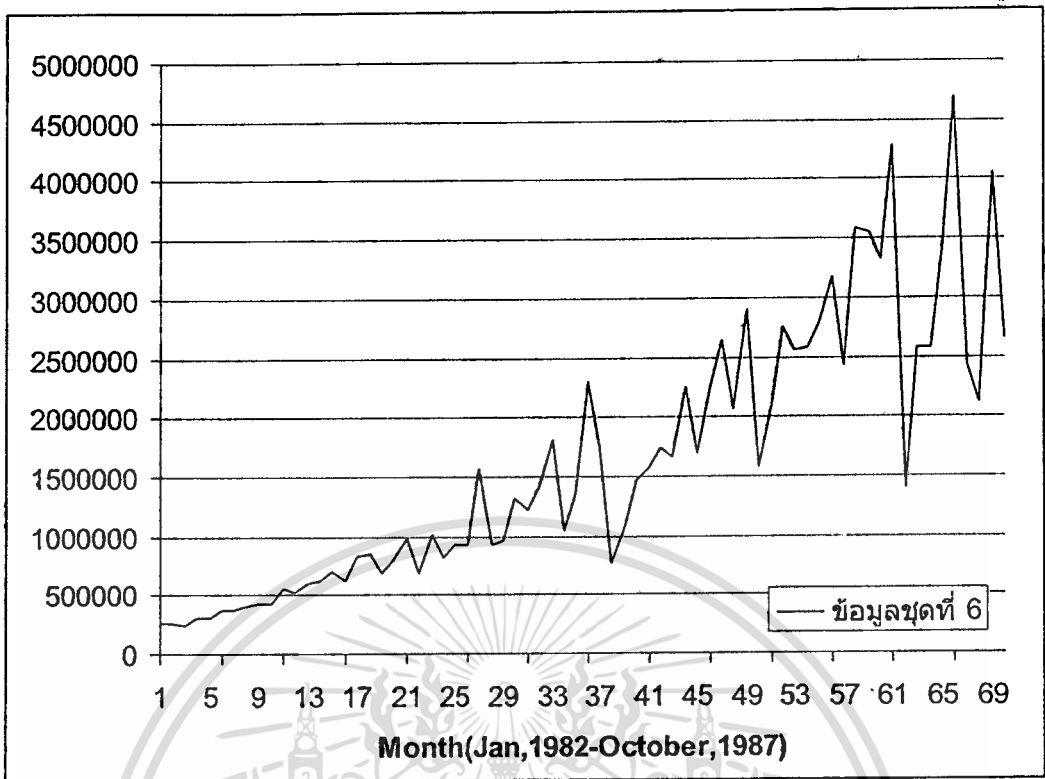


รูปที่ 4.4 ลักษณะของข้อมูลชุดที่ 4

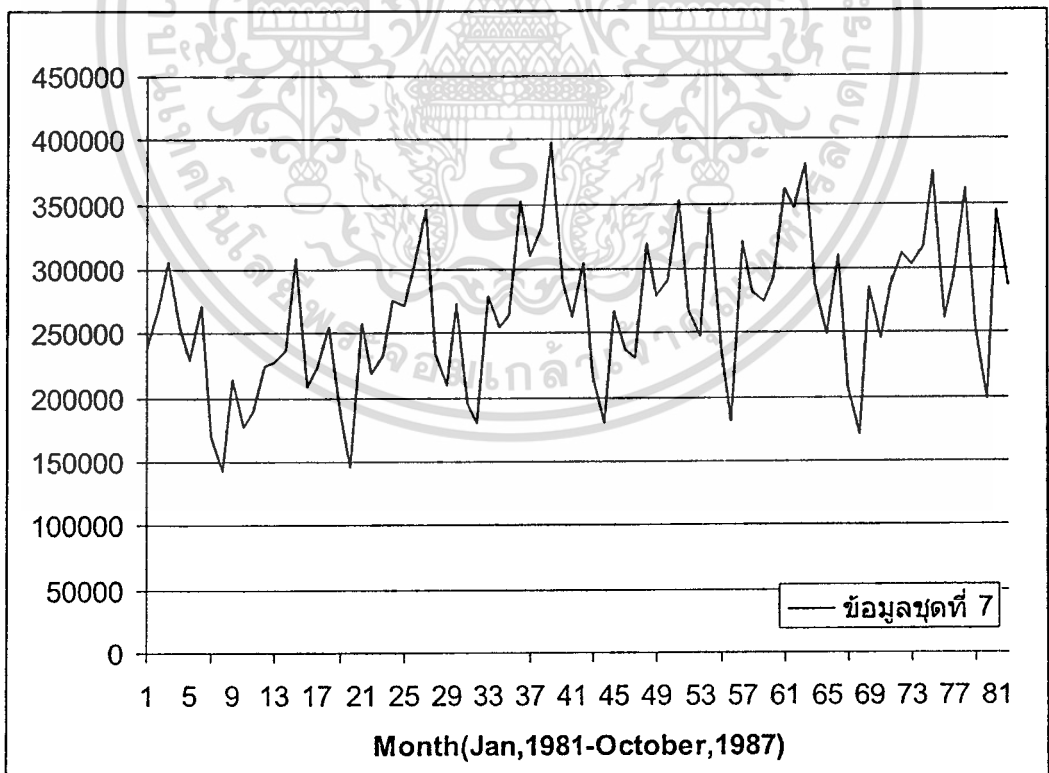


รูปที่ 4.5 ลักษณะของข้อมูลชุดที่ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

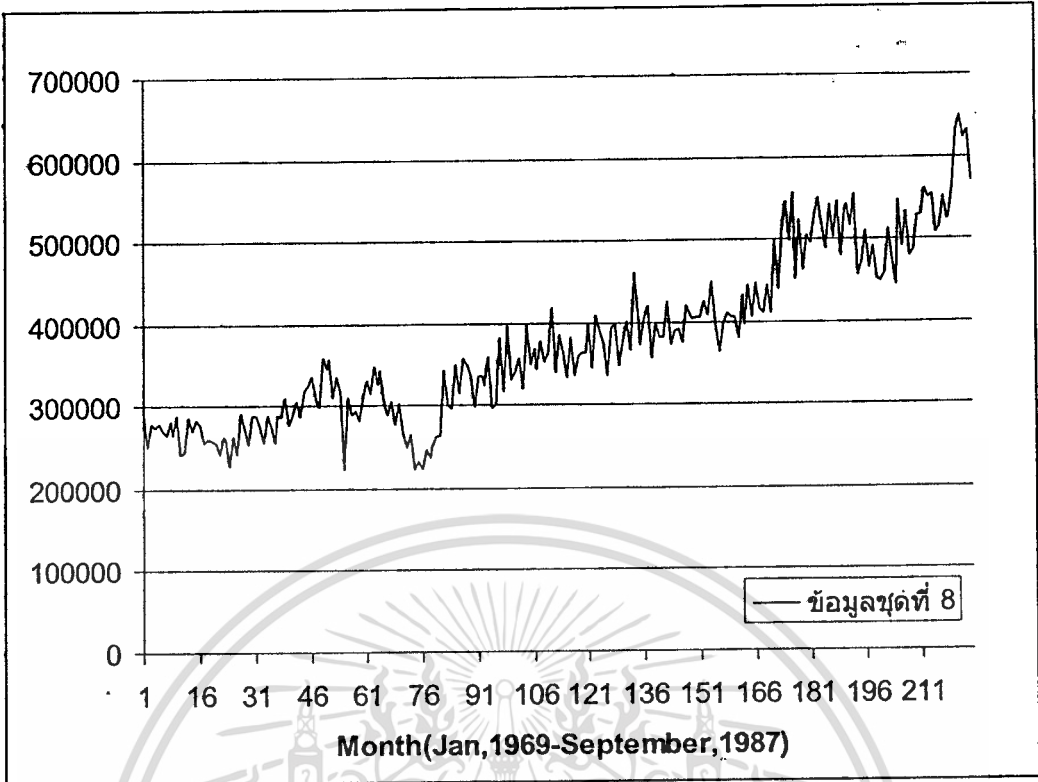


รูปที่ 4.6 ลักษณะของข้อมูลชุดที่ 6

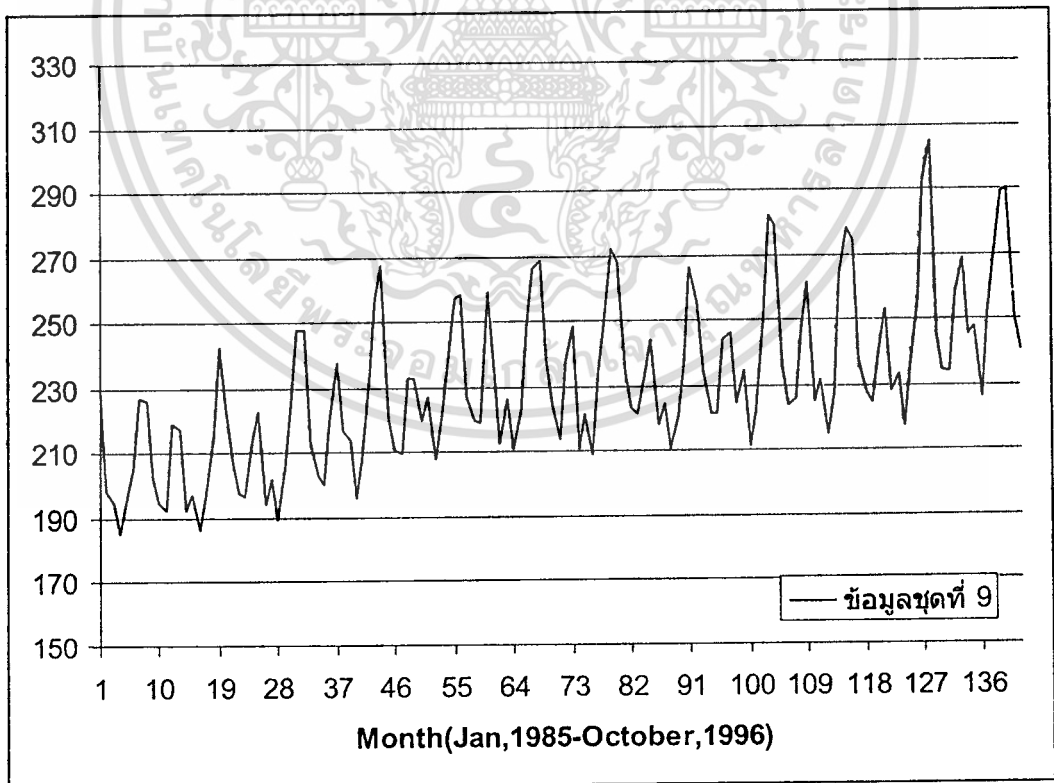


รูปที่ 4.7 ลักษณะของข้อมูลชุดที่ 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

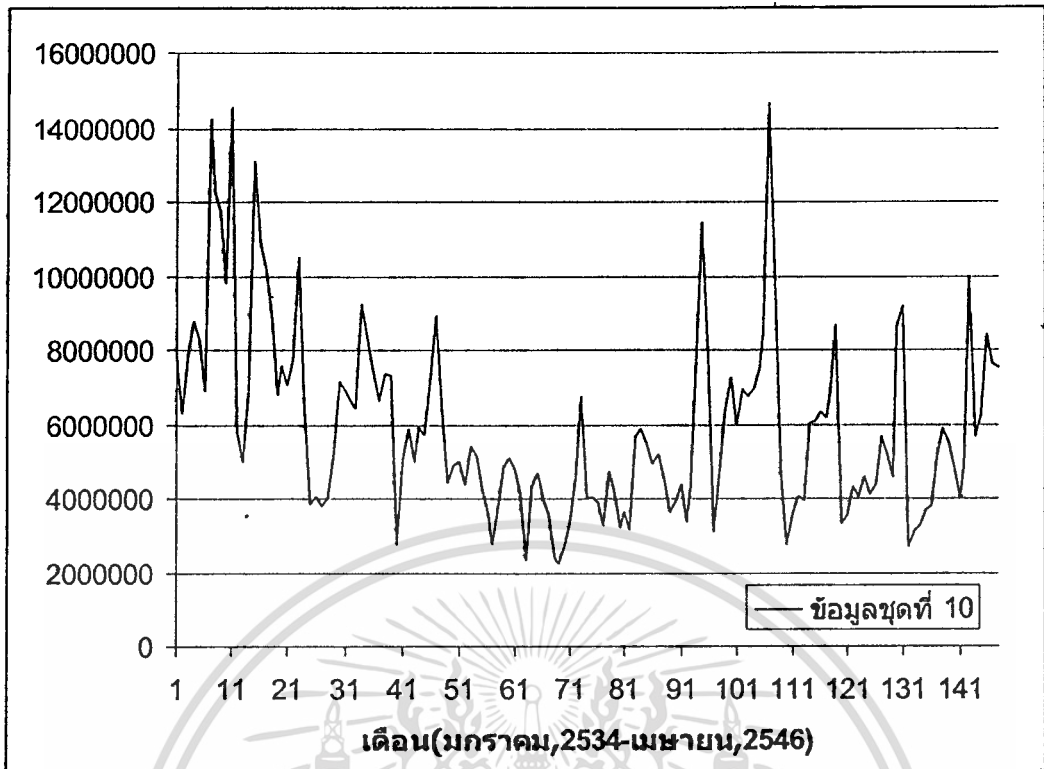


รูปที่ 4.8 ลักษณะของข้อมูลชุดที่ 8



รูปที่ 4.9 ลักษณะของข้อมูลชุดที่ 9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.10 ลักษณะของข้อมูลชุดที่ 10

4.2 การทำนายโดยใช้โมเดลแบบผสม

การทำนายโดยโมเดลแบบผสมจะทำการทดลอง 3 การทดลองย่อยคือการทดลองที่ 1 การทดลองที่ 2 และการทดลองที่ 3 โดยแต่ละการทดลองจะกำหนดค่าพารามิเตอร์ตามตารางที่ 4.1 ในการทดลองนี้จะเปรียบเทียบผลการทดลองที่ 1 และการทดลองที่ 2 โดยค่าพารามิเตอร์ที่กำหนดแตกต่างกันระหว่างการทดลองที่ 1 และการทดลองที่ 2 คือจำนวนฮิดเดนนิวรอนโดยต้องการทดสอบสมมติฐานที่ว่าเมื่อเพิ่มจำนวนฮิดเดนนิวรอนแล้วทำให้ผลการทำนายดีขึ้นหรือไม่ ส่วนการทดลองที่ 3 ต้องการทดสอบสมมติฐานที่ว่าเมื่อใช้ขนาดอินพุตที่มากกว่า period ของข้อมูลจะทำให้ผลการทำนายดีขึ้นหรือไม่ ซึ่งข้อมูลที่มี period คือ ข้อมูลชุดที่ 1, 2, 7 และ 10 ซึ่งเป็นข้อมูลที่มีอิทธิพลของฤดูกาลโดยจะเปรียบเทียบผลการทดลองที่ 3 กับการทดลองที่ 1

4.2.1 การกำหนดค่าพารามิเตอร์

ในการทำนายโดยโมเดลแบบผสมจะแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดข้อมูลสำหรับเรียนรู้และชุดข้อมูลสำหรับทดสอบโดยใช้อัตราส่วน 80 เปอร์เซ็นต์ และ 20 เปอร์เซ็นต์ตามลำดับ การกำหนดค่าพารามิเตอร์ในส่วนโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันกำหนดดังตารางที่ 4.1

ตารางที่ 4.1 แสดงค่าพารามิเตอร์ในการทำนายโดยโมเดลแบบผสม

	การทดลองที่ 1	การทดลองที่ 2	การทดลองที่ 3
พารามิเตอร์ของส่วนซิมพลิไฟด์พีชชี้อาทแมพ			
ขนาดแพทเทินอินพุท	2 - 6	2 - 6	ครอบคลุม period
ขนาดแพทเทินเอาต์พุท	3		
Step ทำนาย	1 step ข้างหน้า		
อัตราการเรียนรู้	0.1		
ค่าวิจิลแลนซ์	ใช้ค่าระหว่าง 0.75 ถึง 1.0		
จำนวนรอบในการเรียนรู้	100 รอบ		
เงื่อนไขในการหยุดเรียนรู้	เมื่อครบจำนวนรอบที่กำหนด		
พารามิเตอร์ของส่วนแบคพรอพพาเกชัน			
ขนาดแพทเทินอินพุท	เท่ากับขนาดแพทเทินอินพุทที่ดีที่สุดได้จากส่วนทำนายทิศทาง		
ขนาดแพทเทินเอาต์พุท	1		
Step ทำนาย	1 step ข้างหน้า		
อัตราการเรียนรู้	0.1		
จำนวนรอบในการเรียนรู้ สำหรับโครงข่ายทำนายขึ้น	15,000 รอบ		
จำนวนรอบในการเรียนรู้ สำหรับโครงข่ายทำนายลง	15,000 รอบ		
จำนวนชั้นอินพุท	1		
จำนวนชั้นฮิดเดน	1		
จำนวนชั้นเอาต์พุท	1		
จำนวนอินพุทนิรอรอน	เท่ากับขนาดแพทเทินอินพุท		
จำนวนฮิดเดนนิรอรอน	2 - 8	2 - 12	2 - 26
จำนวนเอาต์พุทนิรอรอน	1		
เงื่อนไขในการหยุดเรียนรู้	1. เมื่อครบจำนวนรอบที่กำหนด 2. เมื่อค่าความผิดพลาดของชุดข้อมูลเรียนรู้เพิ่มขึ้นหรือลดลงน้อยกว่าหรือเท่ากับ 0.00000001		

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการทดลองที่ 3 ขนาดของอินพุทจะครอบคลุม period ของข้อมูล ซึ่งขนาด period ของข้อมูลและขนาดอินพุทที่ใช้ในการทดลองแสดงในตารางที่ 4.2

ตารางที่ 4.2 แสดง period ของข้อมูลและขนาดแพทเทินอินพุท

ข้อมูลชุดที่	Period ของข้อมูล	ขนาดแพทเทินอินพุท
1	13	15
2	17	20
7	12	15
10	12	15

4.2.2 ข้อมูลที่ใช้ในการทดลอง

การทดลองที่ 1 ใช้ข้อมูลจำนวนทั้งหมด 10 ชุดข้อมูล

การทดลองที่ 2 ใช้ข้อมูลจำนวนทั้งหมด 10 ชุดข้อมูล

การทดลองที่ 3 ใช้ข้อมูลชุดที่ 1, 2, 7, 10 จำนวนทั้งหมด 4 ชุดข้อมูล

4.2.3 การเรียนรู้

ส่วนแรกเป็นการทดลองในส่วนนายทิศทางโดยในส่วนนี้จะทดลองเพื่อหาขนาดแพทเทินอินพุท จำนวนรอบและค่าวิจิลแลนซ์ที่ให้ผลในการทำนายทิศทางของข้อมูลชุดเรียนรู้ได้ถูกต้องมากที่สุด โดยการทดลองจะเริ่มจากขนาดแพทเทินอินพุทเท่ากับ 2 ค่าวิจิลแลนซ์เริ่มจาก 0.75 เรียนรู้จำนวน 100 รอบ เสร็จแล้วจึงเพิ่มค่าวิจิลแลนซ์อีก 0.05 เรียนรู้ 100 รอบ จนทำจนกว่าค่าวิจิลแลนซ์จะมีค่าเท่ากับ 1 เสร็จแล้วจึงนำผลมาเปรียบเทียบกันโดยเปรียบเทียบในแต่ละรอบ เช่น เปรียบเทียบผลของรอบที่ 1 ของค่าวิจิลแลนซ์ 0.75, 0.8, 0.85, 0.9, 0.95 และ 1 และเลือกค่าวิจิลแลนซ์ที่ให้ค่าความถูกต้องในการทำนายทิศทางมากที่สุด สำหรับรอบที่ 2 ถึงรอบที่ 100 ก็ทำการเปรียบเทียบเหมือนกัน ซึ่งผลจากการเปรียบเทียบนี้จะได้ค่าวิจิลแลนซ์ที่ดีที่สุดของแต่ละรอบที่ให้ค่าความถูกต้องในการทำนายทิศทางสูงสุด จากนั้นจึงเพิ่มขนาดแพทเทินอินพุทอีก 1 ทำการเรียนรู้และเปรียบเทียบผลเหมือนขนาดแพทเทินอินพุทเท่ากับ 2 เมื่อทำจบครบทุกขนาดแพทเทินอินพุทแล้วจึงนำผลของแต่ละขนาดแพทเทินอินพุทมาเปรียบเทียบกันในแต่ละรอบอีกครั้ง และเลือกขนาดแพทเทินอินพุทและรอบที่ให้เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทางมากที่สุด ซึ่งผลที่ได้จากการเปรียบเทียบในขั้นสุดท้ายนี้จะได้ขนาดแพทเทินอินพุทที่ดีที่สุด รอบที่ดีที่สุดและค่าวิจิลแลนซ์ที่ดีที่สุดให้ค่าความถูกต้องในการทำนายทิศทางมากที่สุด ซึ่งขนาดของแพทเทินอินพุทที่ดีที่สุดนี้จะถูกนำไปใช้ในการนอร์มัลไลซ์แพทเทินอินพุทสำหรับส่วนทำนายค่าต่อไป ในส่วนนี้โครงข่ายจะหยุดเรียนรู้เมื่อครบจำนวนรอบที่กำหนด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเรียนรู้ในส่วนทำนายค่า ก่อนที่ข้อมูลจะถูกส่งเข้าโครงข่ายข้อมูลจะถูกนอร์มัลไลซ์ โดยใช้ขนาดแพทเทินอินพุทที่ได้จากขั้นตอนทำนายทิศทางและแพทเทินอินพุทจะถูกแบ่งตามทิศทางของคำตอบของแพทเทินอินพุทนั้น ดังอธิบายไว้ในบทที่ 3 การเรียนรู้นี้จะหยุดเรียนรู้เมื่อครบจำนวนรอบที่กำหนดหรือเมื่อค่าความผิดพลาดของชุดข้อมูลเรียนรู้เพิ่มขึ้นหรือลดลงต่ำกว่าหรือเท่ากับค่าที่กำหนด

4.2.4 วิธีการทำนาย

ข้อมูลทดสอบจะถูกนอร์มัลไลซ์โดยใช้ขนาดแพทเทินอินพุทที่ดีที่สุดที่ได้จากขั้นตอนการเรียนรู้ของข้อมูลชุดเรียนรู้จากนั้นจึงถูกส่งเข้าทำนายทิศทางของคำตอบของแพทเทินอินพุท ซึ่งทิศทางของคำตอบที่เป็นไปได้มี 3 ทิศทางคือ ทิศทางขึ้น ทิศทางลงและทิศทางไม่เปลี่ยนแปลง ถ้าคำตอบบอกว่ามีทิศทางขึ้นก็ส่งไปทำนายค่าในโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันส่วนที่ใช้สำหรับทำนายค่าที่คำตอบเป็นทิศทางขึ้น ถ้าแพทเทินอินพุทถูกทำนายทิศทางว่าคำตอบมีทิศทางลงก็ส่งเข้าไปทำนายค่าในโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันที่ใช้สำหรับทำนายค่าที่คำตอบเป็นทิศทางลง ส่วนแพทเทินอินพุทที่ถูกทำนายว่าคำตอบมีทิศทางไม่เปลี่ยนแปลงจะใช้ค่าก่อนจุดที่ทำนายในอนุกรมเวลาเป็นคำตอบ

4.2.5 ผลการทดลอง

ผลการทำนายโดยโมเดลแบบผสมแสดงในตารางที่ 4.3, 4.4 และ 4.5 ตามลำดับ

ตารางที่ 4.3 ผลการทดลองที่ 1

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	6	6	0.0119	0.0122	100
2	5	8	0.0122	0.0196	95.31
3	5	4	3,126.9	8,352.26	69.23
4	2	2	12.91	30.54	84.38
5	6	8	15,332.82	38,008.91	80
6	2	8	294,812.87	1,051,601.86	85.71
7	4	5	25,716.68	33,979.34	100
8	4	7	19,400.98	36,980.61	79.55

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 (ต่อ)

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
9	4	4	6.83	8.33	92.86
10	5	5	1,395,979.50	1,804,231.79	75.86

ตารางที่ 4.4 ผลการทดลองที่ 2

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	6	6	0.0119	0.0122	100
2	5	8	0.0122	0.0196	95.31
3	5	10	3,029.62	8,411.03	69.23
4	2	2	12.91	30.54	84.38
5	6	11	14,630.93	38,285.91	80
6	2	8	294,812.87	1,051,601.86	85.71
7	4	9	25,550.46	34,354.78	100
8	4	7	19,400.98	36,980.61	79.55
9	4	9	6.47	8.23	92.86
10	5	9	1,328,875.16	1,864,920.58	75.86

ตารางที่ 4.5 ผลการทดลองที่ 3

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	9	11	0.0078	0.0081	100
2	5	8	0.0122	0.0196	95.31
7	4	9	25,550.46	34,354.78	100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 (ต่อ)

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
9	11	5	4.84	6.34	100

4.2.6 สรุปผลการทดลอง

จากการเปรียบเทียบผลการทดลองที่ 1 และการทดลองที่ 2 เมื่อให้ขนาดอินพุทเท่าเดิม แต่เพิ่มจำนวนฮิดเดนนิวรอนในการเรียนรู้มากขึ้นผลเป็นดังนี้คือ

1. ค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายของข้อมูลทดสอบไม่เปลี่ยนแปลงจำนวน 5 ชุดข้อมูลคือข้อมูลชุดที่ 1, 2, 4, 6 และ 8
2. ค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายของข้อมูลทดสอบลดลงจำนวน 1 ชุดข้อมูลคือข้อมูลชุดที่ 9
3. ค่าความผิดพลาดของข้อมูลเรียนรู้ลดลงแต่ค่าความผิดพลาดในการทำนายข้อมูลทดสอบเพิ่มขึ้นจำนวน 4 ชุดข้อมูลคือข้อมูลชุดที่ 3, 5, 7 และ 10

โดยที่เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทางของข้อมูลทดสอบทั้งหมดเท่าเดิม เนื่องจากไม่ได้เปลี่ยนแปลงขนาดอินพุท การที่ค่าความผิดพลาดของข้อมูลเรียนรู้ลดลงแต่ค่าความผิดพลาดในการทำนายข้อมูลทดสอบเพิ่มขึ้นใน 4 ชุดข้อมูลหลังจากเกิดจากโมเดลเรียนรู้ข้อมูลชุดเรียนรู้ได้ดีเกินไป (Over Fitting) จากการเปรียบเทียบผลการทดลองทั้งสองแสดงให้เห็นว่า สำหรับข้อมูลที่ใช้ในการทดลองในงานวิจัยนี้การเพิ่มจำนวนฮิดเดนนิวรอนให้มากขึ้นไม่ส่งผลให้ความถูกต้องในการทำนายข้อมูลทดสอบเพิ่มขึ้น

เมื่อเปรียบเทียบผลการทดลองที่ 3 กับผลการทดลองที่ 1 พบว่าเมื่อเพิ่มขนาดอินพุทโดยให้โมเดลเรียนรู้ครอบคลุม period ของชุดข้อมูลจำนวน 4 ชุดข้อมูลคือข้อมูลชุดที่ 1, 2, 7 และ 9 ผลเป็นดังนี้คือ เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทางและค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายข้อมูลทดสอบเท่าเดิมจำนวน 2 ชุดข้อมูลคือข้อมูลชุดที่ 2 และ 7 เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทางเพิ่มขึ้นและค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายข้อมูลทดสอบลดลงจำนวน 2 ชุดข้อมูลคือข้อมูลชุดที่ 1 และ 9 ซึ่งแสดงว่าสำหรับข้อมูลที่ใช้ในการทดลองในงานวิจัยนี้การเพิ่มขนาดอินพุทให้ครอบคลุม period ของชุดข้อมูลทำให้โมเดลสามารถทำนายทิศทางได้ถูกต้องมากขึ้นและทำให้ค่าความผิดพลาดในการทำนายข้อมูลทดสอบลดลง

ตารางที่ 4.6 แสดงผลการทำนายที่ดีที่สุดที่ใช้โมเดลแบบผสมในการทำนาย

ตารางที่ 4.6 สรุปผลการทำนายโดยโมเดลแบบผสม

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	9	11	0.0078	0.0081	100
2	5	8	0.0122	0.0196	95.31
3	5	4	3,126.9	8,352.26	69.23
4	2	2	12.91	30.54	84.38
5	6	8	15,332.82	38,008.91	80
6	2	8	294,812.87	1,051,601.86	85.71
7	4	5	25,716.68	33,979.34	100
8	4	7	19,400.98	36,980.61	79.55
9	11	5	4.84	6.34	100
10	5	5	1,395,979.50	1,804,231.79	75.86

4.3 การทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

4.3.1 การกำหนดค่าพารามิเตอร์

ในงานวิจัยนี้จะเปรียบเทียบผลทำนายโดยโมเดลแบบผสมกับการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดสำหรับเรียนรู้และชุดสำหรับทดสอบโดยใช้อัตราส่วน 80 เปอร์เซ็นต์และ 20 เปอร์เซ็นต์เท่ากับการทำนายโดยใช้โมเดลแบบผสม ซึ่งการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นจะประกอบด้วย 3 การทดลองย่อยคือ การทดลองที่ 4, 5 และ 6 การกำหนดสถาปัตยกรรมและค่าพารามิเตอร์สำหรับแต่ละการทดลองย่อยแสดงในตารางที่ 4.7

ตารางที่ 4.7 แสดงค่าพารามิเตอร์ในการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

	การทดลองที่ 4	การทดลองที่ 5	การทดลองที่ 6
ขนาดแพทเทินอินพุท	2 - 6	2 - 6	ครอบคลุม period
ขนาดแพทเทินเอาท์พุท	1		

ตารางที่ 4.7 (ต่อ)

	การทดลองที่ 4	การทดลองที่ 5	การทดลองที่ 6
Step ทำนาย	1 step ข้างหน้า		
อัตราการเรียนรู้	0.1		
จำนวนรอบในการเรียนรู้	30,000 รอบ		
จำนวนชั้นอินพุต	1		
จำนวนชั้นฮิดเดน	1		
จำนวนชั้นเอาต์พุต	1		
จำนวนอินพุตนิเวรอน	เท่ากับขนาดแพทเทินอินพุต		
จำนวนฮิดเดนนิเวรอน	2 - 8	2 - 12	2 - 26
จำนวนเอาต์พุตนิเวรอน	1		
เงื่อนไขในการหยุดการเรียนรู้	1. เมื่อครบจำนวนรอบที่กำหนด 2. เมื่อค่าความผิดพลาดของชุดข้อมูลเรียนรู้เพิ่มขึ้นหรือลดลงน้อยกว่าหรือเท่ากับ 0.00000001		

4.3.2 ข้อมูลที่ใช้ในการทดลอง

การทดลองที่ 4 ใช้ข้อมูลจำนวนทั้งหมด 10 ชุดข้อมูล

การทดลองที่ 5 ใช้ข้อมูลจำนวนทั้งหมด 10 ชุดข้อมูล

การทดลองที่ 6 ใช้ข้อมูลชุดที่ 1, 2, 7, 10 จำนวนทั้งหมด 4 ชุดข้อมูล

4.3.3 การเรียนรู้

การเรียนรู้ของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันที่ใช้ในการเปรียบเทียบนี้จะเริ่มจากขนาดแพทเทินอินพุตเท่ากับ 2 เรียนรู้จนครบจำนวนฮิดเดนนิเวรอนที่กำหนดและหยุดการเรียนรู้เมื่อครบจำนวนรอบที่กำหนดหรือเมื่อค่าความผิดพลาดของชุดเรียนรู้เพิ่มขึ้นหรือลดลงน้อยกว่าหรือเท่ากับค่าที่กำหนด แล้วจึงเพิ่มขนาดของแพทเทินอินพุตอีก 1 จนทำจนกระทั่งครบทุกขนาดแพทเทินอินพุต

4.3.4 การทำนาย

เมื่อเรียนรู้จนครบจำนวนแพทเทินอินพุตและจำนวนฮิดเดนนิเวรอนแล้วจึงเลือกโครงข่ายที่ให้ค่าความผิดพลาดในการทำนายข้อมูลเรียนรู้น้อยที่สุดแล้วจึงส่งข้อมูลชุดทดสอบเข้าทดสอบพร้อมกับคำนวณเปอร์เซ็นต์ความถูกต้องในการทำนายทิศทางโดยเปรียบเทียบกับทิศทางของข้อมูลอนุกรมเวลาที่เกิดขึ้นจริง

4.3.5 ผลการทดลอง

ผลการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นแสดงในตารางที่

4.8, 4.9 และ 4.10

ตารางที่ 4.8 ผลการทดลองที่ 4

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	6	4	0.0091	0.0093	100
2	6	7	0.0112	0.0132	97.86
3	6	7	2,835.41	15,163.47	61.54
4	6	7	17.44	42.89	75
5	6	8	9,232.48	54,955.32	60
6	6	6	307,509.07	1,188,381.75	84.62
7	6	5	14,110.33	38,262.12	73.33
8	6	8	23,386.21	4,1401.62	70.45
9	6	8	5.26	9.99	85.19
10	6	7	1,565,331.20	2,146,598.01	65.52

ตารางที่ 4.9 ผลการทดลองที่ 5

ข้อมูลชุดที่	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	6	4	0.0091	0.0093	100
2	6	11	0.0083	0.0102	98.39
3	6	12	2,330.12	13,923.07	69.23
4	6	10	16.23	48.18	71.88
5	6	12	8,964.29	56,620.12	60
6	6	12	303,918.11	1,228,130.91	76.92

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 (ต่อ)

ข้อมูลชุดที่	ขนาดอินพุตที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
7	6	12	12,922.02	44,186.84	73.33
8	6	10	22,033.22	41,924.16	65.91
9	6	8	5.26	9.99	85.19
10	6	12	1,538,341.53	2,180,194.32	68.97

ตารางที่ 4.10 ผลการทดลองที่ 6

ข้อมูลชุดที่	ขนาดอินพุตที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	15	9	0.0054	0.0059	100
2	18	24	0.0054	0.0068	98.92
7	15	21	4,656.32	62,292.38	64.29
9	15	21	3.74	7.81	88.46

4.3.6 สรุปผลการทดลอง

จากการเปรียบเทียบผลการทดลองที่ 4 และการทดลองที่ 5 เมื่อให้ขนาดอินพุตเท่าเดิม แต่เพิ่มจำนวนฮิดเดนนิวรอนในการเรียนรู้มากขึ้นผลเป็นดังนี้คือ

1. ค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายของข้อมูลทดสอบไม่เปลี่ยนแปลงจำนวน 2 ชุดข้อมูลคือข้อมูลชุดที่ 1 และ 9
2. ค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายของข้อมูลทดสอบลดลงจำนวน 2 ชุดข้อมูลคือข้อมูลชุดที่ 2 และ 3
3. ค่าความผิดพลาดของข้อมูลเรียนรู้ลดลงแต่ค่าความผิดพลาดในการทำนายข้อมูลทดสอบเพิ่มขึ้นจำนวน 6 ชุดข้อมูลคือข้อมูลชุดที่ 4, 5, 6, 7, 8 และ 10

การที่ค่าความผิดพลาดของข้อมูลเรียนรู้ลดลงแต่ค่าความผิดพลาดในการทำนายข้อมูลทดสอบเพิ่มขึ้นใน 6 ชุดข้อมูลหลังอาจเกิดจากโมเดลเรียนรู้ข้อมูลชุดเรียนรู้ได้ดีเกินไป (Over Fitting) จึงทำให้ผลการทำนายข้อมูลทดสอบผิดพลาดมากขึ้น จากการเปรียบเทียบผลการทดลอง

ทั้งสองแสดงให้เห็นว่าสำหรับข้อมูลที่ใช้ในการทดลองในงานวิจัยนี้การเพิ่มจำนวนฮิดเดนนิวรอนมากขึ้นส่งผลให้ความถูกต้องในการทำนายข้อมูลทดสอบลดลง

เมื่อเปรียบเทียบผลการทดลองที่ 4 กับผลการทดลองที่ 6 พบว่าเมื่อเพิ่มขนาดอินพุตโดยให้โมเดลเรียนรู้ครอบคลุม period ของชุดข้อมูลจำนวน 4 ชุดข้อมูลคือข้อมูลชุดที่ 1, 2, 7 และ 9 ผลเป็นดังนี้คือ

1. ค่าความผิดพลาดของข้อมูลเรียนรู้และค่าความผิดพลาดในการทำนายของข้อมูลทดสอบลดลงเท่าเดิมจำนวน 3 ชุดข้อมูลคือข้อมูลชุดที่ 1, 2 และ 9

2. ค่าความผิดพลาดของข้อมูลเรียนรู้ลดลงแต่ค่าความผิดพลาดในการทำนายของข้อมูลทดสอบเพิ่มขึ้นจำนวน 1 ชุดข้อมูลคือข้อมูลชุดที่ 7

จากการเปรียบเทียบผลการทดลองที่ 4 กับผลการทดลองที่ 6 แสดงให้เห็นว่าสำหรับข้อมูลที่ใช้ในการทดลองในงานวิจัยนี้เมื่อเพิ่มขนาดอินพุตให้ครอบคลุม period ของชุดข้อมูลทำให้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันทำนายข้อมูลทดสอบได้ดีขึ้น

ตารางที่ 4.11 แสดงผลการทำนายที่ดีที่สุดของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

ตารางที่ 4.11 สรุปผลการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

ข้อมูลชุดที่	ขนาดอินพุตที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลเรียนรู้	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	15	9	0.0054	0.0059	100
2	18	24	0.0054	0.0068	98.92
3	6	12	2,330.12	13,923.07	69.23
4	6	7	17.44	42.89	75
5	6	8	9,232.48	54,955.32	60
6	6	6	307,509.07	1,188,381.75	84.62
7	6	5	14,110.33	38,262.12	73.33
8	6	8	23,386.21	41,401.62	70.45
9	15	21	3.74	7.81	88.46
10	6	7	1,565,331.20	2,146,598.01	65.52

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 สรุปผลการเปรียบเทียบการทำนายโดยโมเดลแบบผสมและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

ตารางที่ 4.12 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น

ข้อมูลชุดที่	โมเดลแบบผสม				โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่น			
	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง	ขนาดอินพุทที่ดีที่สุด	ฮิดเดนนิวรอนที่ดีที่สุด	RMSE ของข้อมูลทดสอบ	เปอร์เซ็นต์ความถูกต้องในการทำนายทิศทาง
1	9	11	0.0081	100	15	9	0.0059	100
2	5	8	0.0196	95.31	18	24	0.0068	98.92
3	5	4	8,352.26	69.23	6	7	13,923.07	69.23
4	2	2	30.54	84.38	6	7	42.89	75
5	6	8	38,008.91	80	6	8	54,955.32	60
6	2	8	1,051,601.86	85.71	6	6	1,188,381.75	84.62
7	4	5	33,979.34	100	6	5	38,262.12	73.33
8	4	7	36,980.61	79.55	6	8	41,401.62	70.45
9	11	5	6.34	100	15	21	7.81	88.46
10	5	5	1,804,231.79	75.86	6	7	2,146,598.01	65.52

พิจารณารางที่ 4.12 จากข้อมูลทดลองทั้งหมด 10 ชุดข้อมูลสรุปผลได้ดังนี้

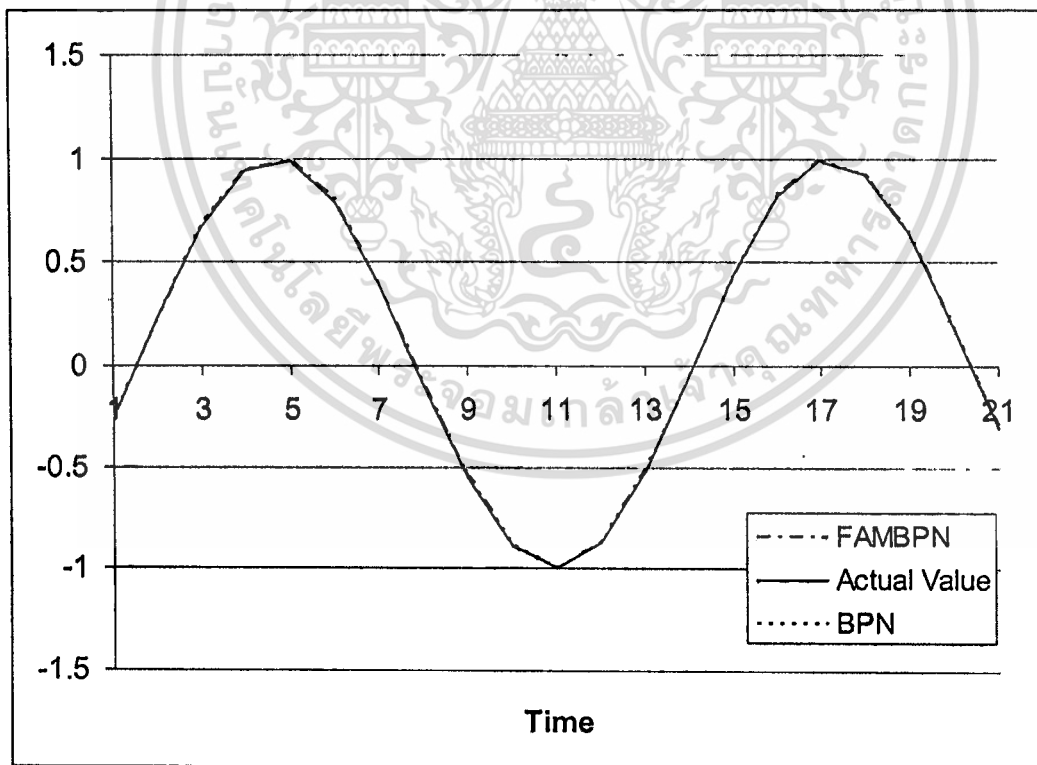
1. โมเดลแบบผสมสามารถทำนายทิศทางได้ถูกต้องมากกว่าและค่าความผิดพลาดในการทำนายข้อมูลทดสอบน้อยกว่าการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชั่นจำนวน 8 ชุดข้อมูลคือข้อมูลชุดที่ 3, 4, 5, 6, 7, 8, 9 และ 10

2. ทั้งสองโมเดลสามารถทำนายทิศทางได้ความถูกต้อง 100 เปอร์เซ็นต์เท่ากันแต่โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันให้ค่าความผิดพลาดในการทำนายข้อมูลทดสอบต่ำกว่าการทำนายโดยโมเดลแบบผสมจำนวน 1 ชุดข้อมูลคือข้อมูลชุดที่ 1

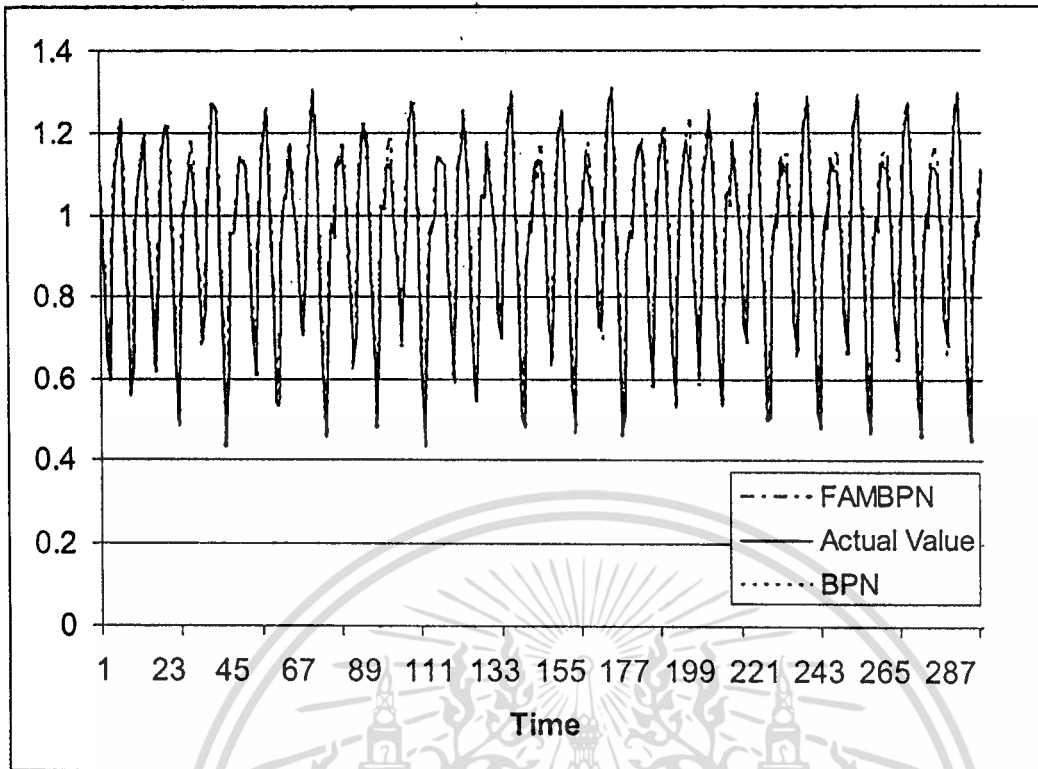
3. โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันสามารถทำนายทิศทางได้ถูกต้องมากกว่าและค่าความผิดพลาดในการทำนายข้อมูลทดสอบน้อยกว่าการทำนายโดยใช้โมเดลแบบผสมจำนวน 1 ชุดข้อมูลคือข้อมูลชุดที่ 2

จากการเปรียบเทียบการทำนายทั้งสองวิธีสรุปได้ว่าการทำนายข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบผสมมีความถูกต้องในการทำนายทิศทางและให้ค่าความผิดพลาดในการทำนายดีกว่าการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันซึ่งแสดงให้เห็นด้วยว่าหากโมเดลแบบผสมสามารถทำนายทิศทางได้ถูกต้องมากกว่าก็จะทำให้สามารถทำนายข้อมูลทดสอบได้ถูกต้องมากขึ้นด้วย

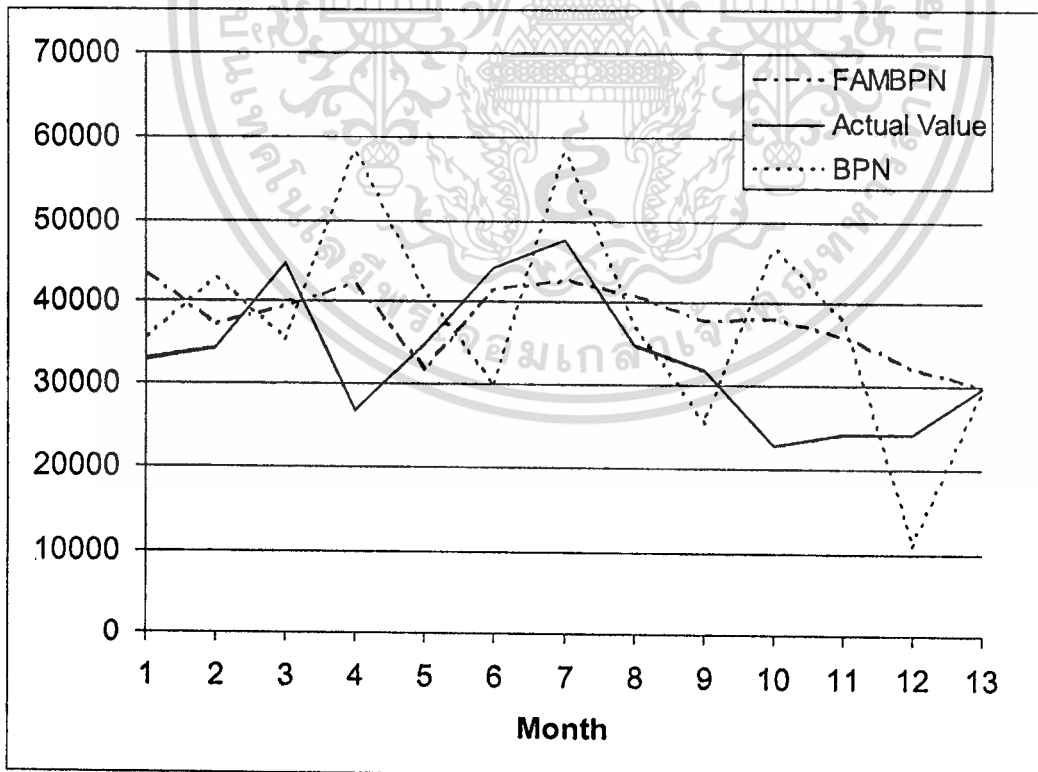
4.5 รูปแสดงผลการทำนายโดยโมเดลแบบผสมเปรียบเทียบกับการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน



รูปที่ 4.11 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPN ของข้อมูลทดสอบชุดที่ 1

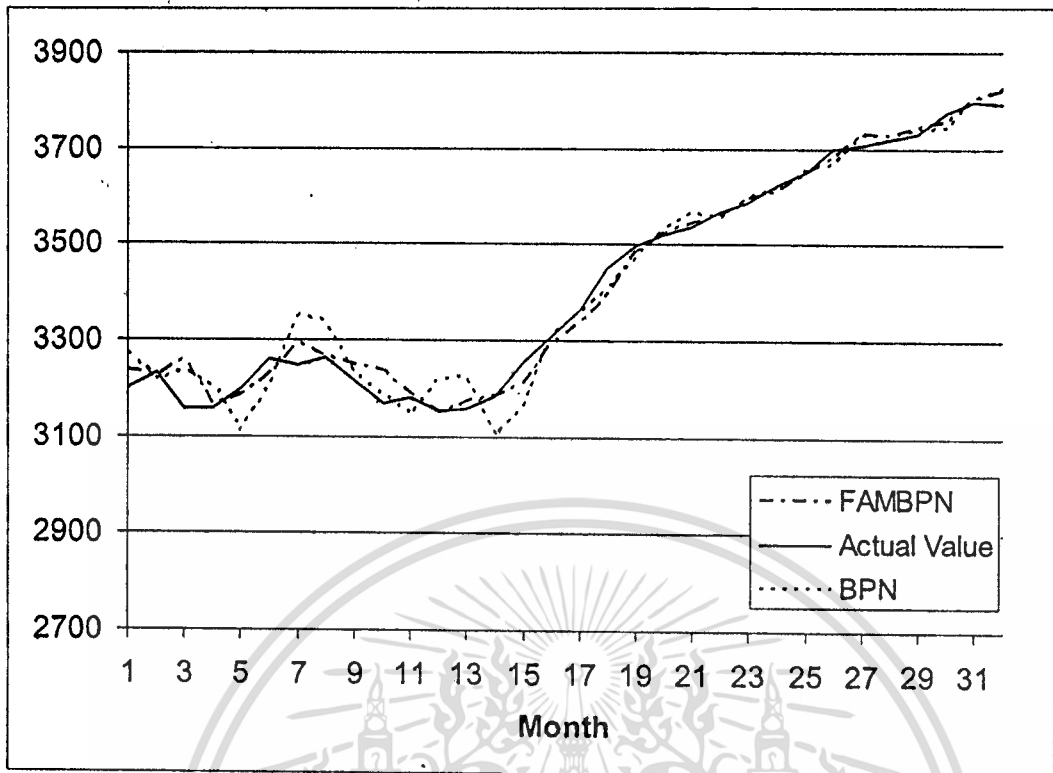


รูปที่ 4.12 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 2

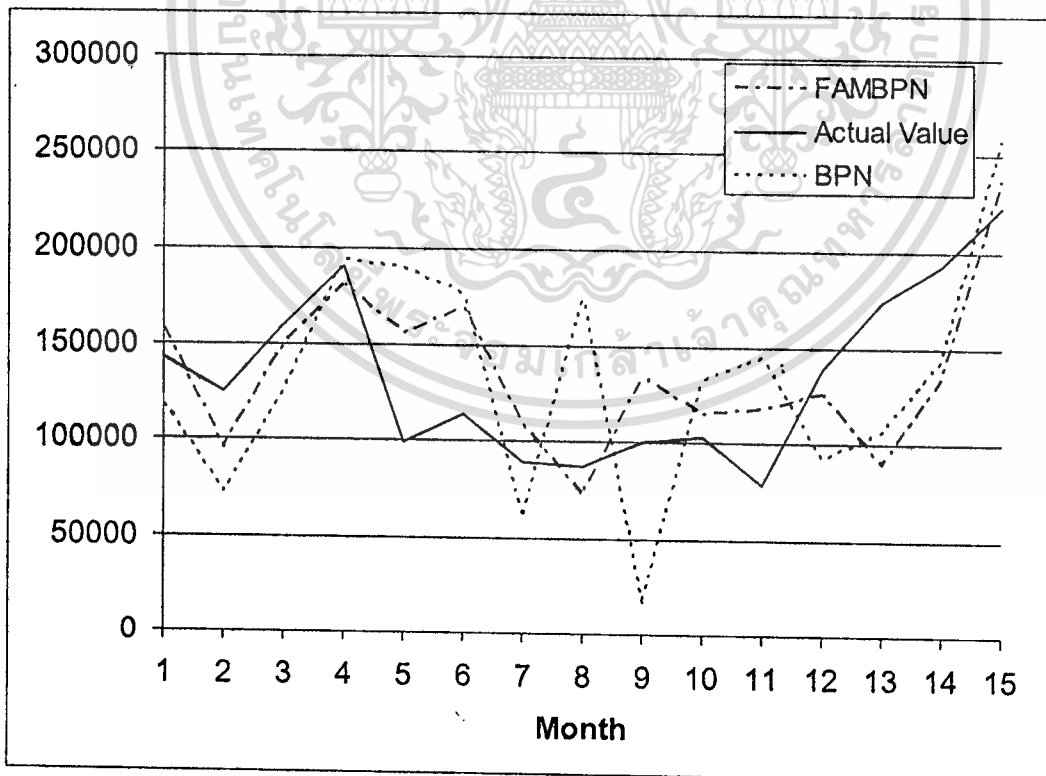


รูปที่ 4.13 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

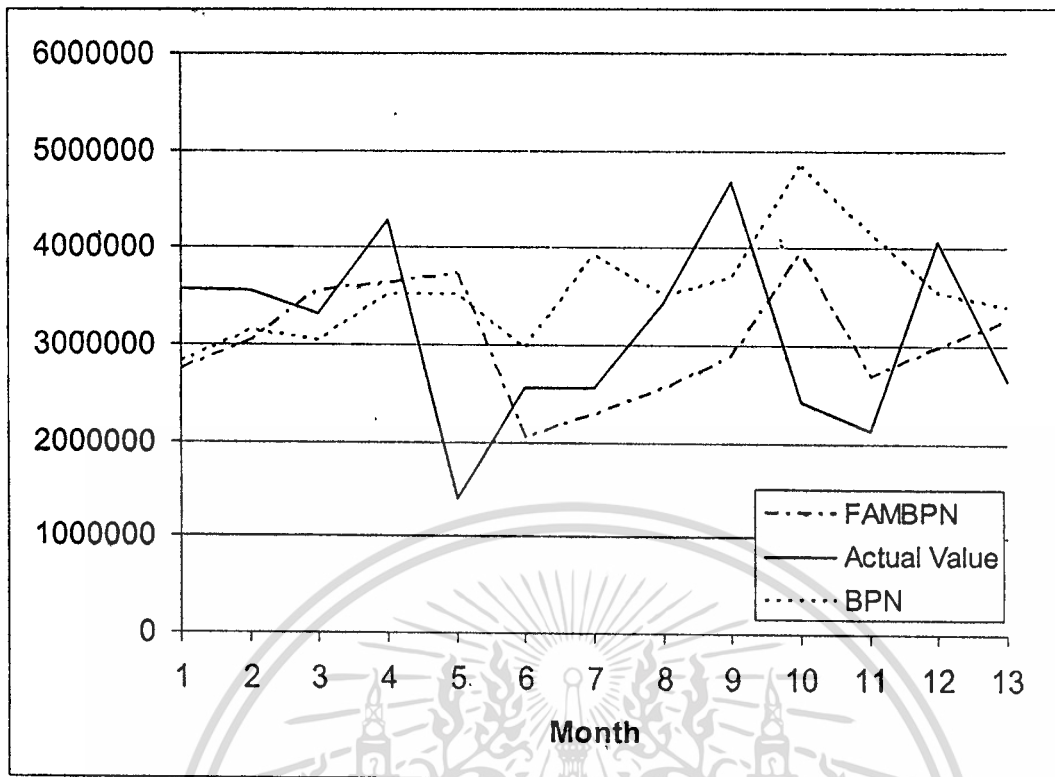


รูปที่ 4.14 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 4

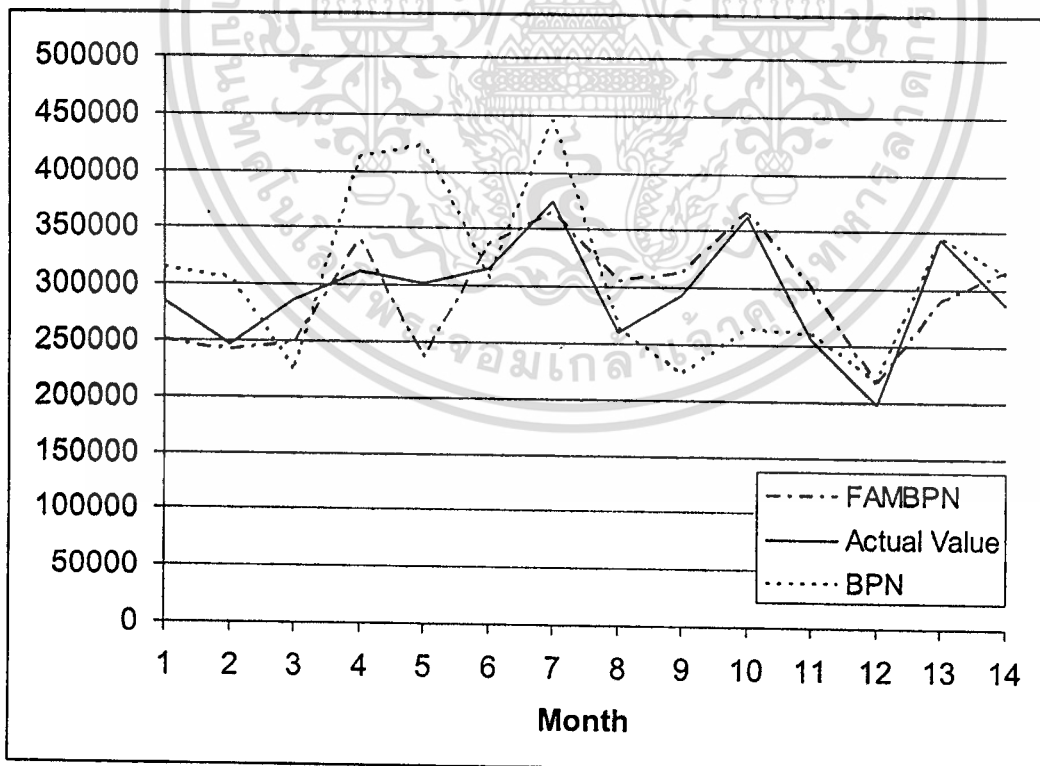


รูปที่ 4.15 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

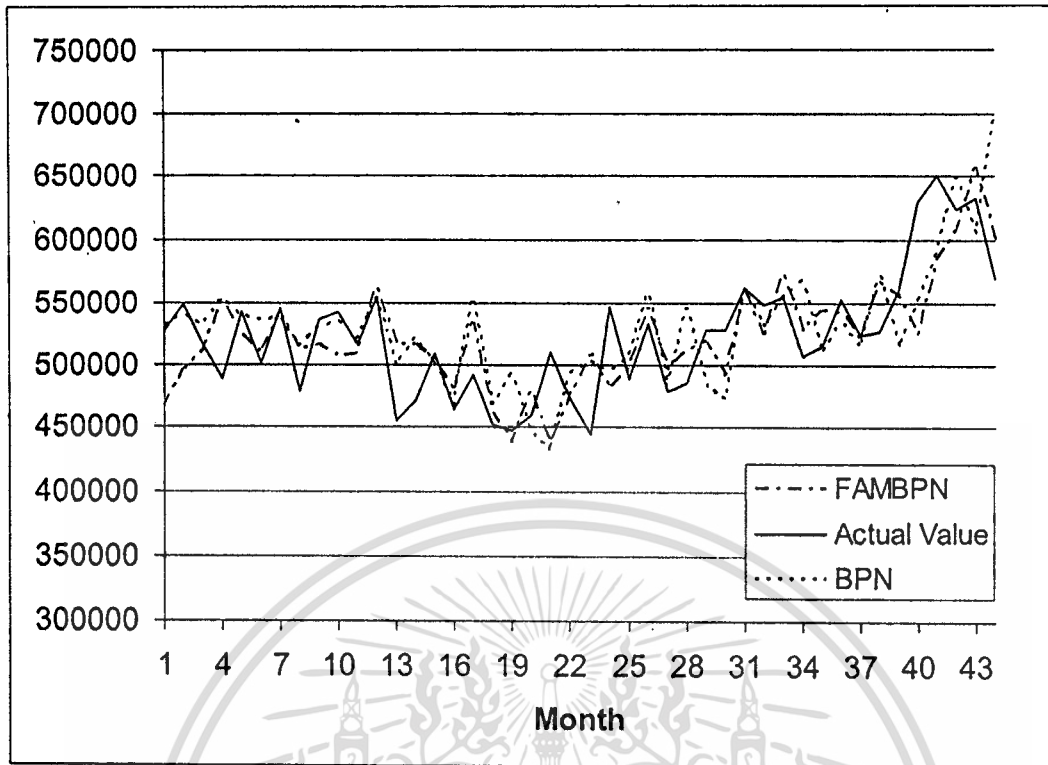


รูปที่ 4.16 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 6

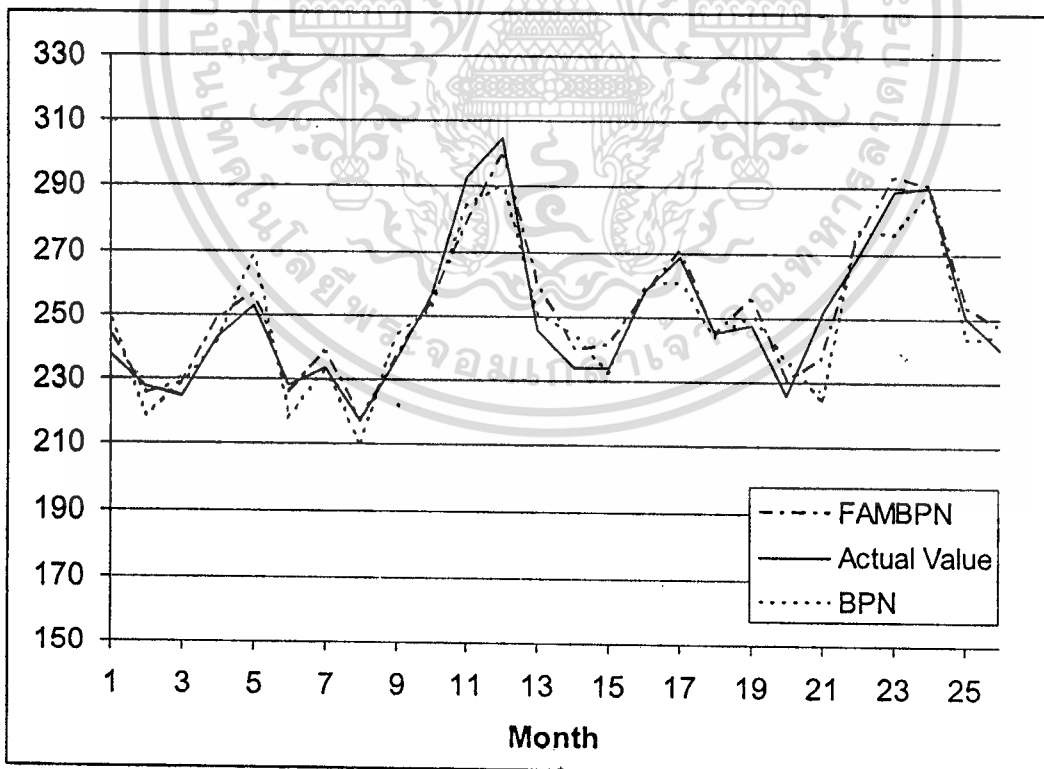


รูปที่ 4.17 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

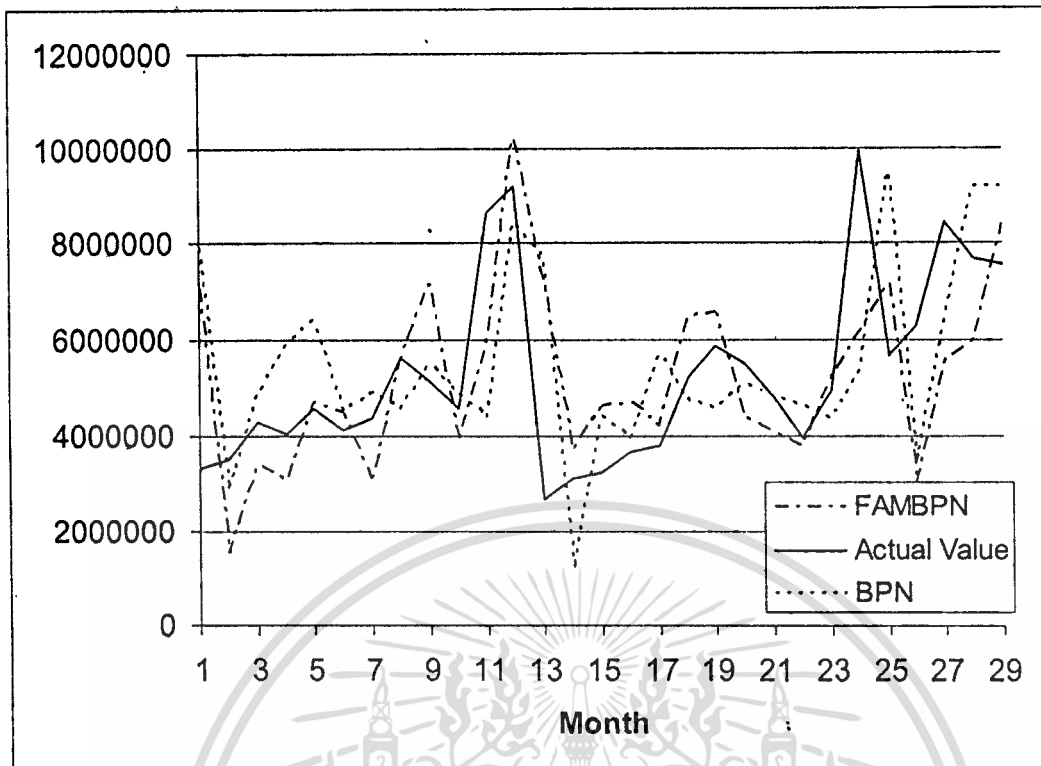


รูปที่ 4.18 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 8



รูปที่ 4.19 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.20 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ BPNN ของข้อมูลทดสอบชุดที่ 10

4.6 การทำนายโดยวิธี Autoregressive Integrated Moving Average (ARIMA)

4.6.1 การกำหนดรูปแบบและค่าพารามิเตอร์

การทำนายโดยวิธี ARIMA จะใช้โปรแกรมสำเร็จรูปคือ Minitab for Windows : version 14.12 ในการกำหนดรูปแบบและประมาณค่าพารามิเตอร์ โดยการทดลองจะแบ่งข้อมูลออกเป็น 2 ส่วนคือส่วนแรกสำหรับกำหนดรูปแบบและประมาณค่าพารามิเตอร์ ส่วนที่ 2 สำหรับทดสอบโดยใช้อัตราส่วน 80 เปอร์เซ็นต์และ 20 เปอร์เซ็นต์ตามลำดับ

4.6.2 การทำนาย

หลังจากได้รูปแบบที่เหมาะสมแล้วจะใช้รูปแบบนั้นในการทำนายข้อมูลทดสอบและเปรียบเทียบผลการทำนายกับข้อมูลจริงที่เกิดขึ้นเพื่อคำนวณหาค่าความผิดพลาด

4.6.3 ข้อมูลที่ใช้ทดลอง

ใช้ข้อมูลทดลองกลุ่มที่ 2 จำนวน 8 ชุดข้อมูล

4.6.4 ผลการทดลอง

ผลการทดลองแสดงในตารางที่ 4.13

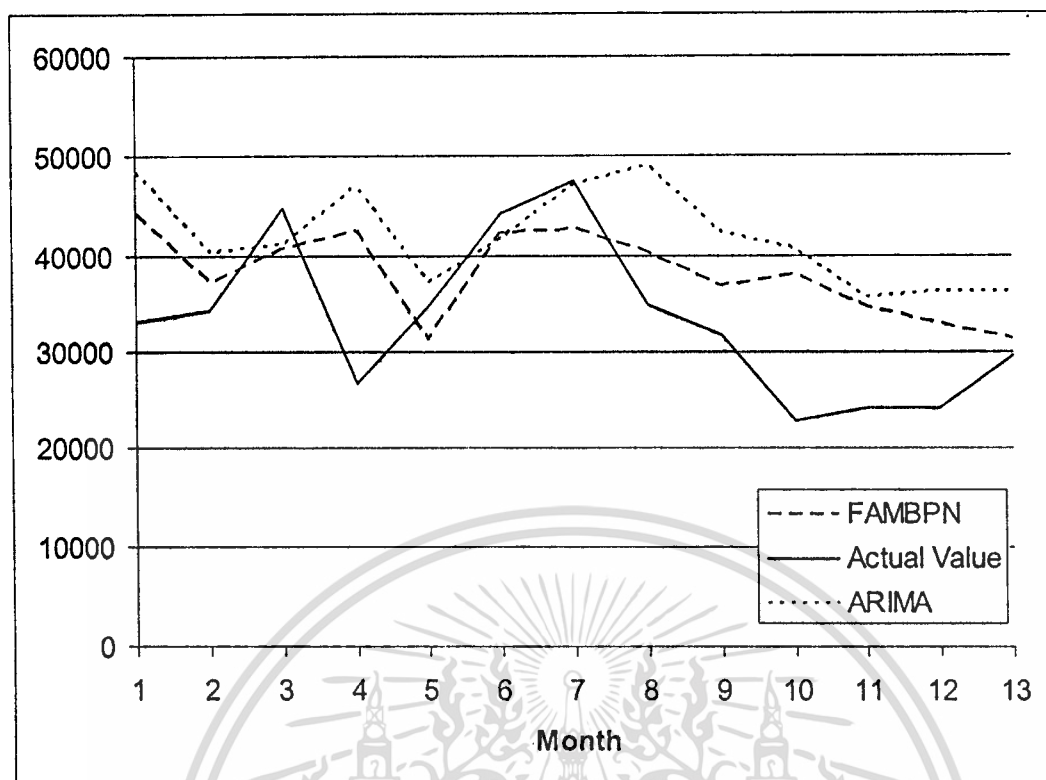
ตารางที่ 4.13 ผลการทำนายโดยวิธี ARIMA

ข้อมูลชุดที่	รูปแบบที่เหมาะสม		MSE	RMSE
	ARIMA	SARIMA		
3	(1,1,1)	-	121,214,268.90	11,009.74
4	(0,1,2)	-	1,333.73	36.52
5	(4,1,5)	-	1,596,788,500	39,959.84
6	2,1,0)	-	1.16E+12	1,077,392.71
7	(2,0,0)	(0,1,0)	4,761,080,495	69,000.58
8	(2,1,0)	(2,1,2)	2,393,337,401	48,921.75
9	(2,0,1)	(0,1,2)	128.30	11.33
10	(0,1,3)	(1,1,0)	3.41E+12	1,845,443.48

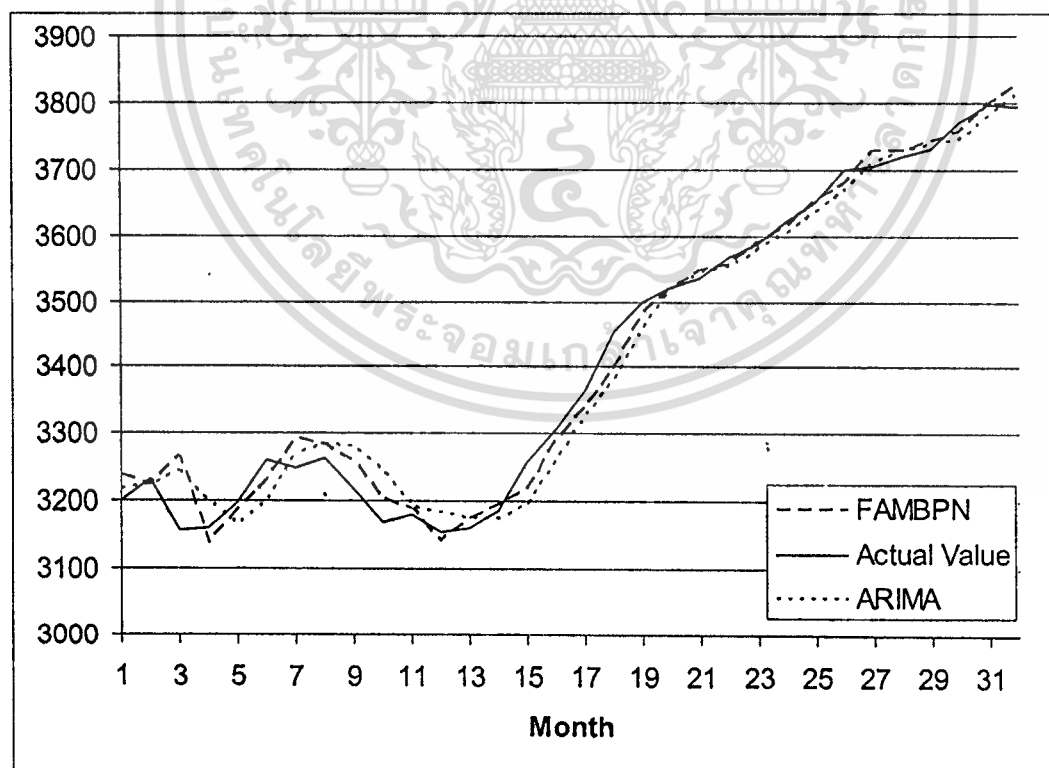
4.7 สรุปผลการเปรียบเทียบการทำนายโดยโมเดลแบบผสมและวิธี ARIMA

จากการเปรียบเทียบผลการทดลองที่ 1 และการทดลองที่ 3 ปรากฏว่าโมเดลแบบผสมให้ความผิดพลาดในการทำนายข้อมูลทดสอบต่ำกว่าการทำนายโดยใช้วิธี ARIMA ทุกชุดข้อมูล ดังนั้น จึงสามารถสรุปได้ว่าการทำนายข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบผสมมีความถูกต้องในการทำนายสูงกว่าการทำนายโดยวิธี ARIMA

4.8 รูปแสดงผลการทำนายโดยโมเดลแบบผสมเปรียบเทียบกับการทำนายโดยวิธี ARIMA

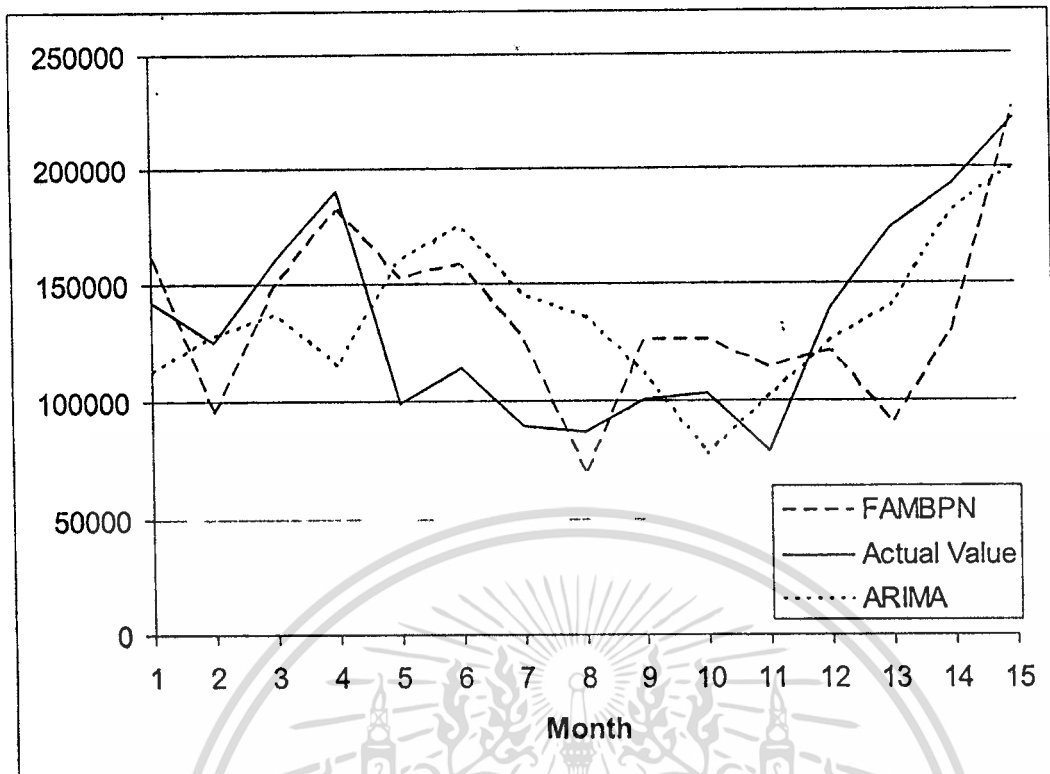


รูปที่ 4.21 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 3

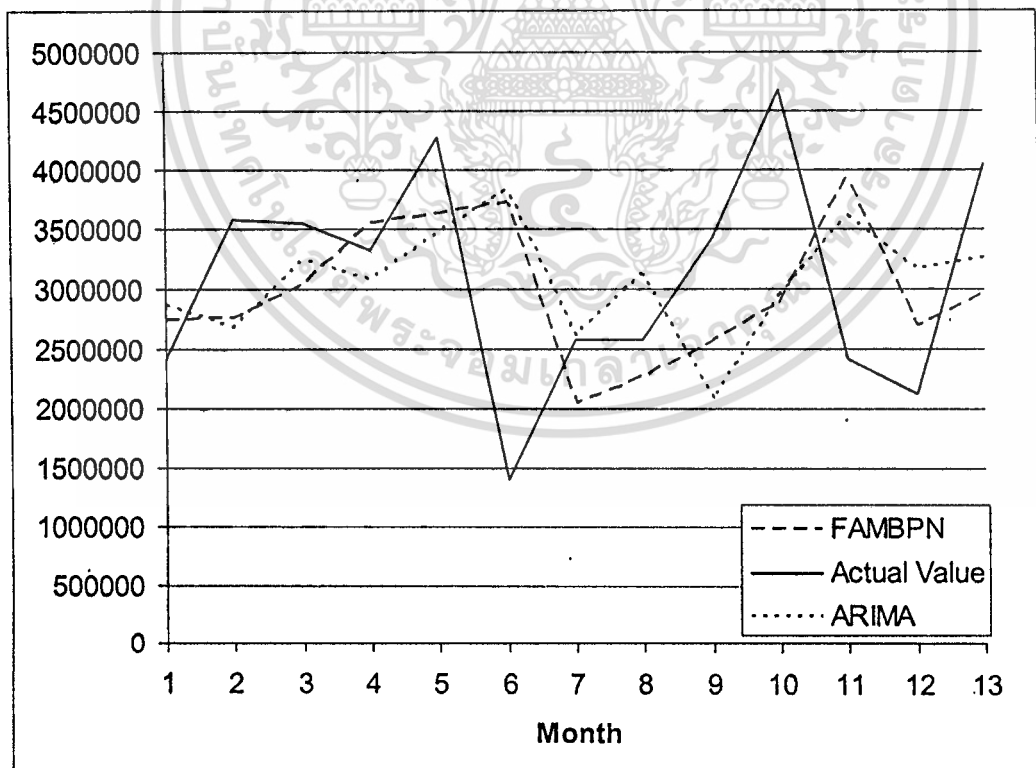


รูปที่ 4.22 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

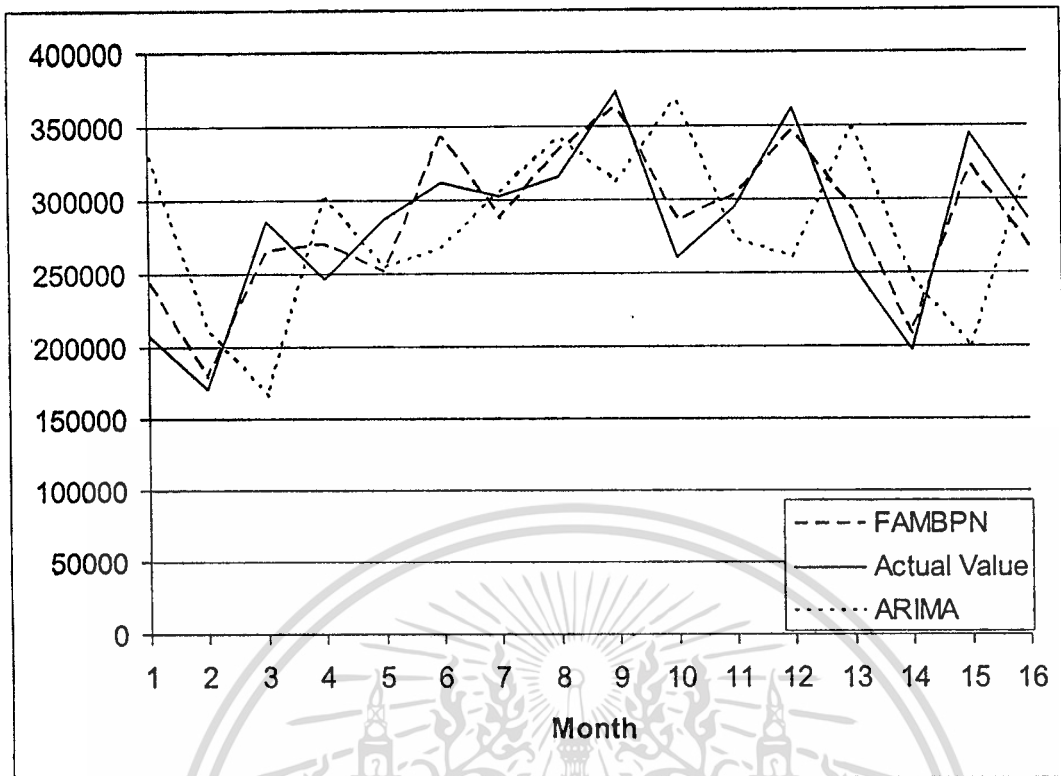


รูปที่ 4.23 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 5

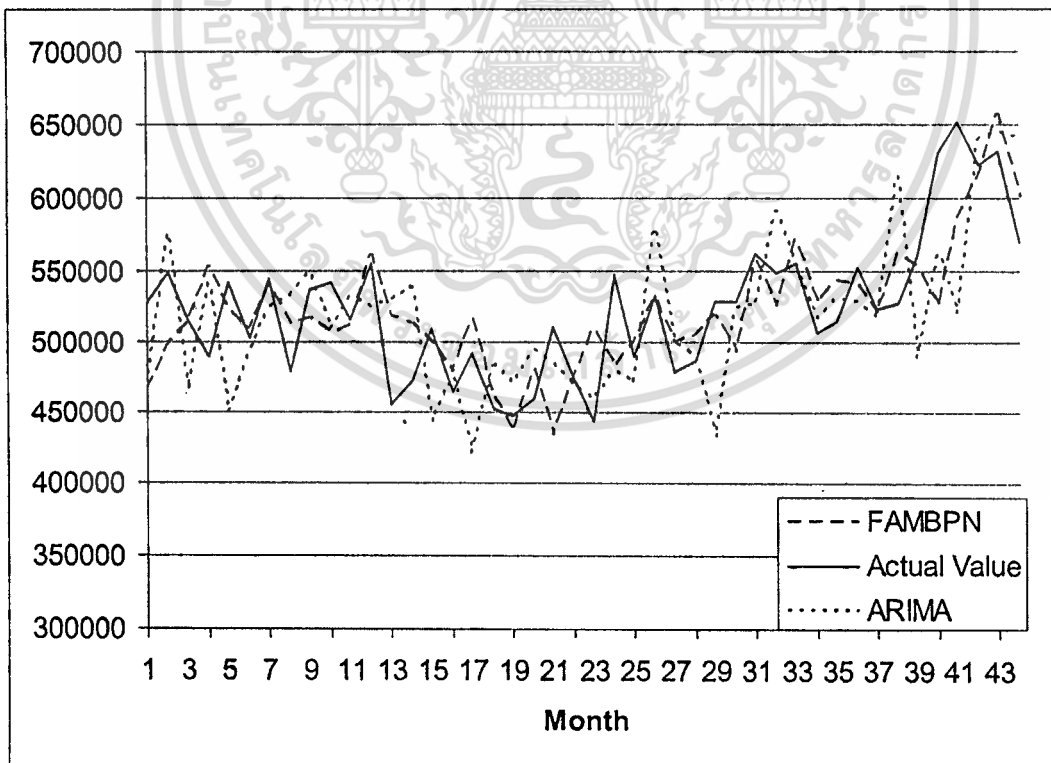


รูปที่ 4.24 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

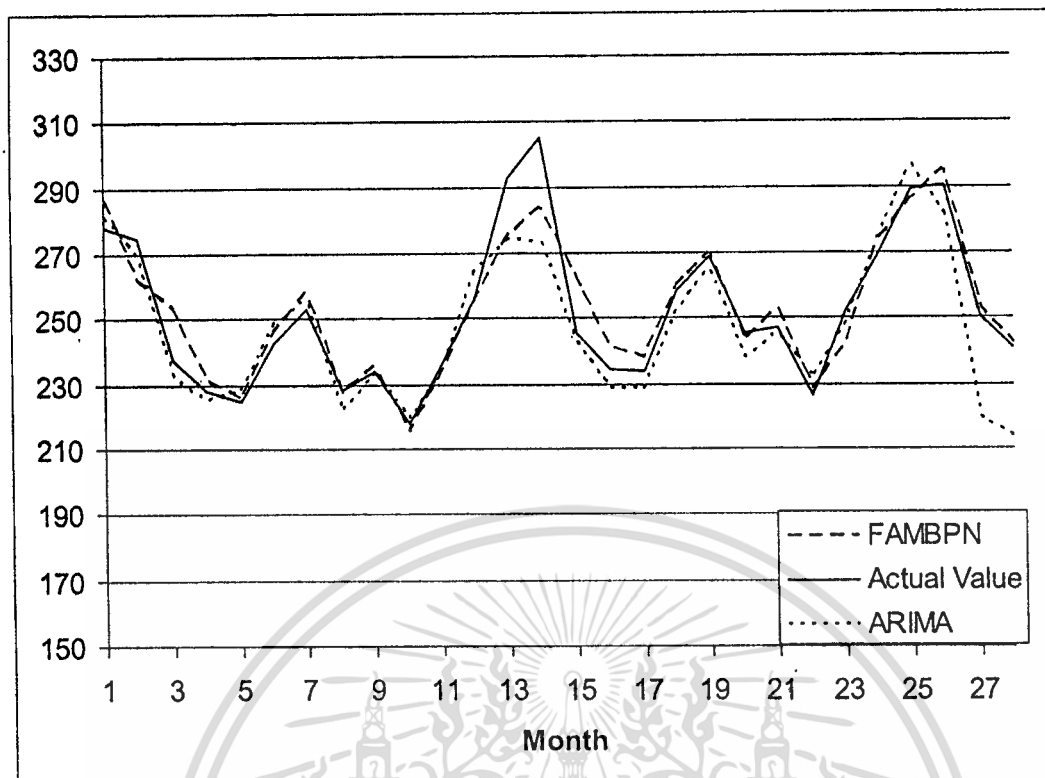


รูปที่ 4.25 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 7

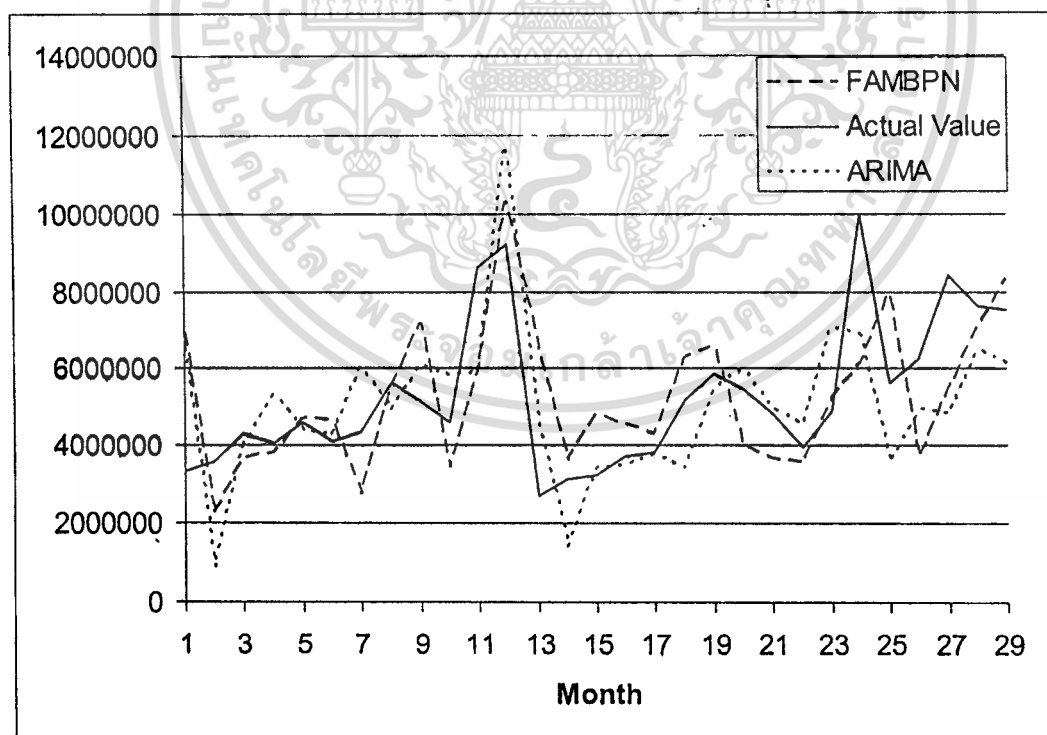


รูปที่ 4.26 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.27 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 9



รูปที่ 4.28 เปรียบเทียบผลการทำนายโดยโมเดลแบบผสมและ ARIMA ของข้อมูลทดสอบชุดที่ 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในการวิจัยนี้ได้นำเสนอวิธีการทำนายข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบผสมที่ได้จากการนำเอาโครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพและโครงข่ายประสาทเทียมแบบแคพรอพพาเกชันมารวมกัน โดยมีแนวคิดที่ว่าหากสามารถทำนายทิศทางกรเปลี่ยนแปลงของอนุกรมเวลาได้ถูกต้องก็น่าจะสามารถทำนายค่าที่เกิดขึ้นในอนาคตได้ถูกต้องมากขึ้นด้วย โดยโมเดลแบบผสมนี้แบ่งการทำนายออกเป็น 2 ส่วนคือ ส่วนแรกใช้เรียนรู้และทำนายทิศทางของแพทเทินอินพุทซึ่งส่วนนี้ใช้โครงข่ายประสาทเทียมแบบซิมพลิไฟด์พีชชี้อาทแมพ ส่วนที่ 2 ใช้ทำนายค่าในอนาคตของแพทเทินอินพุทซึ่งส่วนนี้ใช้โครงข่ายประสาทเทียมแบบแคพรอพพาเกชันจำนวน 2 โครงข่าย โดยโครงข่ายหนึ่งใช้เรียนรู้และทำนายค่าของแพทเทินอินพุทที่คำตอบอยู่ในกลุ่มทิศทางขึ้นและอีกโครงข่ายหนึ่งใช้เรียนรู้และทำนายค่าของแพทเทินอินพุทที่คำตอบอยู่ในกลุ่มทิศทางลง ถ้าส่วนทำนายทิศทางทำนายว่าคำตอบมีทิศทางขึ้น ก็จะส่งไปทำนายในส่วนทำนายค่าที่ใช้ทำนายคำตอบที่มีทิศทางขึ้น และถ้าส่วนทำนายทิศทางให้คำตอบว่าคำตอบมีทิศทางลงก็จะส่งไปทำนายในส่วนที่ใช้ทำนายคำตอบที่มีทิศทางลง แต่ถ้าส่วนทำนายทิศทางทำนายว่าคำตอบมีทิศทางไปไม่เปลี่ยนแปลงจะใช้ค่าก่อนจุดที่ทำนายในอนุกรมเวลาเป็นคำตอบ

ในการทดลองนี้ได้เปรียบเทียบผลการทดลองกับการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแคพรอพพาเกชัน ซึ่งใช้สถาปัตยกรรมและกำหนดค่าพารามิเตอร์เหมือนกับสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบแคพรอพพาเกชันในโมเดลแบบผสม จากการเปรียบเทียบผลการทดลองพบว่าการทำนายข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบผสมมีความถูกต้องในการทำนายทิศทางได้ดีกว่าและให้ค่าความผิดพลาดในการทำนายต่ำกว่าการทำนายโดยใช้โครงข่ายประสาทเทียมแบบแคพรอพพาเกชันถึง 8 ชุดข้อมูลจากทั้งหมด 10 ชุดข้อมูล

นอกจากนี้ยังไปได้เปรียบเทียบผลการทำนายโดยใช้โมเดลแบบผสมกับการทำนายด้วยวิธีทางสถิติคือ Autoregressive Integrated Moving Average (ARIMA) ซึ่งเป็นวิธีทางสถิติที่ถือว่ามีความแม่นยำสูงในการทำนายอนุกรมเวลาระยะสั้น ในการทดลองนี้ได้ใช้โปรแกรม Minitab for Windows : version 14.12 ในการกำหนดรูปแบบและประมาณค่าพารามิเตอร์ จากการเปรียบเทียบผลการทดลองปรากฏว่าทำนายข้อมูลอนุกรมเวลาโดยใช้โมเดลแบบผสมมีความถูกต้องในการทำนายสูงกว่าการทำนายโดยใช้วิธี ARIMA ในทุกชุดข้อมูล

จากผลการทดลองแสดงให้เห็นว่า โมเดลแบบผสมที่นำเสนอมีความถูกต้องแม่นยำในการทำนายข้อมูลอนุกรมเวลามากพอที่จะนำไปใช้ในการทำนายข้อมูลอนุกรมเวลาได้

5.2 ข้อเสนอแนะ

โมเดลแบบผสมที่เสนอนี้ มีจุดที่น่าสนใจคือสามารถทำนายทิศทางของอนุกรมเวลาได้ ซึ่งสามารถบอกแนวโน้มของอนุกรมได้ว่าจะเปลี่ยนไปในทิศทางใด ซึ่งทิศทางที่เป็นไปได้มี 3 ทิศทางคือ ทิศทางขึ้น ทิศทางลงและทิศทางไม่เปลี่ยนแปลง ส่วนข้อเสียโมเดลนี้คือ ค่าความผิดพลาดในการทำนายจะขึ้นอยู่กับความถูกต้องในการทำนายทิศทาง นั่นคือ ถ้าโมเดลสามารถทำนายทิศทางได้ถูกต้องสูงโอกาสของค่าความผิดพลาดในการทำนายก็จะต่ำ แต่ทั้งนี้ค่าความผิดพลาดในการทำนายยังขึ้นอยู่กับข้อกำหนดค่าพารามิเตอร์ของส่วนทำนายค่าหรือส่วนของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันด้วย แนวทางในการพัฒนาต่อไปก็คือพัฒนาให้ส่วนทำนายทิศทางให้มีความถูกต้องมากขึ้น ซึ่งน่าจะส่งผลให้ค่าความผิดพลาดในการทำนายต่ำตามไปด้วยหรืออาจเปลี่ยนวิธีการทำนายค่า

ปัญหาที่พบในการวิจัยคือ การกำหนดค่าพารามิเตอร์สำหรับโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน เช่น จำนวนรอบในการเรียนรู้ซึ่งต้องกำหนดให้เหมาะสม เพราะถ้าใช้จำนวนรอบในการเรียนรู้มากอาจทำให้โครงข่ายเรียนรู้ข้อมูลชุดเรียนรู้ได้ดีเกินไปซึ่งเมื่อนำข้อมูลชุดทดสอบมาทดสอบอาจให้ค่าความผิดพลาดในการทำนายสูง และการกำหนดจำนวนฮิดเดนนิวรอนในชั้นฮิดเดนไม่มีกฎเกณฑ์ในการกำหนด ดังนั้น ต้องอาศัยการลองผิดลองถูกเพียงอย่างเดียวทำให้ใช้เวลาในการทดลองนาน

เอกสารอ้างอิง

- [1] Singh S. and Fieldsend F. "Pattern Matching and Neural Networks based Hybrid Forecasting System." Proc. International Conference on Advance in Pattern recognition (ICAPR2001), Rio de Janeiro, Brazil, March, 2001. pp. 11-14.
- [2] Ricardo B. C. Prudencio and Teresa B. Ludermir. "Design of Neural Networks for Time Series Prediction Using Case-Initialized Genetic Algorithms." Proc. the 8th International Conference on Neural Information Processing, ICONIP' 01, Shanghai, China, 2001.
- [3] Corchado J.M. and Aiken J. "Hybrid Artificial Intelligence Methods in Oceanographic Forecast Models." IEEE Transactions on System, Man and Cybernetics, vol.3, issue 4, November, 2002. pp. 307-313.
- [4] Makridakis, S., Wheelwright, S., and Hyndman, R., Forecasting: Methods and Applications. 3rd Ed. Wiley, 1998
- [5] Rob Hyndman and Muhammad Akram. "Time Series Data Library." [Online]. Available : <http://www-personal.buseco.monash.edu.au/~hyndman/Forecasting/gotodata.html>. 2004.
- [6] Rob Hyndman and Muhammad Akram. "Time Series Data Library." [Online]. Available : <http://www-personal.buseco.monash.edu.au/~hyndman/TSDL>. 2004.
- [7] ทรงศิริ แต่สมบัติ. เทคนิคการพยากรณ์เชิงปริมาณ. กรุงเทพมหานคร : หจก. สำนักพิมพ์ ฟิสิกส์เซ็นเตอร์. 2539.
- [8] Carpenter G.A., Grossberg S., Markuzon N., Reynold J.H., and Rosen D.B. "Fuzzy ARTMAP : A neural Network Architecture for incremental supervised learning of Analog Multidimensional Maps." IEEE Transactions on Neural Networks, vol.3, issue 5, September, 1992. pp. 698-713.
- [9] Carpenter G.A., Grossberg S. and Rosen D.B. "Fuzzy ART : An Adaptive Algorithm for Rapid, Stable Classification of Analog Patterns." International Joint Conference on Neural Networks; vol.2, July, 1991., pp. 411-416.
- [10] Kasuba T. "Simplified Fuzzy ARTMAP", AI Expert, November, 1993. pp. 18-25.
- [11] Negnevitsky M. Artificial Intelligence : A Guide to Intelligence Systems. Harlow : Addison-Wesley publishing Company, Inc. 2002.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [12] Bowerman L. Bruce and O'Connell T. Richard, *Forecasting and Time Series : an Applied Approach*. 3rd ED. Belmont : Duxbury press. 1993.
- [13] Wang, Li-Xin, *Adaptive Fuzzy System and Control*. Eaglewood Cliffs : Prentice-Hall, Inc. 1994.
- [14] International Institute of Forecasters. "Time Series Data." [Online]. Available : <http://www.ms.ic.ac.uk/iif/data/index.htm>. 2004.
- [15] กรมการส่งออก กระทรวงพาณิชย์. "ข้อมูลการส่งออกปลาทูน่า." กรมการส่งออก. 2546



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Arit Thammano and Jindarat Juntaupalee . "Hybrid Time Series Forecasting System Using Simplified Fuzzy ARTMAP and Backpropagation Neural Networks" in Proceedings of The 2004 International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2004). Sendai/Matsushima, Japan. July, 2004.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Hybrid Time Series Forecasting System Using Simplified Fuzzy ARTMAP and Backpropagation Neural Networks

Arit Thammano and Jindarat Juntaupalee
 Faculty of Information Technology
 King Mongkut's Institute of Technology Ladkrabang
 Ladkrabang, Bangkok 10520 Thailand
 Tel.: (662)737-2551-4, Fax.: (662)326-4332
 Email: arit@it.kmitl.ac.th, Jindarat_j@yahoo.com

Abstract: This paper proposes a new hybrid time series forecasting system using simplified fuzzy ARTMAP and backpropagation neural networks. The simplified fuzzy ARTMAP, a modified version of the well-known fuzzy ARTMAP, is used to decide the direction of the time series data whether it is up, down or sideways, and then use two backpropagation neural networks to forecast the magnitude of direction change. This proposed model was tested with the M2-competition time series data. The results obtained are very encouraging.

1. Introduction

Forecasting is the basis for planning and decision making. The more accurate the organization's forecasts are, the better prepared it will be to take advantage of future opportunities and to reduce potential risks. The science of forecasting relies heavily on the models used for forecasting, quantity and quality of data, and the ability to pick the right models for a given data. Artificial neural networks, especially the backpropagation neural network, have dramatically gained a lot of attention from researchers in various areas, such as pattern recognition, classification, and time series forecasting. The backpropagation neural network has proven in practice to be quite efficient and has been widely used in time series forecasting, such as in [1, 2]. Fuzzy ARTMAP has also proven in practice to be quite efficient in pattern recognition and classification applications [3, 4]. However, the trend of the forecasting technique is now moving toward hybrid models. Many researchers have proposed hybrid models for forecasting such as [5, 6, 7]. These researches show that the hybrid systems can improve the accuracy of the forecasting over the systems that use only a single method. In this paper, the simplified fuzzy ARTMAP and the backpropagation neural network are combined to generate a hybrid model for forecasting.

Following this introduction, section 2 briefly describes the simplified fuzzy ARTMAP component, the backpropagation neural network component, and the architecture of the proposed hybrid model. The description of the tested time series data is described in section 3. In section 4, the experimental setup and results are discussed. Finally, section 5 is the conclusion.

2. Hybrid Forecasting System

2.1. Simplified Fuzzy ARTMAP Component

Simplified Fuzzy ARTMAP (SFAM), proposed by Tom Kasuba in 1993, is a simplification of fuzzy ARTMAP. SFAM is developed from the fuzzy ARTMAP in order to reduce the computational overhead and architecture redundancy of the fuzzy ARTMAP. Fuzzy ARTMAP, introduced by Grossberg et al. in 1992, is a supervised neural network with very interesting properties, such as a very fast convergence, a capacity of incremental learning of recognition categories, and multidimensional mapping in response to input vectors [8, 9].

SFAM comprises three layers: the input layer, the cluster layer, and the category layer. Each node in the cluster layer is fully connected to the nodes in the input layer via a set of top-down weights. The nodes in the category layer represent the categories or classes that the network has to learn to recognize. First, the input pattern A is flowed through the complement coder where the input string is stretched to double the size by adding its complement A^c . The complement coded input I , which is equal to (A, A^c) , is then applied to the input layer. At this stage, all the cluster nodes become active, yet with different degrees. The cluster node with the highest activation is then selected as the winning node.

$$T_j(I) = \frac{|I \wedge W_j|}{\alpha + |W_j|} \quad (1)$$

$$\text{Winning Node } J = \max(T_j) \quad (2)$$

where α is a very small number close to zero.

W_j is the weight vector of the cluster node j .

If more than one T_j are maximal, the cluster node with the smallest index is chosen. Next, the match function is calculated to measure the degree of matching between the winning cluster node and the input pattern. The match function is defined as

$$MF(I, W_j) = \frac{|I \wedge W_j|}{|I|} \quad (3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

The match function in conjunction with the vigilance parameter ρ states whether the winning cluster node is good enough to encode the current input pattern or whether another cluster node should be formed to code the input pattern. If the match function is greater than or equal to the vigilance parameter, the network is said to be in the state of resonance. Resonance means that cluster node J is good enough to encode the input I . After resonance, the system then checks whether the category node, which is linked to the winning cluster node, corresponds to the current input pattern. If no error occurs, the weight vector W_j is updated according to the following equation:

$$W_j^{new} = \beta(I \wedge W_j^{old}) + (1 - \beta)W_j^{old} \quad (4)$$

Otherwise, the activation of the winning cluster node is reset to zero and the vigilance parameter is increased just enough to exceed the match function value. The new winning cluster node J is chosen. The search process continues until the criterion $MF(I, W_j) \leq \rho$ is respected and the chosen node J corresponds to the current input pattern. If such node does not exist, a new cluster node is recruited to code the input pattern; the association between a new cluster node and the category node is created. The answer obtained from this component is whether the direction of the series is up, down or sideways.

2.2. Backpropagation Neural Network Component

Backpropagation neural network (BPNN) is a class of multi-layer feedforward neural networks with a supervised learning rule. In this paper, two BPNNs are employed in the second component. The first BPNN is used to forecast the magnitude of change in “up” direction; while the second is used to forecast the magnitude of change in “down” direction. Each BPNN consists of 3 layers: one input layer, one hidden layer, and one output layer. The sigmoidal transfer function is used in the hidden and output neurons.

2.3. The Hybrid Model

The new hybrid model proposed in this paper is composed of 2 components. The first component which is the simplified fuzzy ARTMAP is used to forecast the direction of the time series and the second component which is BPNN is used to forecast the magnitude of change of the time series. If the direction of the next step is forecasted to be up, the input pattern will be sent to the first BPNN, which uses to forecast the magnitude of up direction. If the first component predicts that the next step is down, then the input pattern will be sent to the second BPNN, which uses to forecast the magnitude of down direction. Otherwise, the output of the model is the value of the current step.

3. Time Series Data

To test the performance of the proposed model, the experiments have been conducted on 6 data sets. These data

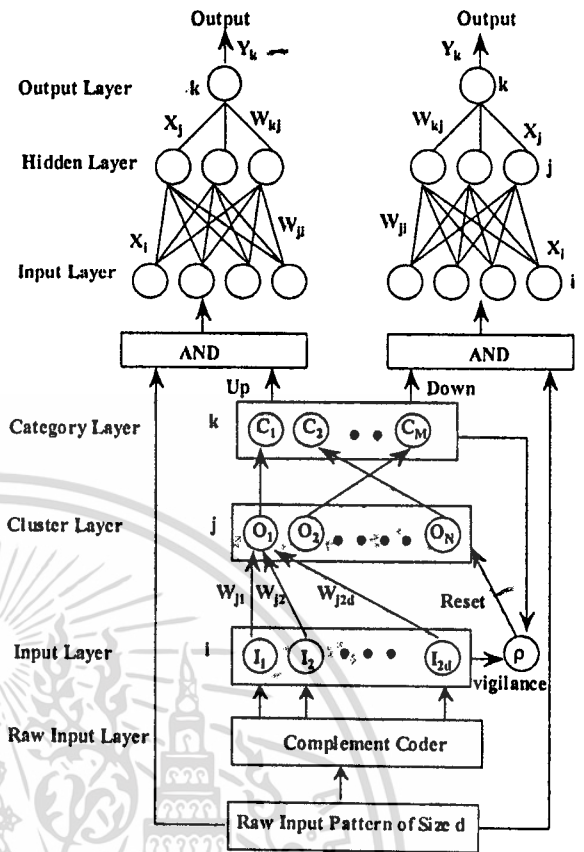


Figure 1. The Architecture of a Hybrid Model

sets are parts of the M2-competition time series data sets available from [10]. The description of the data sets are shown in Table 1 and Figure 2-7.

4. Experimental Results

The data sets used in the experiments are non-stationary. However, the time series data must be transformed to the stationary series before being sent to the forecasting model. The process usually used to remove the trend from the time series data is the differencing. Therefore, the original time series is replaced by its first differences. If a time series is represented by $Y = \{Y_1, Y_2, \dots, Y_{n-1}, Y_n\}$ where n is the total number of data in the series, the first differencing of the time series is defined by $\delta = (\delta_1, \delta_2, \dots, \delta_{n-1})$ where $\delta_i = Y_{i+1} - Y_i, 1 \leq i \leq n-1$. Differencing not only eliminates the trend of the series but also helps reduce the range of the time series.

The first differences are then normalized into values between 0 and 1 by using the following equation:

$$A_i = \frac{\delta_i - \min}{\max - \min} \quad (5)$$

where $\max = (\text{maximum value of } \delta) + 5\%$
 $\min = (\text{minimum value of } \delta) - 5\%$

Table 1. Description of the Time Series

Data set	Description	Observations	Range of Data set
A	Automobile Industry: Car Model: Panther Monthly series (Jan., 1982 – Oct., 1987)	70	3,619.8 - 52,727.2
B	Macro Data: Gross National Product Quarterly series (Q1/ 1947 – Q3/ 1987)	163	1,056.5 - 3,796.4
C	Honeywell: Total OEM Sales Monthly series (Jan., 1981 – Oct., 1987)	82	24,611 - 222,062
D	Squibb: Product Sales Monthly series (Jan., 1982 – Oct., 1987)	70	252,360 - 4,685,079
E	Aussedat-rey: European Coated-paper Orders Monthly series (Jan., 1981 – Oct., 1987)	82	142,827 - 397,034
F	Aussedat-rey: US Uncoated-paper Orders Monthly series (Jan., 1969 – Sept., 1987)	225	222,000 - 651,000

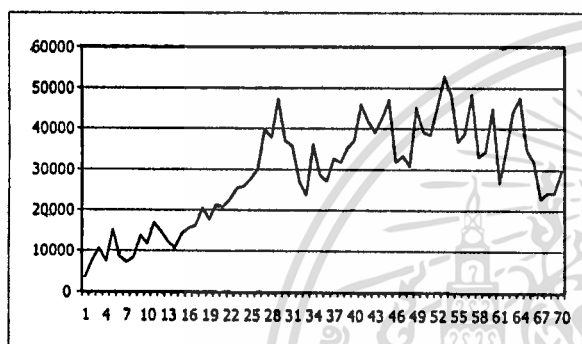


Figure 2. Data Set A

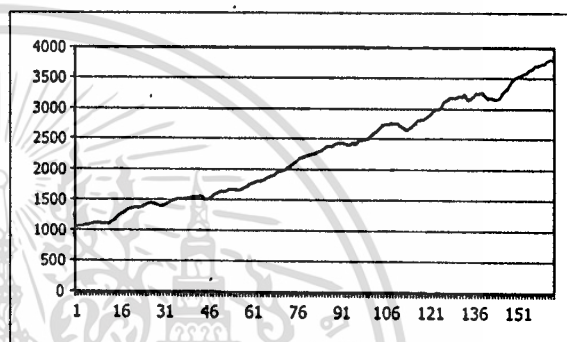


Figure 3. Data Set B

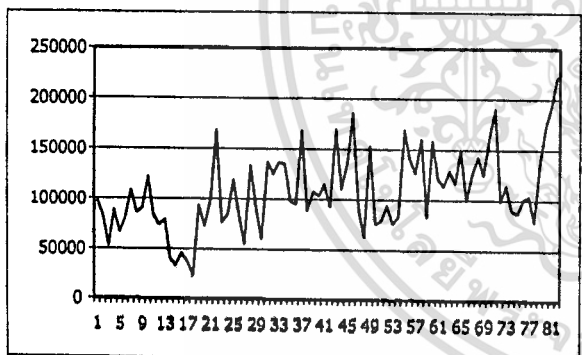


Figure 4. Data Set C

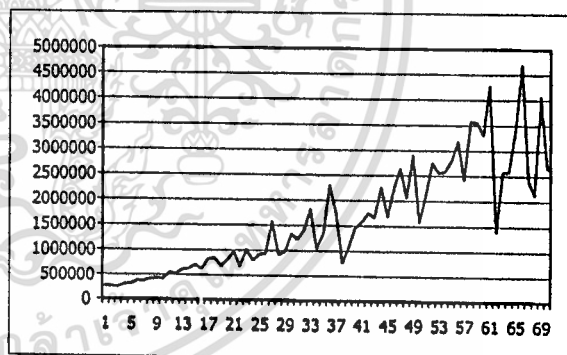


Figure 5. Data Set D

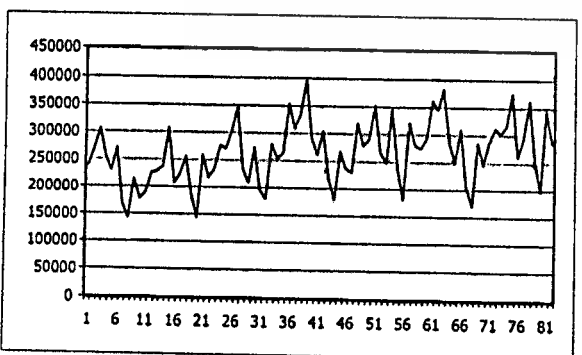


Figure 6. Data Set E

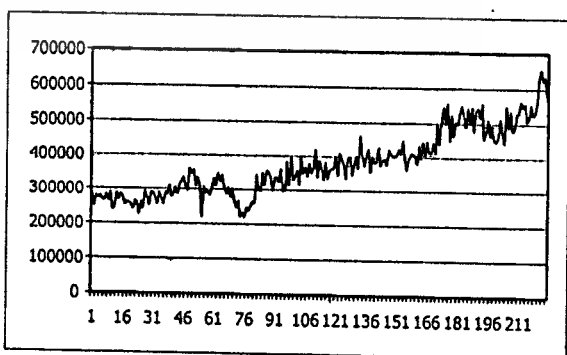


Figure 7. Data Set F

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

After being scaled, the input patterns of size k are constructed by taking k consecutive values from those normalized series. In the experiments, k is varied from 2 to 6 to find the pattern size that gives the best result. The desired output of SFAM is the direction of the series, i.e. [1 0 0] representing "up" direction, [0 0 1] representing "down" direction, and [0 1 0] representing "unchanged" direction. However, the desired output of BPNN is the $k+1$ value.

Next, the data sets are divided into two distinct sets, the first 80% is used as the training data and the remaining 20% is used as the testing data. The training set is trained until the best performance on the training data set is obtained. During the training phase of the first component, the learning rate is set at 0.1, while the vigilance parameter is varied from 0.75 - 1.0 to find the vigilance value that gives the best result on the training data set.

BPNN component consists of one input layer, one hidden layer, and one output layer. According to [11], only one hidden layer is sufficient to approximate the continuous function. The number of input neuron is k , while the number of output neuron is one. The number of hidden neuron is varied from 2 to 8. In the experiments, the learning rate is set at 0.1. The training data sets are divided into 2 sets according to their directions. If the direction of the input pattern is up, it will be sent to the "up" BPNN. However, if the direction of the input pattern is down, it will be sent to the "down" BPNN.

This proposed model is benchmarked against the standard BPNN. To compare the performance of both methods, the Root Mean Square Error (RMSE) and the percentage of direction correct are used as the measures. Table 2 is the results from the experiments; it shows that the results obtained from the hybrid model outperform the results obtained from BPNN on both percentage of direction correct and RMSE.

5. Conclusion

This paper proposes a new hybrid time series forecasting model, which is the combination of simplified fuzzy ARTMAP and backpropagation neural network. The results obtained from the proposed model outperform the results obtained from a basic backpropagation neural network.

References

- [1] Kolarik, T. and Rudorfer, G., "Time Series Forecasting Using Neural Networks," *APL Quote Quad*, 25(1), 1994, pp. 86-94.
- [2] Chenoweth, T., Obradovic, Z., and Lee, S. S., "Embedding Technical Analysis into Neural Network Based Trading Systems," *Applied Artificial Intelligence*, vol. 10, no. 6, 1996, pp. 523-541.
- [3] Song, X. H., Hopke, P. K., Bruns, M. A., Bossio, D. A., and Scow, K. M., "A Fuzzy Adaptive Resonance Theory - Supervised Predictive Mapping Neural Network Applied to the Classification of Multivariate Chemical Data," *Chemom Intel Lab Syst.*, 1998, pp. 161-170.
- [4] Busque, M. and Parizeau, M., "A Comparison of Fuzzy ARTMAP and Multilayer Perceptron for Handwritten Digit Recognition," *Computer Vision and Systems Laboratory*, Université Laval, Canada, 1997.
- [5] Singh, S. and Fieldsend, J. E., "Pattern Matching and Neural Networks Based Hybrid Forecasting System," *Proceedings of the International Conference on Advances in Pattern Recognition*, Rio de Janeiro, Brazil, March, 2001, pp. 11-14.
- [6] Corchado, J. M. and Aiken, J., "Hybrid Artificial Intelligence Methods in Oceanographic Forecast Models," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 32, no. 4, November, 2002, pp. 307-313.
- [7] Ricardo, B. C. P. and Teresa, B. L., "Design of Neural Networks for Time Series Prediction Using Case-Initialized Genetic Algorithms," *Proceedings of the 8th International Conference on Neural Information Processing*, Shanghai, China, 2001.
- [8] Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H., and Rosen, D. B., "Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, September, 1992, pp. 698-713.
- [9] Kasuba, T., "Simplified Fuzzy ARTMAP," *AI Expert*, 8(11), November, 1993, pp. 18-25.
- [10] <http://www.ms.ic.ac.uk/iif/index.html>
- [11] Hornik, K., Stinchcombe, M. and White, H., "Multilayer Feedforward Networks are Universal Approximators," *Neural Networks*, vol. 2, 1989, pp. 359-366.

Table 2. Results of the Experiments

Data Set	Observations	Pattern Size (k)	Hybrid Model			BPNN		
			% Direction Correct	RMSE	No. of Hidden Neurons of BPNN Component	% Direction Correct	RMSE	No. of Hidden Neurons
A	70	5	69.23	8,595.80	4	53.85	12,709.00	6
B	163	4	78.13	36.79	8	75.00	40.23	6
C	82	5	75.00	42,437.00	5	62.50	36,299.00	3
D	70	6	69.23	1.1561e+6	6	84.61	1.1593e+6	6
E	82	5	93.75	31,225.00	3	87.50	35,053.00	3
F	225	4	77.27	37,779.00	8	70.45	38,727.00	6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาวจินดารัตน์ จันทอุปพี
วัน เดือน ปีเกิด	18 สิงหาคม พ.ศ.2513 ที่จังหวัดกาฬสินธุ์
ที่อยู่	130/1 ถนนแก่งสำโรง ตำบลกาฬสินธุ์ อำเภอเมือง จังหวัดกาฬสินธุ์ 46000
ประวัติการศึกษา	2534 ประกาศนียบัตรการศึกษามัธยมศึกษา สาขาการประถมศึกษา (เกียรตินิยม อันดับ 1) มหาวิทยาลัยศรีนครินทรวิโรฒมหาสารคาม 2541 ครุศาสตรบัณฑิต สาขาคอมพิวเตอร์ศึกษา สถาบันราชภัฏมหาสารคาม
ประสบการณ์ทำงาน	อาจารย์ สังกัดวิทยาลัยพลศึกษาจังหวัดมหาสารคาม สำนักงานพัฒนาการ กีฬาและนันทนาการ กระทรวงการท่องเที่ยวและกีฬา
ผลงานวิจัย	- Arit Thammano and Jindarat Juntauppalee. "Hybrid Time Series Forecasting System Using Simplified Fuzzy ARTMAP and Backpropagation Neural Networks" in Proceedings of The 2004 International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2004). Sendai/Matsushima, Japan. July, 2004.