ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

ระบบสนับสนุนการตัดสินใจ
สำหรับลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย
โดยใช้เทคโนโลยีดาต้าไมนิ่ง

Decision Support System
For invests in Stock Exchange of Thailand
by Data Mining Technology

โดย

นายธีรพงษ์ วรวงศ์

รหัส 45061741

H003148

อาจารย์ที่ปรึกษา

รศ.ดร.อาริต ธรรมโน

วัน เดือน ปี 09 พ.ค. 2550
เลขทะเบียน 03148
เลขเรียกหนังสือ วพ. ธ 6315 2547
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการศึกษากรณีพิเศษ
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
ภาคเรียนที่ 1 ปีการศึกษา 2547
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้ สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	ระบบสนับสนุนการตัดสินใจ สำหรับลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้เทคโนโลยีดาต้าไมนิ่ง
นักศึกษา	นายธีรพงษ์ วรวงศ์
อาจารย์ที่ปรึกษา	รศ.ดร.อาริต ธรรมโน
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	การจัดการเทคโนโลยีสารสนเทศ
ปีการศึกษา	2547

บทคัดย่อ

การนำเทคโนโลยีสารสนเทศในแขนงของการทำดาต้าไมนิ่ง มาใช้พยากรณ์อนาคตการเปลี่ยนแปลงของราคาหลักทรัพย์เฉพาะแต่ละตัวหลักทรัพย์ ในตลาดหลักทรัพย์แห่งประเทศไทยนั้น มีวิธีการโดยนำข้อมูลหรือปัจจัยที่มีผลต่อการเปลี่ยนแปลงราคาหลักทรัพย์ในอดีต ซึ่งข้อมูลเหล่านี้มีลักษณะเป็นข้อมูลอนุกรมเวลา มีความสลับซับซ้อน และมีความสัมพันธ์ระหว่างกันในหลายๆมิติ มาผ่านกระบวนการทำดาต้าไมนิ่งเพื่อสร้างแบบจำลองพยากรณ์ฯ โดยใช้อัลกอริทึมแบล็คพร็อบพาเกชั่น หลังจากที่ได้แบบจำลองฯดังกล่าวแล้ว จะนำมาทดสอบกับข้อมูลที่เกิดขึ้นจริง และจะทำกระบวนการดังกล่าวซ้ำๆกัน เพื่อสรรหาแบบจำลองฯที่มีถูกต้องแม่นยำที่สุด

หลังจากนั้นจะนำแบบจำลองฯดังกล่าว ไปเปรียบเทียบกับเครื่องมือหรือการวิเคราะห์ทางเทคนิคอื่นๆ ที่ใช้กันอยู่ในปัจจุบัน เช่น แบบจำลองพยากรณ์ฯที่ใช้อัลกอริทึมแบล็คพร็อบพาเกชั่น โดยใช้เฉพาะข้อมูลราคาหลักทรัพย์ในอดีตที่มีลักษณะอนุกรมเวลาเท่านั้น เพื่อพิสูจน์ว่าเทคโนโลยีฯเดียวกันแต่ใช้ข้อมูลที่ครอบคลุมมากกว่า จะมีประสิทธิภาพสูง และมีความน่าเชื่อถือมากกว่า เพื่อเป็นประโยชน์ต่อผู้ที่สนใจลงทุนในตลาดหลักทรัพย์ ทั้งที่ลงทุนแล้วและผู้ที่สนใจจะลงทุน ได้มีเครื่องมือหรือระบบที่สามารถสร้างข้อมูลสารสนเทศเพื่อสนับสนุนการตัดสินใจลงทุนในตลาดทุนได้อย่างมั่นใจ รวมถึงผู้ที่สนใจศึกษาเทคโนโลยีสารสนเทศในแขนงดังกล่าว ได้เข้าใจและทราบถึงประโยชน์อย่างชัดเจนยิ่งขึ้น อีกทั้งสามารถนำไปประยุกต์ใช้กับระบบอื่นๆได้ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	Decision Support System for invests in Stock Exchange of Thailand by Data Mining Technology
Student	Mr.Teerapong Woarawong
Advisor	Assoc.Prof.Dr. Arit Thammano
Level of Study	Master of Science in Information Technology
Major	Information Technology Management
Academic Year	2004

ABSTRACT

The process of applying data mining software as a crucial tool to project the price change of the stock, registed in the Stock Exchange of Thailand, starts with inputting the sets of time series data into the program. The sets of data, which, we think, effect the price change of the stock. Each set of the data may or may not have the relation to each other. After the program that's algorithm Backpropagation, proceeds those sets of data, it will create both predictive model, and the output. Then we have to which will be tested by comparing the output with the real sets of data Then we going to repeat program process for the best predictive model.

After that we will compare the best predictive model with other technique tools example Backpropagation 's predictive model was created form only stock price time series data, It is prove itself very effective and reliable more than. It will be a new tools or decision support system for investor invests with confidently and effectively into stock exchange of Thailand. And we hope this project. It is very usefully for people develop and integrate for them systems.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ในการพัฒนาระบบสนับสนุนการตัดสินใจการลงทุนหลักทรัพย์ ในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้เทคโนโลยีดาต้าไมนิ่งฉบับนี้ ผู้จัดทำขอขอบพระคุณ รศ.ดร.อาริต ธรรมโน ที่ได้กรุณาให้แนวทางการศึกษา ตลอดจนให้คำปรึกษาตลอดช่วงการดำเนินงาน จนสามารถพัฒนาระบบได้สำเร็จลุล่วงได้ด้วยดี

ขอขอบคุณนายประสิทธ์ เสาวรัติธาดา กรรมการผู้จัดการบริษัท อินชัวร์เนต จำกัด ที่ให้โอกาสและสนับสนุนการศึกษาตลอดหลักสูตร

สุดท้ายขอขอบคุณทุกๆกำลังใจทั้งจากบุคคลในครอบครัว และเพื่อนๆ ITM 12 ที่เป็นส่วนช่วยให้สามารถแก้ไขปัญหา และผ่านอุปสรรคต่างๆได้ด้วยดี

ธีรพงษ์ วรวงศ์

11 กันยายน 2547

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย I

บทคัดย่อภาษาอังกฤษ II

กิตติกรรมประกาศ III

สารบัญ IV

สารบัญรูป VII

สารบัญตาราง IX

บทที่

1. บทนำ 1

1.1 ความเป็นมาของโครงงาน 1

1.2 วัตถุประสงค์ของโครงงาน 2

1.3 ขอบเขตของโครงงาน 3

1.4 แนวทางการดำเนินงาน 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

2. ทฤษฎีเบื้องต้นของดาต้าไมนิ่ง 4

2.1 หลักการทั่วไปของดาต้าไมนิ่ง 4

2.1.1 การกำหนดวัตถุประสงค์ทางธุรกิจ 5

2.1.2 การคัดเลือกข้อมูล 6

2.1.3 การกรองข้อมูล 6

2.1.4 การแปลงรูปแบบข้อมูล 6

2.1.5 การทำไมนิ่งข้อมูล 7

2.1.6 การวิเคราะห์และประเมินผลลัพธ์ 8

2.2 ประเภทของการใช้งานดาต้าไมนิ่ง 9

2.2.1 งานด้านการพยากรณ์ (Predictive Data Mining) 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

บทที่

2.2.2 งานด้านการแบ่งกลุ่ม (Descriptive Data Mining) 10

2.3 การประยุกต์ใช้ดาต้าไมนิ่งกับงานพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ 11

2.3.1 การวิเคราะห์ข้อมูลในอนุกรมเวลา (Time Series Analysis) 11

2.3.2 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing) 13

3. หลักการของโครงข่ายประสาทเทียม (Artificial Neural Network) 15

3.1 ลักษณะการทำงานของโครงข่ายประสาทเทียม 16

3.2 โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม 18

3.2.1 การคำนวณในโครงสร้างในโครงข่ายประสาทเทียม(Neural Computation) 19

3.2.2 การคำนวณของโหนดในโครงข่ายประสาทเทียม 19

3.2.3 อัลกอริทึมที่ใช้ในการสอน (Training Algorithm) 21

3.2.4 การวัดค่าความผิดพลาด (Error) 24

3.2.5 การพิจารณาจำนวนโหนดในชั้นอินพุตและชั้นกลาง 25

3.3 ประเภทของโครงข่ายประสาทเทียม 26

3.3.1 โครงสร้างแบบส่งสัญญาณไปข้างหน้า (Feed Forward Neural Network) 26

3.3.2 โครงสร้างแบบมีการป้อนย้อนกลับ (Recurrent Neural Network:RNN) 27

3.4 ขั้นตอนการพัฒนาโครงข่ายประสาทเทียมสำหรับระบบพยากรณ์ราคาหลักทรัพย์ 29

4. การพัฒนาระบบพยากรณ์ราคาหลักทรัพย์ 31

4.1 องค์ประกอบของระบบพยากรณ์ราคาหลักทรัพย์ 31

4.2 การใช้งานระบบพยากรณ์ราคาหลักทรัพย์ 33

4.3 การใช้งานระบบพยากรณ์ราคาหลักทรัพย์ สำหรับลงทุนในตลาดหลักทรัพย์ฯ 41

5. การทดสอบระบบ และสรุป 43

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

หน้า

บทที่

5.1 การทดสอบระบบ 43

5.2 สรุปผลการทดสอบ 46

บรรณานุกรม 47

ประวัติผู้เขียน 48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

หน้า

รูปที่

2.1 แสดงขั้นตอนต่างๆของกระบวนการทำดาต้าไมนิ่ง 5
2.2 แสดงตัวอย่างของโครงข่ายประสาทเทียม (Neural Network) 7
2.3 แสดงตัวอย่างของโครงสร้างการตัดสินใจแบบต้นไม้ (Decision Tree) 8
3.1 แสดงโครงสร้างเซลล์ประสาทของมนุษย์ 17
3.2 แสดงโครงสร้างพื้นฐานของแบบจำลองโครงข่ายประสาทเทียม 18
3.3 แสดงแบบจำลองการคำนวณของโหนดภายในโครงข่ายประสาทเทียม 20
3.4 แสดงรูปแบบของฟังก์ชันกระตุ้นที่ใช้ในการคำนวณค่าในโหนดของโครงข่าย 20
3.5 แสดงโครงสร้างของโครงข่ายประสาทเทียมแบบส่งสัญญาณไปข้างหน้า 26
3.6 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบง่าย-Jordan Network 27
3.7 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบง่าย-Elman Network 28
3.8 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบสมบูรณ์ 29
4.1 แสดงตัวอย่างข้อมูลอินพุตเพื่อใช้ในการสร้างแบบจำลองพยากรณ์ฯ 34
4.2 แสดงตัวอย่างข้อมูลอินพุตที่ผ่านการแปลงค่าข้อมูล 35
4.3 แสดงตัวอย่างข้อมูลราคาปิดหลักทรัพย์อนุกรมเวลา 10 วันที่ผ่านการแปลงค่าข้อมูล 35
4.4 แสดงหน้าจอการนำเข้าข้อมูลเพื่อสร้างแบบจำลองพยากรณ์ฯ 36
4.5 แสดงหน้าจอสำหรับกำหนดจำนวนอินพุตและเอาต์พุต 36
4.6 แสดงหน้าจอการกำหนดค่าตัวแปรต่างๆเพื่อสร้างแบบจำลองพยากรณ์ฯ 37
4.7 แสดงหน้าจอการสร้างแบบจำลองพยากรณ์ฯ 37
4.8 แสดงหน้าจอการทดสอบแบบจำลองพยากรณ์ฯ 38
4.9 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลเรียนรู้ 38
4.10 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลทดสอบ 39
4.11 แสดงตัวอย่างหน้าจอการเรียกใช้ค่าถ่วงน้ำหนักของแบบจำลองพยากรณ์ฯ 39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป(ต่อ)

หน้า

รูปที่

4.12 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลทดสอบในแบบรายงาน....40
4.13 แสดงตัวอย่างไฟล์ที่เก็บรายละเอียด ตัวแปรต่างๆของแบบจำลองพยากรณ์ฯ.....40
4.14 กราฟแสดงแนวโน้มการเปลี่ยนแปลงของราคาหลักทรัพย์ จากแบบจำลองพยากรณ์ฯ....41
5.1 แสดงผลการเปรียบเทียบค่าข้อมูลจริงกับค่าข้อมูลที่ได้จากแต่ละแบบจำลองพยากรณ์ฯ 46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

หน้า

ตารางที่

2.1 แสดงประเภทของงานที่มีการนำเทคนิคต่างๆของดาต้าไมนิ่งมาใช้งาน 9
5.1 แสดงรายละเอียดข้อมูลอินพุตเพื่อใช้ในการสร้างระบบพยากรณ์ราคาหลักทรัพย์ 43
5.2 แสดงผลลัพธ์ที่ได้จากแบบจำลองพยากรณ์ต่างๆ 44
5.3 แสดงราคาปิดที่เกิดขึ้นจริง และราคาที่แต่ละแบบจำลองพยากรณ์ฯคำนวณได้ 45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาของโครงงาน

ในยุคของอัตราดอกเบี้ยเงินฝากต่ำสุดในประวัติการณ์ ผู้มีเงินออม นักลงทุนคงหนีไม่พ้นการลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย ยิ่งได้รับการสนับสนุนจากทั้งภาครัฐและเอกชน อีกทั้งช่องทางในการซื้อขายหลักทรัพย์มีมากและอำนวยความสะดวกมากขึ้น จึงทำให้การลงทุนในหลักทรัพย์เป็นเรื่องที่ไม่ยุ่งยากอีกต่อไป แต่การตัดสินใจลงทุนในหลักทรัพย์ที่ดีนั้นเป็นสิ่งที่ยากยิ่ง

เนื่องจากราคาหลักทรัพย์นั้นจะสะท้อนถึงมูลค่าของบริษัทเจ้าของหลักทรัพย์นั้นๆ ซึ่งการจะวิเคราะห์หรือประเมินมูลค่าของบริษัทใดๆ นั้นเป็นเรื่องที่มีความสลับซับซ้อนอย่างมาก ในปัจจุบันวิธีการที่นิยมใช้คือ การวิเคราะห์โดยใช้ปัจจัยพื้นฐานซึ่งเป็นการวิเคราะห์เชิงปริมาณโดยใช้ทั้งข้อมูลภายใน เช่น การวิเคราะห์งบการเงินของบริษัท และข้อมูลภายนอกบริษัท เช่น สภาวะเศรษฐกิจทั้งภายในและภายนอกประเทศ สภาวะทางการเมือง สภาวะอุตสาหกรรม เพื่อประเมินมูลค่าของหลักทรัพย์ว่าสูงหรือต่ำกว่ามูลค่าที่แท้จริงหรือไม่ อีกวิธีที่นิยมใช้คือการวิเคราะห์ปัจจัยทางด้านเทคนิค ซึ่งเป็นการนำข้อมูลราคาหลักทรัพย์และปริมาณการซื้อขายในอดีต มาสร้างเป็นกราฟเพื่อคำนวณหาราคาหลักทรัพย์ในอนาคต โดยอาศัยทฤษฎีทางด้านสถิติ ความน่าจะเป็น เพราะมีความเชื่อว่าราคาหลักทรัพย์มีลักษณะเป็นข้อมูลเชิงอนุกรมเวลา มีการเปลี่ยนแปลงแบบมีแบบแผนเนื่องจากพฤติกรรมการลงทุนทั่วไปมักจะกระทำซ้ำรอยเดิม

หากเปรียบเทียบความน่าเชื่อถือของการวิเคราะห์โดยใช้ปัจจัยพื้นฐานกับการวิเคราะห์ปัจจัยทางด้านเทคนิคนั้น กลับมีข้อดีและข้อเสียต่างกันคือ หากลงทุนในระยะสั้นการวิเคราะห์ปัจจัยทางด้านเทคนิคมีความน่าเชื่อถือสูงกว่า แต่หากเป็นการลงทุนในระยะปานกลางถึงระยะยาวความน่าเชื่อถือของการวิเคราะห์ปัจจัยโดยใช้ปัจจัยพื้นฐานมีความน่าเชื่อถือมากกว่า ถึงกระนั้นก็ตามการจะใช้การวิเคราะห์ใดๆก็ยังเป็นเรื่องที่ยุ่งยาก สับสน ตลอดจนการเข้าถึงและคัดเลือกข้อมูลจากแหล่งข้อมูลที่แตกต่างกัน และมีความสัมพันธ์กันในหลายๆมิติ จึงเป็นการยากยิ่งที่นักลงทุนจะสามารถวิเคราะห์ราคาหลักทรัพย์ได้เอง จึงได้เพียงแต่รับคำแนะนำ หรือบทวิเคราะห์จากบริษัท นายหน้าค้าหลักทรัพย์ หรือจากข่าวสารหรือข่าวลือที่เกิดขึ้นทั่วๆไปเท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจุบันระบบเทคโนโลยีสารสนเทศในแขนงของดาต้าไมนิ่ง โดยใช้อัลกอริทึมของโครงข่ายประสาทเทียม ได้รับการพัฒนาเพื่อใช้ในงานวิเคราะห์การเปลี่ยนแปลงของราคาหลักทรัพย์จนทำให้ได้รับการยอมรับว่าเป็นระบบที่มีความสามารถในการเรียนรู้ได้ทั้งการวิเคราะห์โดยใช้ปัจจัยพื้นฐานและการวิเคราะห์โดยใช้ปัจจัยทางด้านเทคนิค ตลอดจนระบบสารสนเทศในปัจจุบันมีประสิทธิภาพสูง ทำให้สามารถเลือกใช้อัลกอริทึมต่างๆของโครงข่ายประสาทเทียม เพื่อนำมาประกอบกันเป็นระบบสนับสนุนการตัดสินใจที่มีความถูกต้อง แม่นยำ และน่าเชื่อถือมากยิ่งขึ้น

ดังนั้นโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้ จะพัฒนาระบบสนับสนุนการตัดสินใจเพื่อการลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยโดยสร้างแบบจำลองพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้เทคโนโลยีดาต้าไมนิ่งจากอัลกอริทึมแบล็คพร็อบพาเกชั่น สำหรับนำไปพยากรณ์ราคาหลักทรัพย์ในอนาคต โดยนำข้อมูลที่มีความครอบคลุมถึงปัจจัยพื้นฐานและปัจจัยทางด้านเทคนิคโดยมีลักษณะเป็นข้อมูลอนุกรมเวลา และมีความสัมพันธ์กันในหลายๆมิติ เช่น ราคาปิดของหลักทรัพย์ อัตราแลกเปลี่ยนเงินตราต่างประเทศ อัตราดอกเบี้ยเงินฝากภายในประเทศ ดัชนีราคาหลักทรัพย์ทั้งในประเทศและต่างประเทศ ตลอดจนราคาน้ำมันภายในประเทศ เพื่อให้ได้ระบบที่ดีที่สุด สำหรับนักลงทุนนำไปเป็นเครื่องมือที่ใช้สนับสนุนการตัดสินใจในการลงทุนในตลาดหลักทรัพย์แห่งประเทศไทยได้ด้วยความมั่นใจ และมีประสิทธิภาพสูงสุดต่อไป

1.1 วัตถุประสงค์ของโครงงาน

1.2.1 เพื่อศึกษาทฤษฎีและหลักการของเทคโนโลยีดาต้าไมนิ่ง เพื่อนำมาประยุกต์ใช้ในงานด้านการพยากรณ์ข้อมูล

1.2.2 เพื่อศึกษาทฤษฎีและหลักการของอัลกอริทึมโครงข่ายประสาทเทียม และอัลกอริทึมแบล็คพร็อบพาเกชั่น ในการพยากรณ์ข้อมูล

1.2.3 เพื่อสร้างแบบจำลองพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้เทคโนโลยีดาต้าไมนิ่งจากอัลกอริทึมแบล็คพร็อบพาเกชั่น

1.2.4 เพื่อพิสูจน์ เปรียบเทียบ และนำเสนอแบบจำลองพยากรณ์ฯ กับข้อมูลที่เกิดขึ้นจริง

1.2.5 เพื่อใช้เป็นพื้นฐานสำหรับการนำเทคโนโลยีดาต้าไมนิ่งไปประยุกต์ใช้ในด้านอื่นๆได้อย่างมีประสิทธิภาพต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 ขอบเขตของโครงงาน

1.3.1 เป็นการพัฒนาระบบตามหลักการโดยใช้เทคโนโลยีดาต้าไมนิ่งจากอัลกอริทึมแบล็คพร็อบพาเกชั่น ในการสร้างแบบจำลองพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย

1.3.2 ข้อมูลราคาปิดของหลักทรัพย์ และข้อมูลปัจจัยความเสี่ยงต่างๆที่มีผลต่อราคาปิดของหลักทรัพย์ ตลอดจนมีผลต่อสภาวะเศรษฐกิจ ซึ่งข้อมูลที่นำมาใช้ทดลองเริ่มตั้งแต่วันที่ 4 ม.ค. 2531 สิ้นสุดวันที่ 31 พ.ค. 2547

1.3 แนวทางการดำเนินงาน

1.4.1 ศึกษาทฤษฎีและหลักการของเทคโนโลยีดาต้าไมนิ่ง ที่ใช้ในงานด้านการพยากรณ์ข้อมูล

1.4.2 ศึกษาทฤษฎีและหลักการของอัลกอริทึมแบล็คพร็อบพาเกชั่น ในการพยากรณ์ข้อมูล เพื่อใช้สร้างแบบจำลองพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย

1.4.3 . รวบรวมข้อมูลราคาปิดหลักทรัพย์ และข้อมูลปัจจัยความเสี่ยงต่างๆที่มีผลต่อราคาปิดของหลักทรัพย์ เพื่อใช้ในการสร้างแบบจำลองพยากรณ์ฯโดยใช้เทคโนโลยีดาต้าไมนิ่ง

1.4.4 ดำเนินการพัฒนาระบบงาน โดยแบ่งส่วนการทำงานเป็น 4 ขั้นตอน ดังนี้

1.4.4.1 การนำข้อมูลเข้าสู่ระบบ และการประมวลผลเบื้องต้น

1.4.4.2 การสร้างแบบจำลองพยากรณ์ฯโดยใช้เทคโนโลยีดาต้าไมนิ่ง

1.4.4.3 ทดสอบและเปรียบเทียบ แบบจำลองพยากรณ์ฯโดยใช้เทคโนโลยีดาต้าไมนิ่งที่สร้างจากอัลกอริทึมแบล็คพร็อบพาเกชั่น กับข้อมูลที่เกิดขึ้นจริง

1.4.4.4 สรุปและนำเสนอผลลัพธ์ที่ได้จากการทดสอบและเปรียบเทียบ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 . สามารถเรียนรู้ และเข้าใจทฤษฎี หลักการ และกระบวนการทำงานของเทคโนโลยีดาต้าไมนิ่ง อัลกอริทึมของโครงข่ายประสาทเทียม และอัลกอริทึมแบล็คพร็อบพาเกชั่น สำหรับการประยุกต์ใช้งานด้านการพยากรณ์ข้อมูล

1.5.2 สามารถสร้างเครื่องมือหรือระบบเพื่อสนับสนุนการตัดสินใจลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทย

1.5.3 สามารถนำทฤษฎีและหลักการของเทคโนโลยีดาต้าไมนิ่งไปสร้างระบบ พัฒนาระบบ เพื่อสนับสนุนการตัดสินใจ และนำไปประยุกต์ใช้กับงานด้านอื่นๆได้ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีเบื้องต้นของดาต้าไมนิ่ง

การดำเนินธุรกิจในโลกที่เปลี่ยนแปลงอยู่ตลอดเวลานั้น นอกจากจะต้องมีภาคการผลิตที่มีประสิทธิผลสูง ต้นทุนต่ำ และภาคการบริการที่มีประสิทธิภาพนั้น ไม่เพียงพอต่อการแข่งขันในยุคการค้าเสรี เพราะปัจจุบันนี้เป็นยุคของข้อมูลข่าวสาร องค์กรที่สามารถเข้าถึงข้อมูลข่าวสารได้มากกว่า รวดเร็วกว่า และสามารถนำมาใช้ประกอบการตัดสินใจได้ในเวลาที่เหมาะสมนั้น จะทำให้สามารถแข่งขันได้อย่างมีประสิทธิภาพ

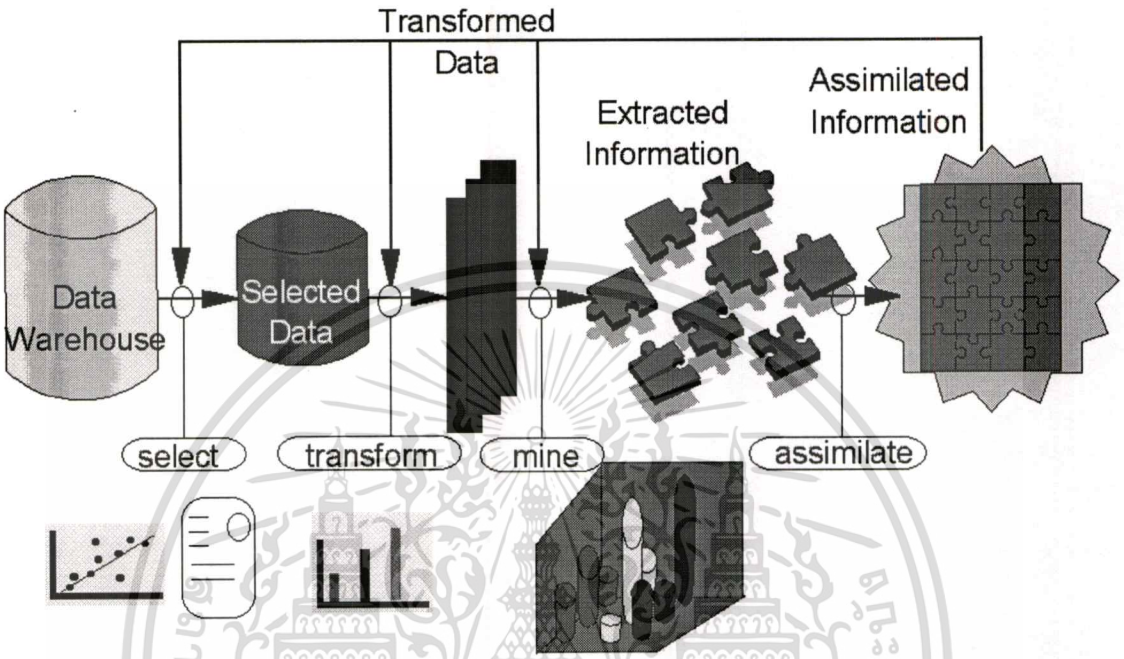
ปัจจุบันเทคโนโลยีสารสนเทศได้รับการพัฒนาอย่างรวดเร็ว และมีประสิทธิภาพสูงมาก ทำให้องค์กรสามารถจัดเก็บข้อมูลจากการดำเนินธุรกิจให้อยู่ในรูปอิเลคทรอนิกส์ได้ง่ายดาย และสามารถจัดเก็บข้อมูลได้อย่างมากมาย มหาศาล อีกทั้งความก้าวหน้าของระบบเครือข่ายอินเทอร์เน็ตทำให้องค์กรสามารถถึงเข้าข้อมูลจากโลกภายนอกได้มากยิ่งขึ้นแต่อย่างไรก็ตามในระบบฐานข้อมูลที่มีขนาดใหญ่ขึ้น มีจำนวนข้อมูลมากขึ้นการที่จะทำการวิเคราะห์หาข้อมูลสารสนเทศจากฐานข้อมูลดังกล่าวด้วยมือเป็นเรื่องที่ทำได้ยากมาก ดังนั้นจึงมีการพัฒนาเทคโนโลยีดาต้าไมนิ่งขึ้น เพื่อใช้เป็นเครื่องมือสำหรับสร้างข้อมูลสารสนเทศ จากฐานข้อมูลดังกล่าว เพื่อให้องค์กรสามารถนำผลลัพธ์ที่ได้ไปใช้ประกอบการตัดสินใจเพื่อแข่งขันในโลกธุรกิจได้อย่างมีประสิทธิภาพต่อไป

2.1 หลักการทั่วไปของดาต้าไมนิ่ง

ดาต้าไมนิ่ง เป็นกระบวนการที่สำคัญกระบวนการหนึ่งในการค้นหาลักษณะแฝงของข้อมูลที่อยู่ในกลุ่มข้อมูลจำนวนมาก (Knowledge Discovery in Database : KDD) ซึ่งลักษณะที่น่าสนใจของข้อมูลเหล่านี้ ได้แก่ รูปแบบ ความสัมพันธ์ การเปลี่ยนแปลงของข้อมูลโครงสร้างที่เด่นชัด หรือลักษณะที่ผิดปกติของข้อมูลจากข้อมูลจำนวนมาก ที่เก็บอยู่ในฐานข้อมูลหรือแหล่งที่เก็บข้อมูลอื่นๆ เป็นต้น ซึ่งวิธีการที่นำมาใช้ในการทำไมนิ่งข้อมูลเหล่านี้ก็มีวัตถุประสงค์แตกต่างกันไปขึ้นอยู่กับผลลัพธ์ของกระบวนการโดยรวมที่ต้องการ ดังนั้นจึงมีการนำเสนอวิธีการหลายอย่างสำหรับแต่ละงานที่มีเป้าหมายที่แตกต่างกันไปเพื่อให้ได้ผลลัพธ์ที่เหมาะสมตามที่ต้องการ และเนื่องจากความแพร่หลายของการจัดเก็บข้อมูลในลักษณะที่อยู่ในรูปแบบอิเล็กทรอนิกส์ และความต้องการเปลี่ยนแปลงข้อมูลเหล่านี้ให้เป็นข้อมูลสารสนเทศ ที่มีประโยชน์เพื่อนำไปประยุกต์ใช้ในงานด้านต่างๆ เช่น การจัดกลุ่มข้อมูลลูกค้า การค้นหารูปแบบความสัมพันธ์ในตัวสินค้า เพื่อช่วยสนับสนุน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

องค์กรในการตัดสินใจทำการตลาด ดังนั้นจึงทำให้การทำดาต้าไมนิ่งได้รับความสนใจมากยิ่งขึ้น กระบวนการทำดาต้าไมนิ่งประกอบด้วยขั้นตอนต่างๆตามรูปที่ 2.1 ดังนี้



รูปที่ 2.1 แสดงขั้นตอนต่างๆ ของกระบวนการทำดาต้าไมนิ่ง

ขั้นตอนการทำดาต้าไมนิ่ง ประกอบไปด้วย

2.1.1 การกำหนดวัตถุประสงค์ทางธุรกิจ

2.1.2 การคัดเลือกข้อมูล

2.1.3 การกรองข้อมูล

2.1.4 การแปลงรูปแบบข้อมูล

2.1.5 การทำไมนิ่งข้อมูล

2.1.6 การวิเคราะห์และประเมินผลลัพธ์

2.1.1 การกำหนดวัตถุประสงค์ทางธุรกิจ

ถือเป็นหัวใจสำคัญยิ่ง เนื่องจากเป็นสิ่งที่กำหนดทิศทางในการทำดาต้าไมนิ่ง การกำหนดวัตถุประสงค์นั้นจะต้องเข้าใจถึงปัญหาและความต้องการของธุรกิจได้อย่างชัดเจน และถูกต้อง เพื่อนำไปสู่ความสำเร็จในการแก้ปัญหาได้อย่างมีประสิทธิภาพสูงสุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 การคัดเลือกข้อมูล

เป็นการคัดเลือกเฉพาะข้อมูลที่ต้องการตามวัตถุประสงค์ที่ตั้งไว้ การคัดเลือกข้อมูลจะต้องเข้าถึงแหล่งกำเนิดข้อมูล เข้าใจความหมายของข้อมูล ประเภทข้อมูล ความเป็นไปได้ของข้อมูล รวมถึงความทันสมัยของข้อมูล เนื่องจากการเปลี่ยนแปลงของเวลาและสถานการณ์ภายนอกที่เกิดขึ้นตลอดเวลา จึงมีผลทำให้ข้อมูลลดประสิทธิภาพลง นอกจากนี้จะต้องพิจารณาปัจจัยเพิ่มเติม ดังนี้

- ระดับข้อมูล จะขึ้นอยู่กับวัตถุประสงค์ว่าต้องการข้อมูลที่สรุป หรือข้อมูลที่อยู่ในระดับรายการทั่วไป
- ลักษณะการจัดเก็บข้อมูลอยู่ในลักษณะใด เช่น ECDIC Code, ASCII Code, Floating Point, Packed Decimal etc. ซึ่งจะขึ้นอยู่กับภาษาคอมพิวเตอร์และระบบปฏิบัติการที่ใช้
- ความแตกต่างกันของข้อมูลจากแหล่งข้อมูลที่แตกต่างกัน รวมถึงรูปแบบ ความหมายและลักษณะการจัดเก็บข้อมูลด้วย
- ความหมายของข้อมูล ประเภทข้อความ ซึ่งค่าของข้อมูลที่จัดเก็บอาจจะแตกต่างกัน เนื่องจากความผิดพลาดสามารถเกิดขึ้นได้ในระหว่างการบันทึกข้อมูล

2.1.3 การกรองข้อมูล

เป็นกระบวนการที่ทำให้เกิดความมั่นใจในคุณภาพของข้อมูลที่จะนำมาใช้วิเคราะห์ ซึ่งโดยทั่วไปในขั้นตอนนี้จะคำนึงถึงข้อมูลใน 2 ลักษณะ ดังนี้

- **ข้อมูลรบกวน (Noisy Data)** ซึ่งหมายถึงข้อมูลที่มีค่าผิดพลาดไปจากค่าปกติ หรือค่าที่เป็นไปได้
- **ข้อมูลผิดพลาด (Missing Value)** ซึ่งหมายถึงข้อมูลที่ขาดหายไป โดยจะต้องพิจารณาว่าควรจะตัดข้อมูลเหล่านั้นทิ้งไป หรือควรหาวิธีที่เหมาะสมในการคาดคะเนค่าสำหรับข้อมูลในตำแหน่งนั้นๆ เพื่อนำมาใช้วิเคราะห์ร่วมกับข้อมูลอื่นๆ ได้ต่อไป

2.1.4 การแปลงรูปแบบข้อมูล

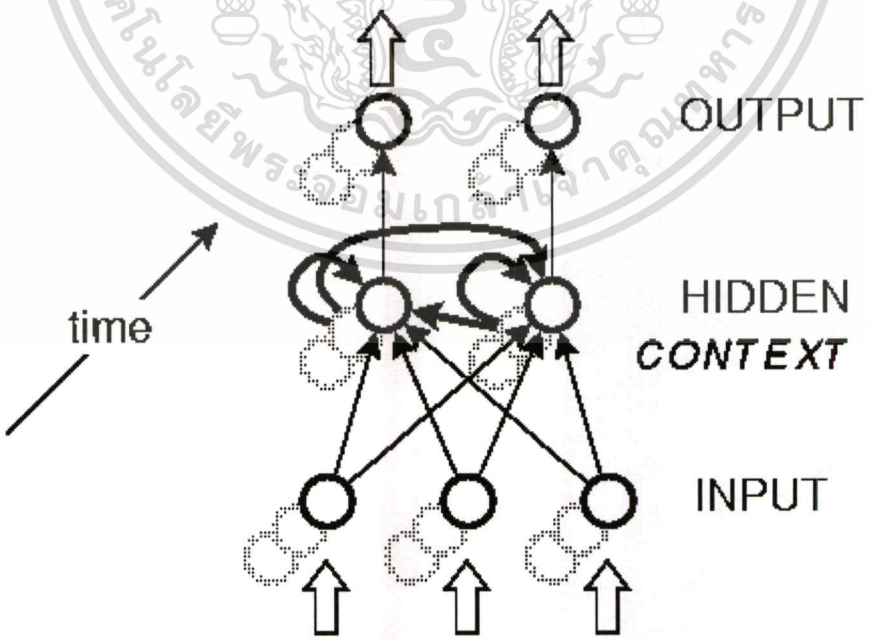
เป็นการแปลงข้อมูลที่เลือกมาให้อยู่ในรูปแบบที่สามารถนำไปใช้วิเคราะห์ตามอัลกอริทึมของแบบจำลองที่ใช้ในการทำดาต้าไมนิ่งได้อย่างมีประสิทธิภาพต่อไป เช่น โครงข่ายแพร่กระจายย้อนกลับ ใช้ได้กับข้อมูลที่เป็นตัวเลขโดยมีค่าในระหว่าง -1 ถึง 1 เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.5 การทำไมนิ่งข้อมูล

จากลักษณะงานหลายประเภทในการทำดาต้าไมนิ่งนั้นทำให้มีการพัฒนาเทคนิค หรือวิธีการที่เหมาะสมเพื่อใช้ในงานแต่ละประเภท ซึ่งวิธีการที่ต่างกันก็จะมีความเหมาะสมกับวัตถุประสงค์ของผู้ใช้งานในด้านต่างๆกันด้วย โดยแต่ละวิธีก็มีข้อดีและข้อด้อยแตกต่างกันไป นอกจากนี้ในงานแต่ละประเภทนั้นไม่ได้ระบุว่าต้องใช้เทคนิคใดเทคนิคหนึ่งโดยเฉพาะเท่านั้น แต่ขึ้นอยู่กับลักษณะของปัญหาที่นำมาวิเคราะห์ ซึ่งในบางครั้งการนำเทคนิคหลายๆอย่างมาใช้เพื่อเปรียบเทียบความเหมาะสมกับแต่ละงาน หรือการนำหลายๆวิธีมาใช้ร่วมกันอาจจะมีส่วนช่วยในการค้นหาความหมาย หรือความสัมพันธ์จากข้อมูลออกมาได้อย่างมีประสิทธิภาพมากกว่าการใช้เพียงวิธีการใดวิธีการหนึ่ง ซึ่งตัวอย่างของเทคนิคที่สามารถนำมาใช้ในการทำดาต้าไมนิ่ง ได้แก่

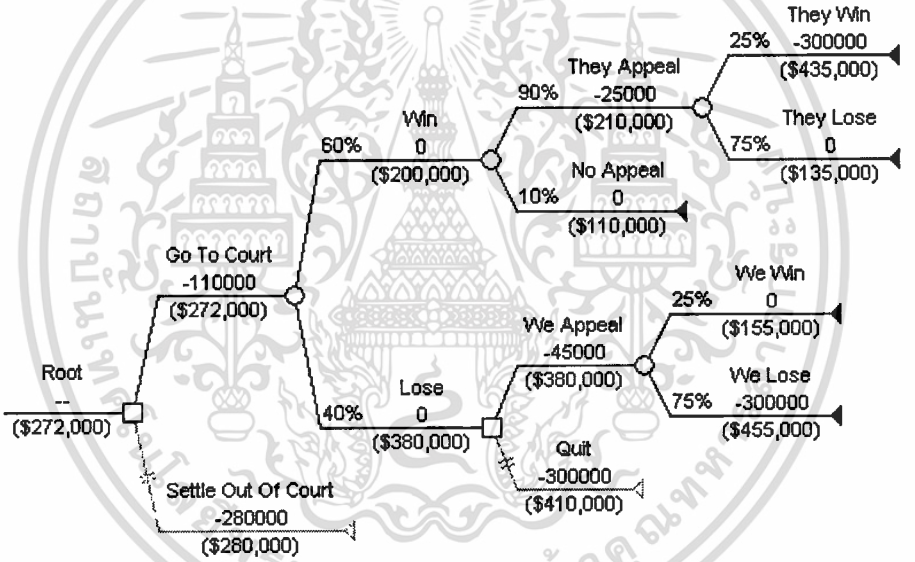
- **โครงข่ายประสาทเทียม (Neural Networks)** เป็นวิธีการที่ใช้ในการระบุรูปแบบของข้อมูล หรือ จัดกลุ่มข้อมูล โดยในแบบจำลองจะประกอบไปด้วยโหนดที่มีการเชื่อมต่อกันเป็นโครงข่าย ทำงานในลักษณะคล้ายกับระบบประสาทของมนุษย์ ดังแสดงตัวอย่างในรูปที่ 2.2 โดยที่แต่ละโหนดมีการตอบสนองต่อตัวแปร หรือข้อมูลที่เข้ามากระตุ้นในระบบ โดยในช่วงเวลาใดๆอาจมีโหนดที่ได้รับการกระตุ้นให้มีการทำงานได้พร้อมๆกันหลายๆโหนด ซึ่งลักษณะของการทำงานที่ตอบสนองต่อข้อมูลที่เข้าสู่ระบบในแต่ละครั้งนั้น จะเสมือนเป็นตัวแทนของข้อมูลที่สร้างให้ระบบมีการเรียนรู้ลักษณะดังกล่าวของแต่ละข้อมูล ต่อจากนั้นจึงทำการจัดกลุ่มข้อมูลตามลักษณะบางอย่างที่คล้ายคลึงกัน



รูปที่ 2.2 แสดงตัวอย่างของโครงข่ายประสาทเทียม (Neural Network)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **โครงสร้างการตัดสินใจแบบต้นไม้ (Decision Tree)** เป็นการวิเคราะห์เพื่อค้นหากฎเกณฑ์ หรือความสัมพันธ์ในลักษณะที่แบ่งย่อยลงไปตามโครงสร้างแบบต้นไม้ โดยที่แต่ละโหนดที่ไม่ใช่โหนดสุดท้ายของโครงสร้างจะแทนจุดที่มีการตรวจสอบเงื่อนไข หรือเป็นการตัดสินใจเกี่ยวกับลักษณะของข้อมูลที่นำมาพิจารณาและจากผลลัพธ์ที่ได้ในแต่ละครั้งจะมีการเลือกเส้นทางที่แน่นอน โดยในการจัดแบ่งประเภทของข้อมูลนั้นจะเริ่มต้นพิจารณาจากโหนดที่เป็นโหนดแรก (Root Node) และอ้างอิงลงมาตามโครงสร้างต้นไม้เรื่อยๆไป จนกระทั่งถึงโหนดสุดท้ายของโครงสร้าง (Leaf Node) ซึ่งแสดงว่ากระบวนการตัดสินใจ หรือการวิเคราะห์ข้อมูลนั้นๆ เสร็จสิ้นลงแล้ว และสามารถบอกลักษณะหรือประเภทของข้อมูลนั้นได้ตามโหนดที่ข้อมูลผ่านการตรวจสอบ ดังแสดงตัวอย่างโครงสร้างในรูปที่ 2.3



รูปที่ 2.3 แสดงตัวอย่างของโครงสร้างการตัดสินใจแบบต้นไม้ (Decision Tree)

- **การหาความสัมพันธ์ของข้อมูล (Association Rule)** เป็นวิธีการเพื่อหาข้อมูลที่มีความสัมพันธ์กันหรือข้อมูลที่มีลักษณะการเกิดขึ้นร่วมกันภายในกลุ่มข้อมูลที่นำมาพิจารณา แต่จากผลที่ได้นั้นไม่ได้หมายความว่าความสัมพันธ์ของข้อมูลที่พบนั้น จะเป็นสิ่งที่มีความสำคัญเสมอไป ดังนั้นจึงต้องนำผลที่ได้ไปพิจารณาโดยใช้วิธีที่เหมาะสมต่อไป

2.1.6 การวิเคราะห์และประเมินผลลัพธ์

ถือเป็นขั้นตอนที่มีความสำคัญ โดยจะต้องทำการแปลความหมายและประเมินผลลัพธ์ที่ได้จากการทำดาต้าไมนิ่ง ซึ่งต้องอาศัยทักษะในการวิเคราะห์ข้อมูลในทางเทคนิค และการวิเคราะห์ทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ด้านธุรกิจประกอบกัน รวมไปถึงการนำเสนอผลการวิเคราะห์โดยใช้เครื่องมือช่วยนำเสนอต่างๆ (Graphical Visualization) นอกจากนี้ต้องพิจารณาถึงความสมบูรณ์ ความน่าสนใจ และต้องนำไปใช้เป็นประโยชน์ตอบสนองวัตถุประสงค์ที่ตั้งไว้ ซึ่งหากไม่ได้ผลตามนั้น ก็จะต้องวนกลับไปทำดาต้าไมนิ่งและพิจารณาขั้นตอนต่างๆใหม่เพื่อปรับปรุงแก้ไขให้สามารถใช้งานได้ดีขึ้น

2.2 ประเภทของการใช้งานดาต้าไมนิ่ง

ดาต้าไมนิ่งเป็นวิธีการที่ถูกนำมาใช้อย่างกว้างขวางในการวิเคราะห์หาความหมาย หรือลักษณะเด่นที่แฝงอยู่ในข้อมูลจำนวนมากๆ โดยการเลือกเทคนิค หรือแบบจำลองต่างๆ เพื่อนำมาประยุกต์ใช้ก็ขึ้นอยู่กับประเภทของงานและวัตถุประสงค์ของผู้ใช้เป็นหลัก ซึ่งลักษณะงานที่มีการนำดาต้าไมนิ่งมาใช้สามารถจำแนกได้ดังตารางที่ 2.1 ดังนี้

ตารางที่ 2.1 แสดงประเภทของงานที่มีการนำเทคนิคต่างๆของดาต้าไมนิ่งมาใช้งาน

	Market Management	Risk Management	Fraud Management
Applications	✓ Traget Marketing ✓ Market Basket Analysis ✓ Cross Selling ✓ Market Segmentation ✓ Customer Relationship Manangement	✓ Forecasting ✓ Customer Retention ✓ Quality Control ✓ Competitive Analysis	✓ Fraud Detection

	Predictive Model	Database Segmentation	Link Analysis	Deviation Detection
Operations				
Techniques	✓ Classification ✓ Value Prediction	✓ Demographic Clustering ✓ Neural Clustering	✓ Association Discovery ✓ Sequential Pattern Discovery ✓ SimilarTime Sequence Discovery	✓ Visualization ✓ Statistics

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยทั่วไปประเภทของงานตามลักษณะของแบบจำลองที่ใช้ในการทำดาต้าไมนิ่งนั้น สามารถแบ่งกลุ่มได้เป็น 2 ประเภท ดังนี้

2.2.1 งานด้านการพยากรณ์ (Predictive Data Mining)

เป็นการคาดคะเนลักษณะ หรือประมาณค่าที่ชัดเจนของข้อมูลที่จะเกิดขึ้น โดยใช้พื้นฐานจากข้อมูลที่ผ่านมาในอดีต สามารถแบ่งลักษณะงาน ได้ดังต่อไปนี้

- **การจัดหมวดหมู่ (Classification)** เป็นการจัดกลุ่มให้กับแต่ละข้อมูลในฐานข้อมูล โดยมีการระบุค่าหรือลักษณะที่เป็นไปได้ของข้อมูลภายในแต่ละกลุ่ม เช่น การจัดกลุ่มของลูกค้าที่ใช้บริการโทรศัพท์เคลื่อนที่เพื่อระบุรูปแบบสำหรับออกแผนการตลาดสำหรับลูกค้ารายใหม่ เป็นต้น
- **การพยากรณ์ (Prediction)** เป็นการพยากรณ์ค่าที่เป็นไปได้ หรือการกระจายค่าที่เป็นไปได้ของตัวแปรใดๆในกลุ่มข้อมูล

2.2.2 งานด้านการแบ่งกลุ่ม (Descriptive Data Mining)

เป็นการหาแบบจำลองเพื่ออธิบายลักษณะบางอย่างของข้อมูลที่มีอยู่ ซึ่งโดยส่วนมากจะเป็นลักษณะการแบ่งกลุ่มให้กับข้อมูล สามารถแบ่งลักษณะงานได้ ดังต่อไปนี้

- **การหาความสัมพันธ์ (Link Analysis)** เป็นการหาความสัมพันธ์ของข้อมูลภายในกลุ่มข้อมูล เพื่อใช้ลักษณะของข้อมูลหนึ่งๆ ในการบอกลักษณะที่จะเกิดขึ้นกับข้อมูลอีกตัวหนึ่ง ซึ่งอาจจะเป็นการหาความสัมพันธ์ของข้อมูลในกลุ่มเดียวกัน เช่น การระบุว่าในกลุ่มของลูกค้าที่ซื้อสบู่นั้นจะมีลูกค้าที่ 60% ที่ซื้อยาสระผมด้วย หรือ อาจจะเป็นการหาความสัมพันธ์ของตัวแปรระหว่างกลุ่มข้อมูลก็ได้ เช่น ในทุกๆครั้งที่ดัชนีของตลาดหุ้นหนึ่งลดลง 8% ดัชนีของอีกตลาดหุ้นหนึ่งจะเพิ่มขึ้น 10% ภายในช่วง 3 เดือนหลังจากนั้น ซึ่งลักษณะของการหาความสัมพันธ์นั้นอาจแบ่งได้เป็น 3 กลุ่ม คือ การหาความสัมพันธ์ระหว่างข้อมูล (Association Discovery) การหาความสัมพันธ์ในลักษณะที่เป็นลำดับของข้อมูล (Sequential Pattern Discovery) และการหาความสัมพันธ์ของข้อมูลกับช่วงเวลาใดๆ (Similar Time Sequence Discovery)
- **การแบ่งกลุ่ม (Clustering)** เป็นการแบ่งกลุ่มของข้อมูลโดยที่ในแต่ละกลุ่มนั้นประกอบด้วยข้อมูลที่มีลักษณะคล้ายคลึงกัน โดยในการแบ่งกลุ่มของข้อมูลที่ดีนั้น จะทำให้ได้กลุ่มที่มีคุณภาพ นั่นแสดงว่ามีความแตกต่างกันระหว่างกลุ่มมาก (High Heterogeneity) และมีความแตกต่างกันภายในกลุ่มน้อย (High Homogeneity) เช่น การแบ่งกลุ่มของลูกค้าโดยใช้พฤติกรรมการซื้อเป็นเกณฑ์กำหนด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **การหาค่าเบี่ยงเบนไปจากค่ามาตรฐาน (Deviation Detection)** เป็นเทคนิคที่ใช้ในการแสดงลักษณะของข้อมูลที่ผิดปกติ หรือผิดไปจากที่คาดหวังไว้ โดยมีการแสดงผลอยู่ในลักษณะที่สามารถทำความเข้าใจและแปลความหมายได้ง่าย เช่น รูปแบบกราฟ เป็นต้น

2.3 การประยุกต์ใช้ดาต้าไมนิ่ง กับการพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์

ในโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้ จะพัฒนาระบบสนับสนุนการตัดสินใจเพื่อการลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยโดยสร้างแบบจำลองพยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้เทคโนโลยีดาต้าไมนิ่งจากอัลกอริทึมแบล็คพร็อบพาเกชั่น สำหรับนำไปพยากรณ์ราคาหลักทรัพย์ในอนาคต โดยนำข้อมูลที่มีความครอบคลุมถึงปัจจัยพื้นฐานและปัจจัยทางด้านเทคนิคโดยมีลักษณะเป็นข้อมูลอนุกรมเวลา และมีความสัมพันธ์กันในหลายๆมิติ เช่น ราคาปิดของหลักทรัพย์ อัตราแลกเปลี่ยนเงินตราต่างประเทศ อัตราดอกเบี้ยเงินฝากภายในประเทศ ดัชนีราคาหลักทรัพย์ทั้งในประเทศและต่างประเทศ ตลอดจนราคาน้ำมันภายในประเทศ มาใช้เป็นอินพุตร่วมกัน และหลังจากที่โครงข่ายประสาทเทียมได้เรียนรู้แล้ว ก็จะสามารถนำแบบจำลองพยากรณ์ฯ มาใช้ในการพยากรณ์ค่าของราคาหลักทรัพย์ที่เกิดขึ้นในช่วงเวลาต่อๆไปได้ ซึ่งทฤษฎีและหลักการของโครงข่ายประสาทเทียมนั้นจะได้อธิบายไว้ในบทที่ 3 ต่อไป

2.3.1 การวิเคราะห์ข้อมูลในอนุกรมเวลา (Time Series Analysis)

ลักษณะของข้อมูลที่แสดงความสัมพันธ์กับเวลานั้น โดยทั่วไปจะอยู่ในรูปของ t(1), t(2), t(3), ...t(i), ...t(n) โดยที่ n เป็นค่าล่าสุดของข้อมูลในช่วงเวลาใดๆ ซึ่งจากข้อมูลที่รวบรวมได้ในช่วงเวลาที่ผ่านมา การวิเคราะห์ข้อมูลในลักษณะชุดข้อมูลตามช่วงเวลาดังกล่าวนี้ มุ่งเน้นในการพยายามหาแบบจำลองใดๆ เพื่อใช้พยากรณ์ค่าของข้อมูลในลำดับที่ t(n+1) จากค่าของข้อมูลในลำดับก่อนหน้า

ในการวิเคราะห์ชุดข้อมูลที่มีความสัมพันธ์กันในอนุกรมเวลานั้น เป็นการสร้างแบบจำลองอีกแบบหนึ่งที่มีการนำไปประยุกต์ใช้ในเชิงธุรกิจอย่างกว้างขวาง เนื่องจากมีการพิจารณาว่าค่าของข้อมูลหรือลักษณะต่างๆ ที่เกิดขึ้นกับข้อมูลนั้นมีรูปแบบเฉพาะที่มีความสัมพันธ์กันในเชิงเวลา ดังนั้นจึงมีการนำข้อมูลจากช่วงเวลาที่ผ่านมาก่อนหน้า มาทำการวิเคราะห์เพื่อให้สามารถหาวิธีพยากรณ์ค่าข้อมูลที่จะเกิดขึ้นในอนาคตได้ ซึ่งจากลักษณะของข้อมูลราคาหลักทรัพย์โดยทั่วไปนั้นก็จะมีการรวบรวมไว้ในลักษณะของชุดข้อมูลต่างๆ เทียบกับช่วงเวลาเช่นเดียวกัน และเมื่อพิจารณาลักษณะของการวิเคราะห์ข้อมูลแบบที่สัมพันธ์กับเวลานี้แล้ว สามารถแบ่งได้ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **การวิเคราะห์โดยพิจารณาตัวแปรเดียว (Univariate Time – Series Analysis)** เป็นการนำข้อมูลของตัวแปรที่สนใจเพียงตัวแปรเดียวมาใช้ในการวิเคราะห์ เช่น การพยากรณ์ค่าราคาหลักทรัพย์ในอนาคต โดยใช้ข้อมูลในอดีตของราคาหลักทรัพย์ที่ต้องการพยากรณ์มาพิจารณาเพียงตัวแปรเดียวเท่านั้น
- **การวิเคราะห์โดยพิจารณาจากหลายตัวแปร (Multivariate Time – Series Analysis)** เป็นการพัฒนามาจากการพยากรณ์โดยใช้ตัวแปรเดียว โดยแทนที่จะพิจารณาเพียงตัวแปรเดียว ก็จะมีการนำตัวแปรอื่นๆ ณ ช่วงเวลาเดียวกันมาพิจารณาร่วมด้วย เช่น ในการพยากรณ์ราคาหลักทรัพย์อาจจะมีการนำค่าอัตราแลกเปลี่ยนของเงินสกุลต่างๆ ราคาน้ำมัน ดัชนีราคาหลักทรัพย์ในแต่ละประเทศ มาพิจารณาร่วมกับราคาหลักทรัพย์ในแต่ละตัวด้วย

สำหรับในการพัฒนาโครงการนี้จะสร้างระบบเพื่อการพยากรณ์ราคาหลักทรัพย์ตามลักษณะที่ 2 คือเป็นการพยากรณ์ค่าของราคาหลักทรัพย์ในช่วงเวลาถัดไปของอนุกรมเวลา และใช้การวิเคราะห์โดยพิจารณาจากหลายตัวแปร ดังต่อไปนี้

- ราคาหลักทรัพย์ของบริษัทที่สนใจ
- ดัชนีราคาหลักทรัพย์ในประเทศ
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศไทย
  - ดัชนีตลาดหลักทรัพย์ 50
  - ดัชนีตลาดหลักทรัพย์ใหม่
- ดัชนีราคาหลักทรัพย์ต่างประเทศ
  - ดัชนีตลาดหลักทรัพย์แนสเด็ค แห่งสหรัฐอเมริกา
  - ดัชนีตลาดหลักทรัพย์ดาวน์โจนส์ แห่งสหรัฐอเมริกา
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศอังกฤษ
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศเกาหลี
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศสิงคโปร์
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศไต้หวัน
  - ดัชนีตลาดหลักทรัพย์นิเคอิ แห่งประเทศญี่ปุ่น
  - ดัชนีตลาดหลักทรัพย์แห่งประเทศฟิลิปินส์
  - ดัชนีตลาดหลักทรัพย์ฮั่งเส็ง แห่งประเทศฮ่องกง
  - ดัชนีตลาดหลักทรัพย์จากาต้าแห่งประเทศอินโดนีเซีย
  - ดัชนีตลาดหลักทรัพย์กัวลาลัวเปอร์ แห่งประเทศมาเลเซีย
- มูลค่าการซื้อขายในตลาดหลักทรัพย์แห่งประเทศไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- มูลค่าการซื้อขายของนักลงทุนรายย่อย
- มูลค่าการซื้อขายของนักลงทุนสถาบัน
- มูลค่าการซื้อขายของนักลงต่างประเทศ

- อัตราแลกเปลี่ยนเงินตราต่างประเทศไทย-สหรัฐอเมริกา
- อัตราดอกเบี้ยเงินฝากธนาคารแห่งประเทศไทย
- ราคาน้ำมันสำเร็จรูปในประเทศไทย
  - ราคาน้ำมันดีเซลสำเร็จรูป ในประเทศไทย
  - ราคาน้ำมันเบนซิน 91สำเร็จรูป ในประเทศไทย
  - ราคาน้ำมันเบนซิน 95สำเร็จรูป ในประเทศไทย

โดยใช้ข้อมูลในช่วงที่ผ่านมาระยะหนึ่งเป็นเกณฑ์ในการสร้างแบบจำลองพยากรณ์ให้กับระบบ ซึ่งกระบวนการในการสร้างแบบจำลองพยากรณ์จะได้อธิบายในบทที่ 4 ต่อไป

2.3.2 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing)

ในการประมวลผลข้อมูลเบื้องต้นก่อนการสร้างแบบจำลองพยากรณ์ให้เหมาะสมกับระบบพยากรณ์ราคาหลักทรัพย์นั้น มีการพิจารณาใน 2 ลักษณะ คือ การพิจารณาข้อมูลส่วนที่ขาดหายไปในอนุกรมเวลา (Data Cleaning) และการแปลงค่าข้อมูลให้เหมาะสมกับรูปแบบอัลกอริทึมที่ใช้ในการสร้างแบบจำลองพยากรณ์ (Data Transformation) ซึ่งมีวิธีการหลักที่นำมาใช้ ดังต่อไปนี้

2.3.2.1 การแทนค่าข้อมูล (Data Cleaning)

สำหรับเทคนิคที่นิยมนำมาใช้ในการหาข้อมูลเพื่อแทนค่าข้อมูลที่ขาดหายไป ได้แก่

- **การหาค่าเฉลี่ยเคลื่อนที่ (Moving Average)** ซึ่งเป็นการคำนวณหาค่าเฉลี่ยของข้อมูลจากช่วงเวลาหนึ่งก่อนหน้าแล้วจึงนำมาใช้เป็นอินพุตแทนข้อมูลส่วนที่หายไปให้กับระบบแทน เพื่อช่วยเพิ่มประสิทธิภาพให้กับการสร้างแบบจำลองในการพยากรณ์ข้อมูล โดยส่วนใหญ่จะใช้ช่วงในการหาค่าเฉลี่ยที่ประมาณ 2 – 5 ช่วงข้อมูล โดยมีสูตรที่ใช้ในการคำนวณ ดังต่อไปนี้

$$MA(i,m) = \frac{1}{m}\sum_{j=1-m+1}^{i} t(j) \qquad \text{.....2.1}$$

- **การใช้สมการถดถอยแบบเชิงเส้น (Linear Regression Model)** เป็นการหาสมการเส้นตรงที่เหมาะสมให้กับชุดของข้อมูล โดยที่สามารถนำสมการที่ได้มาใช้คำนวณค่าของตัวแปรหนึ่ง (Y) ได้จากค่าตัวแปรอีกตัวหนึ่ง (X) โดยมีสมการแสดงความสัมพันธ์ของข้อมูล คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$y = a + \beta x \qquad \text{.....2.2}$$

โดย y = ตัวแปรตาม

x = ตัวแปรนำ

α และ β ค่าสัมประสิทธิ์ของสมการถดถอย โดยสามารถคำนวณได้จากสมการ ดังต่อไปนี้

$$\beta = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad \text{.....2.3}$$

$$\alpha = \bar{y} - \beta\bar{x} \qquad \text{.....2.4}$$

2.3.2.2 การแปลงค่าข้อมูล (Data Transformation)

เป็นกระบวนการในการปรับขอบเขตของข้อมูลให้อยู่ในช่วงที่เหมาะสมต่อการนำไปใช้งานในการสอนให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ สำหรับในแบบจำลองที่ใช้หลักการของโครงข่ายประสาทเทียมนั้น กระบวนการแปลงค่าข้อมูลที่ต้องนำมาใช้ คือ การนอมัลไลซ์ข้อมูล (Normalization) ซึ่งเป็นการลดค่าของข้อมูลให้อยู่ในขอบเขตที่น้อยลง เพื่อให้เหมาะสมกับฟังก์ชันที่ใช้งานในอัลกอริทึมที่นำมาใช้ในการเรียนรู้ของเน็ตเวิร์ค ซึ่งวิธีการนอมัลไลซ์ข้อมูลนั้นมีหลายวิธี แต่ที่มีการนำมาใช้กันอย่างกว้างขวางได้แก่ การแปลงค่าข้อมูลในลักษณะเป็นเชิงเส้น (Min-max Normalization)

โดยถ้ากำหนดให้ $\min_A$ และ $\max_A$ เป็นค่าต่ำสุด และค่าสูงสุดของตัวแปร A ตามลำดับ วิธีการของการแปลงค่าข้อมูลแบบเชิงเส้น จะทำการแปลงค่าข้อมูลใดๆ (V) ของตัวแปร A ให้เป็นค่าข้อมูลใหม่ (V') ที่อยู่ในช่วง (new_min_A, new_max_A) โดยการคำนวณจากสูตร ดังต่อไปนี้

$$V' = [(\frac{V - \min_A}{\max_A - \min_A})(new_\max_A - new_\min_A)] + new_\min_A \qquad \text{.....2.5}$$

ซึ่งหลังจากการประมวลผลข้อมูลเบื้องต้นแล้ว จะได้ข้อมูลที่มีความสมบูรณ์มากยิ่งขึ้น และสามารถนำไปใช้สร้างแบบจำลองพยากรณ์ได้อย่างมีประสิทธิภาพต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

หลักการของโครงข่ายประสาทเทียม (Artificial Neural Network)

การสร้างแบบจำลองพยากรณ์ เพื่อใช้ในงานด้านการพยากรณ์ข้อมูลลักษณะที่เป็นอนุกรมเวลานั้นได้มีการนำเครื่องมือทางสถิติมาใช้ เช่น การวิเคราะห์สมการแสดงความถดถอย (Linear auto-regressive model) ซึ่งวิธีที่นำมาใช้ส่วนใหญ่นั้น จะเป็นรูปแบบของการตั้งสมมติฐานจากลักษณะที่ข้อมูลเหล่านั้นมีความสัมพันธ์กันในลักษณะที่เป็นเชิงเส้น แต่ในความเป็นจริงนั้นข้อมูลอนุกรมเวลาส่วนมากนั้นจะไม่มีความสัมพันธ์ในลักษณะเชิงเส้นตามที่ได้ตั้งสมมติฐานไว้ ดังนั้นการวิเคราะห์และพยากรณ์ข้อมูลโดยใช้แบบจำลองพยากรณ์ทางสถิตินั้นจึงทำได้ค่อนข้างยาก ซึ่งผลจากการศึกษาได้มีการระบุข้อด้อยของแบบจำลองพยากรณ์ที่เป็นเชิงเส้นไว้ เช่น แบบจำลองพยากรณ์ในลักษณะเชิงเส้นนั้น ไม่สามารถจัดการกับข้อมูลที่มีค่าเปลี่ยนแปลงอย่างมากในช่วงเวลาที่ไม่แน่นอนได้ และยังไม่สามารถจัดการกับข้อมูลอินพุตที่มีลักษณะความแปรปรวนมากๆ ได้

ดังนั้นจึงได้มีความพยายามนำแบบจำลองพยากรณ์ ในลักษณะที่ไม่เป็นเชิงเส้นมาประยุกต์ใช้งาน เพื่อแก้ปัญหาข้อด้อยของแบบจำลองพยากรณ์ในลักษณะเชิงเส้น ตามที่ได้กล่าวมาแล้วข้างต้น เช่น แบบจำลองพยากรณ์ที่ใช้ทฤษฎีของโครงข่ายประสาทเทียม ซึ่งเป็นแบบจำลองพยากรณ์ที่สามารถปรับเปลี่ยนโครงสร้าง หรือตัวแปรในแบบจำลองพยากรณ์ให้ใช้กับสภาพข้อมูลหรือสภาพการใช้งานที่แตกต่างกัน และใช้งานได้ง่ายกว่าการปรับแบบจำลองพยากรณ์ทางสถิติที่จะต้องอาศัยผู้ที่มีความเชี่ยวชาญเฉพาะด้านนี้ หรือผู้ที่สร้างแบบจำลองพยากรณ์นั้นๆมาปรับเปลี่ยน และจากผลการพัฒนาและประยุกต์ใช้งานอย่างต่อเนื่อง พบว่าโครงข่ายประสาทเทียมนั้นให้ผลลัพธ์ที่ดีกว่ากับการนำไปใช้วิเคราะห์ข้อมูลที่มีลักษณะไม่เป็นเชิงเส้น ถึงแม้ว่าจะไม่สามารถใช้ทฤษฎีที่เกี่ยวกับสภาวะนั้นๆ อธิบายผลลัพธ์ที่ได้จากแบบจำลองพยากรณ์นั้นๆ โดยตรงก็ตาม ดังนั้นจึงอาจจะมีการพิจารณาแบบจำลองพยากรณ์ได้ในลักษณะเดียวกับ กล่องดำ (Black Box) ที่ไม่สามารถอธิบายกระบวนการทำงานภายในระบบได้

ถึงแม้ในช่วงเวลาต่อมาได้มีความพยายามพัฒนาวิธีการทางสถิติ เพื่อสร้างแบบจำลองพยากรณ์ที่ไม่เป็นเชิงเส้นสำหรับไว้ใช้งานในด้านการพยากรณ์ขึ้นอีกมากมาย แต่จากการเปรียบเทียบผลลัพธ์ของโครงข่ายประสาทเทียม กับผลลัพธ์จากแบบจำลองที่เป็นเชิงเส้นที่ถูกพัฒนาแล้ว พบว่า ในการพยากรณ์ข้อมูลในระยะยาวแบบจำลองพยากรณ์ทั้งสองมีลักษณะใกล้เคียงกัน แสดงว่ามีประสิทธิภาพใกล้เคียงกัน แต่สำหรับผลลัพธ์จากการพยากรณ์ข้อมูลในระยะสั้นๆนั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โครงข่ายประสาทเทียมสามารถตอบสนองการใช้งานได้ดีกว่ามาก อย่างไรก็ตามการนำโครงข่ายประสาทเทียมมาใช้งานนั้น ไม่ได้รับรองว่าจะให้ผลลัพธ์ที่ดีกว่าแบบจำลองพยากรณ์ที่เป็นเชิงเส้นทุกครั้ง ในบางครั้งกรณีที่เป็นการพยากรณ์ในระยะยาว หรือในกรณีที่โครงสร้างของโครงข่ายประสาทเทียมไม่เพียงพอที่จะครอบคลุมเพื่อแก้ปัญหานั้นๆ แบบจำลองพยากรณ์ทางสถิติก็อาจถูกนำมาใช้โดยให้ผลลัพธ์ หรือประสิทธิภาพได้ดีกว่า ดังนั้นในการใช้งานจึงต้องพิจารณาลักษณะของการพยากรณ์ตลอดจนโครงสร้างของโครงข่ายประสาทเทียม ให้เหมาะสมกับงานหรือปัญหาที่นำไปใช้ด้วย

สำหรับในบทนี้จะอธิบายถึงหลักการทั่วไปของทฤษฎีโครงข่ายประสาทเทียม ประเภทของโครงข่ายประสาทเทียม ที่มีการนำไปใช้งานด้านการพยากรณ์ข้อมูล โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม ที่นำมาใช้ในการพัฒนาระบบพยากรณ์การเปลี่ยนแปลงของราคาหลักทรัพย์ทั้งในด้านองค์ประกอบของการสร้างแบบจำลองพยากรณ์ การคำนวณของโหนดภายในโครงข่ายประสาทเทียม และการปรับโครงสร้างของโครงข่ายประสาทเทียมตามอัลกอริทึม ที่นำมาใช้ที่ทำให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ และสร้างแบบจำลองพยากรณ์ที่เหมาะสมสำหรับระบบที่จะพัฒนาขึ้น

3.1 ลักษณะการทำงานของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมมีวิธีการที่นำพื้นฐานการคำนวณของแบบจำลองทางคณิตศาสตร์มาใช้ในการสร้างให้เน็ตเวิร์ค (Network) เกิดการเรียนรู้ ซึ่งจะมีการทำงานในลักษณะเช่นเดียวกับโครงข่ายระบบประสาทของมนุษย์ ซึ่งประกอบด้วยเซลล์ประสาท หรือนิวรอล (Neural) ตามโครงสร้างในรูปที่ 3.1 จำนวนมากเชื่อมต่อกันเป็นโครงข่าย โดยในแต่ละเซลล์จะประกอบไปด้วย

3.1.1 เดนไดรต์ (Dendrite) ทำหน้าที่รับสัญญาณอินพุต

3.1.2 ตัวเซลล์ (Soma) ทำหน้าที่เป็นจุดรวมสัญญาณ

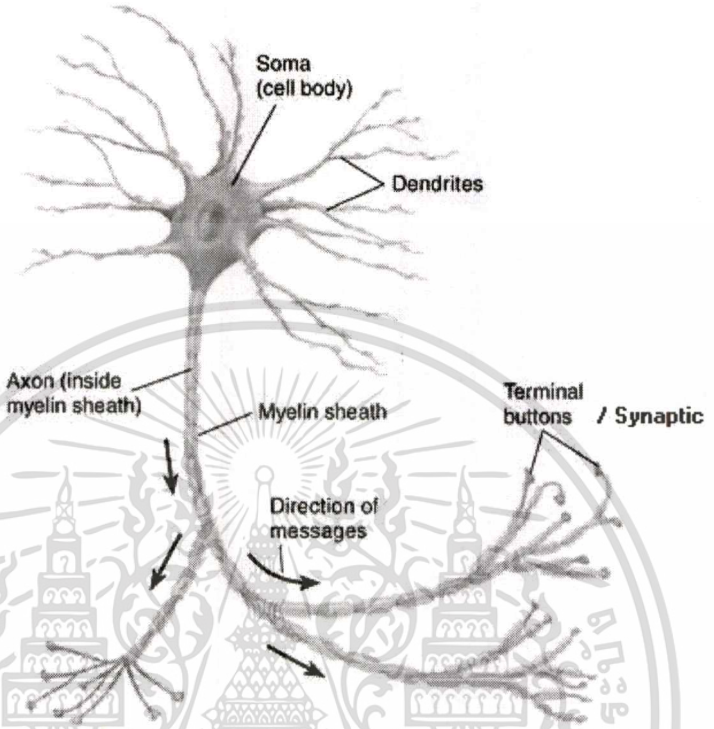
3.1.3 แอกซอน (Axon) ทำหน้าที่ส่งสัญญาณเอาต์พุต

3.1.3 ซินแนปติก (Synaptic) ทำหน้าที่เป็นจุดเชื่อมต่อระหว่างเซลล์

โดยการทำงานของเซลล์เริ่มจาก เดนไดรต์ (Dendrite) ที่ทำหน้าที่เหมือนอินพุตจะรับข้อมูลมาจากช่องว่างซินแนปติก (Synaptic) โดยวิธีการทางเคมี ซึ่งเราจะเรียกว่าเป็นการถ่วงน้ำหนัก (Weight) จากนั้นตัวเซลล์ (Soma) จะทำหน้าที่รวมอินพุตต่างๆ ที่ผ่านการถ่วงน้ำหนักแล้วมาประมวลผลสัญญาณโดยใช้วิธีการกระตุ้น ซึ่งเรียกว่าเป็นฟังก์ชันกระตุ้น (Activation Function) ของเซลล์ ทำให้ได้เอาต์พุตออกมา และเอาต์พุตที่ได้จากตัวเซลล์นี้ก็ส่งเป็นอินพุตให้กับเซลล์ประสาทอีกตัวหนึ่งต่อเนื่องกันต่อไป ซึ่งภายในร่างกายมนุษย์จะเกิดปฏิกิริยาเช่นนี้นับเป็นล้านๆต่อ 1 วินาที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

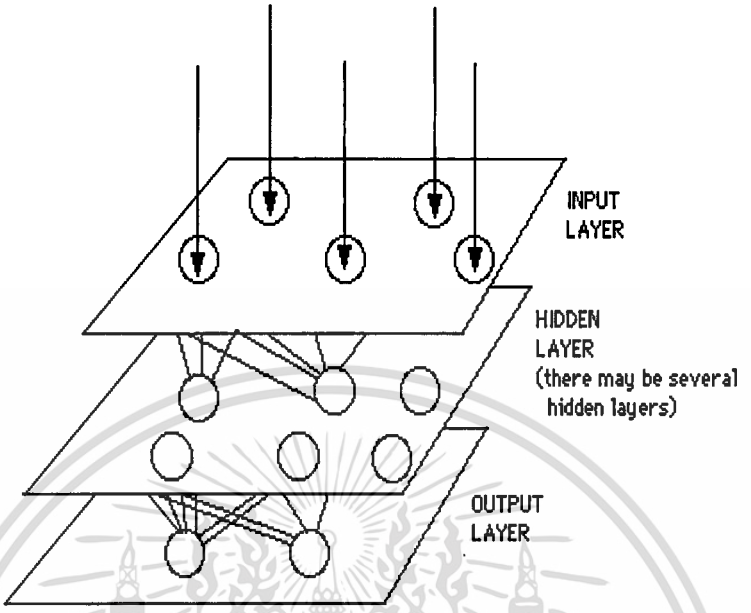
ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.



รูปที่ 3.1 แสดงโครงสร้างเซลล์ประสาทของมนุษย์

สำหรับในโครงข่ายประสาทเทียมนั้น ก็มีการทำงานในลักษณะเช่นเดียวกัน คือ มีการส่งข้อมูลอินพุตให้โครงข่ายเกิดการเรียนรู้จากระบบที่เชื่อมต่อกันเป็นโครงข่ายจากอินพุตจนถึงเอาต์พุต โดยในแบบจำลองของโครงข่ายประสาทเทียมนั้นจะมีการเชื่อมต่อกันของข้อมูลในลักษณะที่เป็นลำดับชั้น (Layer) ดังแสดงโครงสร้างพื้นฐานได้ดังรูปที่ 3.2 แต่อย่างไรก็ตามถึงแม้ว่าแบบจำลองโครงข่ายประสาทเทียมนี้จะมีข้อดีตรงที่มีความยืดหยุ่นกับข้อมูลที่มีลักษณะผิดปกติไปจากกลุ่มมาก และมีความสามารถในการสร้างแบบจำลองพยากรณ์จากข้อมูลที่ไม่เคยมีการเรียนรู้มาก่อนได้เป็นอย่างดี แต่ก็ยังมีส่วนที่เป็นข้อด้อย คือ ต้องใช้เวลามากในการสอนให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ก่อนที่จะนำแบบจำลองพยากรณ์ไปใช้พยากรณ์ค่าข้อมูลได้ และนอกจากนี้ยังไม่สามารถอธิบายการสร้างแบบจำลองของโครงข่ายประสาทเทียมได้อย่างชัดเจน ดังที่ได้กล่าวมาแล้วอีกด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.2 แสดงโครงสร้างพื้นฐานของแบบจำลองโครงข่ายประสาทเทียม

3.2 โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม

จากรูปที่ 3.2 แสดงให้เห็นว่าโครงข่ายประสาทเทียม เป็นลักษณะของโครงข่ายที่มีการเชื่อมต่อกันของกลุ่มอินพุตในชั้นแรก และกลุ่มของเอาต์พุตในชั้นสุดท้าย โดยผ่านโหนดต่างๆ ตามลำดับชั้นของการเชื่อมต่อ ซึ่งในการเชื่อมต่อกันระหว่างแต่ละโหนดจากอินพุตจนถึงเอาต์พุตนั้นจะมีค่าการถ่วงน้ำหนัก (Weight) เฉพาะที่แตกต่างกันไป โดยในระหว่างการฝึกสอนให้โครงข่ายประสาทเทียมได้เรียนรู้นั้น ค่าถ่วงน้ำหนักเหล่านี้จะถูกปรับเปลี่ยนไป ตามค่าที่ได้จากการคำนวณตามอัลกอริทึมที่ใช้จนกระทั่งได้ค่าที่เหมาะสมที่สุด ซึ่งองค์ประกอบที่สำคัญของการพัฒนาแบบจำลองโครงข่ายประสาทเทียม มีดังนี้

- **โครงสร้างของโครงข่าย (Architecture)** ประกอบด้วยจำนวนของชั้นภายในโครงข่าย จำนวนของโหนดในชั้นที่เป็นอินพุต (Input Layer) และชั้นเอาต์พุต (Output Layer) รวมถึงจำนวนโหนดในชั้นกลางแต่ละชั้น (Hidden Layer) ที่เชื่อมต่ออยู่ระหว่างชั้นอินพุต และชั้นเอาต์พุตด้วย
- **ฟังก์ชันกระตุ้น (Transfer Function, Activation Function)** ที่ใช้คำนวณค่าที่ได้จาก แต่ละโหนดภายในโครงข่ายจากอินพุตจนถึงเอาต์พุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **ฟังก์ชันที่ใช้ทดสอบหาค่าความผิดพลาดของผลลัพธ์ที่ได้ (Cost Function)** โดยเป็นการเปรียบเทียบค่าผลลัพธ์ที่คำนวณได้จากแบบจำลองพยากรณ์ กับค่าของข้อมูลที่เกิดขึ้นจริง
- **อัลกอริทึมที่ใช้สอนให้โครงข่ายเกิดการเรียนรู้ (Training Algorithm)** โดยมีการกำหนดวิธีการปรับค่าตัวแปรต่างๆ ภายในโครงข่าย เพื่อให้ค่าความผิดพลาดที่คำนวณได้จากฟังก์ชันทดสอบ มีค่าความผิดพลาดน้อยที่สุด

3.2.1 การคำนวณในโครงสร้างในโครงข่ายประสาทเทียม (Neural Computation)

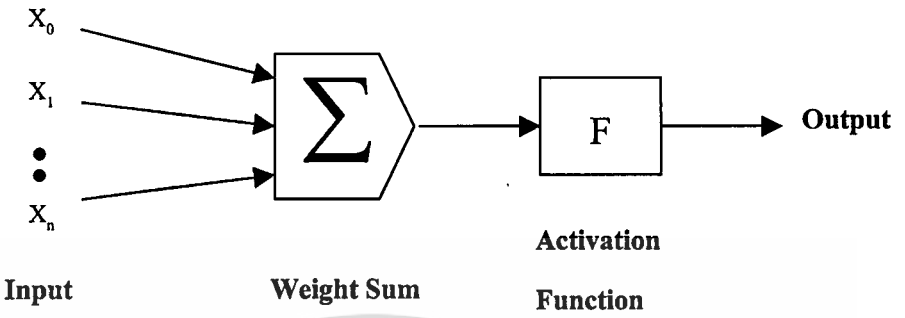
คุณสมบัติที่น่าสนใจของโครงข่ายประสาทเทียม คือ ความสามารถในการเรียนรู้ของโครงข่าย ซึ่งปกติพฤติกรรม หรือตัวแปรต่างๆ ของโครงข่ายนั้นจะถูกพัฒนาในระหว่างที่มีกระบวนการเรียนรู้หรือกระบวนการสอน โดยในระหว่างกระบวนการนี้จะมีการคำนวณในลักษณะซ้ำๆ กับข้อมูลที่นำมาใช้เป็นอินพุตหลายๆชุด ร่วมกับลักษณะของเอาต์พุตที่เกิดขึ้นสำหรับอินพุตแต่ละชุด โดยโครงข่ายจะมีการเปลี่ยนแปลงโดยทำการปรับค่าของตัวแปรต่างๆ ภายในโครงสร้างตามอัลกอริทึมที่นำมาใช้ในการเรียนรู้ จนกระทั่งสามารถสร้างเอาต์พุตได้ตามที่ออกแบบไว้ได้ หรือภายใต้เงื่อนไขที่ยอมรับได้ ซึ่งกลุ่มข้อมูลอินพุตและเอาต์พุตที่นำมาใช้นี้เรียกว่า กลุ่มข้อมูลที่ใช้สอน (Training Set)

ในการสอนให้โครงข่ายเกิดการเรียนรู้นั้น จะเกิดประโยชน์หรือเกิดประสิทธิภาพสูงสุด ก็ต่อเมื่อข้อมูลที่นำมาใช้สอนต้องมีลักษณะครอบคลุมรูปแบบที่เป็นไปได้ของข้อมูลให้มากที่สุด ดังนั้นจึงต้องให้ความสำคัญในการเลือกข้อมูลที่เป็นตัวแทนของทุกรูปแบบของงานที่จะนำแบบจำลองพยากรณ์ที่ได้ไปใช้งาน โดยลักษณะ 2 ประการที่ควรคำนึงถึงในการพิจารณาเลือกข้อมูลที่นำมาใช้ในการสอน คือ ข้อมูลที่ใช้สอนต้องเป็นส่วนหนึ่งของข้อมูลที่คาดหวังว่าจะนำแบบจำลองไปใช้ประมวลผล และข้อมูลที่ใช้สอนต้องถูกเลือกจากช่วงเวลาที่มีการกระจายของข้อมูลครอบคลุมรูปแบบของอินพุตที่เป็นไปได้ทั้งหมด เนื่องจากโครงข่ายไม่สามารถพยากรณ์ข้อมูลได้อย่างมีประสิทธิภาพ ถ้ามีการสอนด้วยข้อมูลที่มีค่าผิดปกติจำนวนมาก

3.2.2 การคำนวณของโหนด ในโครงข่ายประสาทเทียม

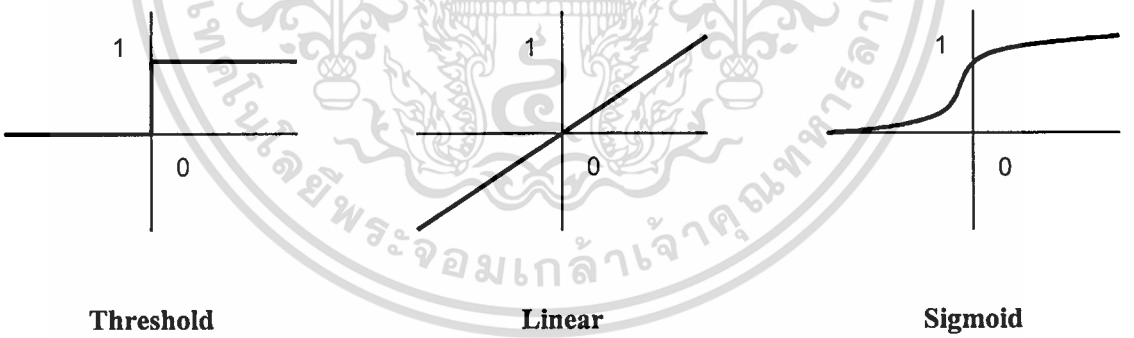
แบบจำลองอย่างง่ายของการคำนวณในแต่ละโหนดภายในโครงข่ายประสาทเทียม แสดงได้ดังรูปที่ 3.3 ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 แสดงแบบจำลองการคำนวณของโหนดภายในโครงข่ายประสาทเทียม

ในการคำนวณนั้นสามารถแบ่งประเภทของโหนด ตามรูปแบบของฟังก์ชันกระตุ้น (Transfer Function) ที่ใช้ได้เป็น 3 แบบ แสดงตามรูปที่ 3.4 ดังนี้



รูปที่ 3.4 แสดงรูปแบบของฟังก์ชันกระตุ้นที่ใช้ในการคำนวณค่าในโหนดของโครงข่าย

- **แบบที่ 1 Threshold Unit** เป็นฟังก์ชันกระตุ้นที่ถูกจำกัดค่าเอาต์พุตไว้เป็น 0 หรือ 1 ตามสถานะก่อนหน้า
- **แบบที่ 2 Linear Unit** ใช้ฟังก์ชัน f(n) = h
- **แบบที่ 3 Non-Linear Unit** โดยทั่วไปจะใช้การเรียนรู้แบบ Gradient Descent ซึ่งส่วนมากโหนดประเภทนี้จะใช้ฟังก์ชันกระตุ้นของซิกมอยด์(Sigmoid Function)เช่น $f(n) = 1/(1+e^{-h})$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยทั่วไปฟังก์ชันกระตุ้นที่นิยมใช้มากที่สุด คือ ฟังก์ชันของซิกมอยด์ (Sigmoid) เนื่องจากสามารถหาค่า Differentiate และคำนวณค่า Gradient ได้ง่าย ส่วนฟังก์ชันแบบเชิงเส้นนั้นนิยมใช้ในส่วนที่เป็นโหนดของชั้นเอาต์พุตของโครงข่ายในการนำไปประยุกต์ ใช้กับข้อมูลที่มีลักษณะเป็นอนุกรมเวลา

3.2.3 อัลกอริทึมที่ใช้ในการสอน (Training Algorithm)

อัลกอริทึมที่ใช้สอนนั้นเป็นการกำหนดวิธีการคำนวณภายในเน็ตเวิร์ค เพื่อใช้ในการปรับค่าตัวแปรภายในโครงสร้างให้ได้แบบจำลองพยากรณ์ที่เหมาะสมที่สุดสำหรับข้อมูลแต่ละชุด โดยอัลกอริทึมที่ได้รับการยอมรับในการนำมาใช้กับงานด้านการพยากรณ์ข้อมูล ได้แก่ อัลกอริทึมแบล็คพร็อบพาเกชั่น (Backpropagation)

ลักษณะของอัลกอริทึมแบล็คพร็อบพาเกชั่น (Backpropagation) นี้เป็นวิธีการที่ทำให้โครงข่ายมีกระบวนการเรียนรู้ซ้ำๆจากข้อมูลในอดีต โดยที่เอาต์พุตที่ได้จากโครงข่ายในแต่ละครั้งที่ป้อนชุดข้อมูลอินพุตให้กับโครงข่ายนั้นจะถูกนำมาเปรียบเทียบกับเอาต์พุตที่เกิดขึ้นจริงในอดีต เพื่อหาค่าความผิดพลาดสำหรับแต่ละชุดของข้อมูลอินพุต และค่าความผิดพลาดที่ได้นี้จะนำมาใช้ในการคำนวณเพื่อปรับค่าถ่วงน้ำหนักระหว่างแต่ละโหนดในโครงสร้างของโครงข่ายต่อไป โดยที่การปรับค่าถ่วงน้ำหนักนี้จะมีการคำนวณในทิศทางย้อนกลับจากทางด้านเอาต์พุตผ่านชั้นกลาง (Hidden Layer) จนปรับค่ามาถึงชั้นแรกที่ต่อจากชั้นอินพุต จากนั้นจึงคำนวณค่าของเอาต์พุตสำหรับข้อมูลอินพุตชุดต่อๆไป โดยใช้ค่าถ่วงน้ำหนักที่คำนวณใหม่นี้ เพื่อเปรียบเทียบหาค่าความผิดพลาดในลักษณะเช่นเดียวกัน โดยในกระบวนการเรียนรู้จะมีการคำนวณเช่นนี้เรื่อยไป จนกระทั่งค่าความผิดพลาดที่ได้มีค่าน้อยที่สุด หรือเป็นค่าที่ยอมรับได้เมื่อเทียบกับค่าเอาต์พุตที่คำนวณได้กับค่าข้อมูลที่เกิดขึ้นจริง ซึ่งจากโครงสร้างในรูปที่ 3.2 สามารถสรุปขั้นตอนการคำนวณของอัลกอริทึมได้ ดังต่อไปนี้

- ขั้นตอนที่ 1 การกำหนดค่าถ่วงน้ำหนักเริ่มต้น (Initial Weight)

เป็นขั้นตอนการกำหนดค่าตัวเลขสุ่มที่มีค่าน้อยๆ ให้กับค่าถ่วงน้ำหนัก (Weight) สำหรับการเชื่อมต่อระหว่างแต่ละโหนดจากชั้นอินพุตถึงชั้นเอาต์พุต รวมถึงการสุ่มค่า Bias ให้กับแต่ละโหนดด้วย

- ขั้นตอนที่ 2 การหาค่าของอินพุตในทิศทางไปข้างหน้า (Propagate the input forward)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นขั้นตอนในการคำนวณค่าอินพุต และค่าเอาต์พุตให้กับแต่ละโหนดในโครงข่ายโดยข้อมูลในอดีตที่นำมาใช้สอนให้เกิดการเรียนรู้ภายในโครงสร้างนั้นจะถูกป้อนเข้าสู่ชั้นอินพุตของโครงข่าย และจะมีการคำนวณค่าอินพุตรวม และค่าเอาต์พุตของแต่ละโหนด ดังต่อไปนี้

<u>การคำนวณค่าในชั้นอินพุต</u> สำหรับค่าเอาต์พุตของโหนด j ใดๆ ในชั้นของอินพุตมีค่าเท่ากับค่าของอินพุตของโหนดนั้นๆ

$$Oj = Ij \qquad \text{.....3.1}$$

โดยที่ Oj = ค่าเอาต์พุตของโหนด j

Ij = ค่าอินพุตของโหนด j

<u>การคำนวณค่าในชั้นกลาง (Hidden Layer) และชั้นเอาต์พุต</u> สำหรับค่าอินพุตของโหนด j ใดๆที่อยู่ในชั้นกลางหรือชั้นเอาต์พุตจะมีค่าเท่ากับการบวกค่า Bias ของโหนด j กับผลรวมทั้งหมดของการคูณค่าเอาต์พุตของทุกๆโหนดที่เชื่อมต่อกับโหนด j กับค่าถ่วงน้ำหนักของการเชื่อมต่อนั้นๆ ดังสมการต่อไปนี้

$$Ij = \sum WijOi + \theta j \qquad \text{.....3.2}$$

โดยที่ θj = ค่า bias ของโหนด j

Wij = ค่าถ่วงน้ำหนักของการเชื่อมต่อระหว่างโหนด i กับโหนด j

ส่วนค่าของเอาต์พุตสำหรับโหนดใดๆในชั้นกลาง หรือชั้นเอาต์พุตนั้น ได้จากการคำนวณค่าอินพุตรวม สำหรับแต่ละโหนดผ่านฟังก์ชันกระตุ้นที่ใช้ทำการคำนวณผลที่ได้จากโหนดนั้นๆ ซึ่งในอัลกอริทึมนี้จะใช้ฟังก์ชันของซิกมอยด์โดยได้ค่าของเอาต์พุตตามสมการต่อไปนี้

$$Oj = \frac{1}{1+e^{-Ij}} \qquad \text{.....3.3}$$

- ขั้นตอนที่ 3 การคำนวณค่าความผิดพลาดย้อนกลับ (Backpropagation the error)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อได้ค่าของเอาต์พุตครบทุกๆ โหนดแล้ว ก็จะคำนวณค่าความผิดพลาดของเอาต์พุตที่ได้ในทิศทางย้อนกลับเพื่อปรับค่าถ่วงน้ำหนักสำหรับการเชื่อมต่อระหว่างโหนดใดๆ ในโครงข่ายรวมทั้งปรับค่า bias ให้กับแต่ละโหนดด้วย

<u>การคำนวณหาค่าความผิดพลาดในชั้นเอาต์พุต</u> สำหรับโหนด k ใดๆ ในชั้นเอาต์พุต สามารถคำนวณค่าความผิดพลาด (Err_k) ได้ดังสมการต่อไปนี้

$$Err_k = O_k(1-O_k)(T_k - O_k) \qquad \text{.....3.4}$$

โดยที่ O_k = ค่าเอาต์พุตของโหนด k ที่ได้จากโครงข่าย

T_k = ค่าเอาต์พุตจริงของโหนด k ที่ได้จากข้อมูลในอดีต

<u>การคำนวณค่าผิดพลาดในชั้นกลาง (Hidden Layer)</u> สำหรับโหนดในชั้นกลางนี้จะมีการนำค่าถ่วงน้ำหนักของการเชื่อมต่อระหว่างโหนด j ใดๆที่พิจารณา กับโหนดอื่นๆที่อยู่ในลำดับชั้นที่สูงกว่าในโครงข่ายมาคำนวณร่วมกัน ดังสมการต่อไปนี้

$$Err_j = O_j(1-O_j)\sum Err_k W_{jk} \qquad \text{.....3.5}$$

โดยที่ Err_k = ค่าความผิดพลาดของโหนดkในลำดับชั้นที่สูงกว่า

W_{jk} = ค่าถ่วงน้ำหนักของการเชื่อมต่อระหว่างโหนด j กับโหนด k ใดๆในชั้นถัดไปที่สูงกว่าในโครงข่าย

ค่าถ่วงน้ำหนักของการเชื่อมต่อระหว่างโหนดใดๆจะถูกปรับเปลี่ยนด้วยค่าความผิดพลาดที่เกิดขึ้นต่อเนื่องกันไปตามสมการต่อไปนี้

$$\Delta W_{ij} = (l)Err_j O_i \qquad \text{.....3.6}$$

$$W_{ij} = W_{ij} + \Delta W_{ij} \qquad \text{.....3.7}$$

และค่า bias ใหม่ของโหนดใดๆคำนวณได้จาก

$$\Delta\theta_j = (l)Err_j \qquad \text{.....3.8}$$

$$\theta_j = \theta_j + \Delta\theta_j \qquad \text{.....3.9}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ ΔW_{ij} = ค่าการเปลี่ยนแปลงของ W_{ij}

$\Delta\theta_j$ = ค่าการเปลี่ยนแปลงของ θ_j

l = ค่าอัตราการเรียนรู้ (Learning Rate)

สำหรับ "l" หมายถึง อัตราการเรียนรู้ (Learning Rate) ของโครงข่ายซึ่งเป็นค่าที่ใช้ควบคุมความเร็วในการตอบสนองต่อค่าความผิดพลาดที่เกิดขึ้นในแต่ละครั้ง โดยในกรณีที่ "l" มีค่ามากแสดงว่าเน็ตเวิร์คมีการตอนสนองต่อค่าความผิดพลาดมาก โดยการปรับค่าถ่วงน้ำหนักให้เปลี่ยนแปลงไปอย่างชัดเจน แต่ถ้า "l" มีค่าน้อยๆแสดงว่ามีการปรับค่าถ่วงน้ำหนักอย่างค่อยเป็นค่อยไป จนกระทั่งได้ค่าที่เหมาะสมที่ทำให้เกิดค่าความผิดพลาดของเอาต์พุตน้อยที่สุด

3.2.4 การวัดค่าความผิดพลาด (Error)

การเรียนรู้ของโครงข่ายนั้นขึ้นอยู่กับค่าความผิดพลาด (Error) ที่วัดได้โดยยึดหลักแนวความคิดที่ว่า ยิ่งค่าความผิดพลาดที่วัดได้ หรือค่าจากฟังก์ชันทดสอบความผิดพลาด (Cost Function) ที่คำนวณมาได้มีน้อยเท่าไร นั่นหมายความว่าค่าถ่วงน้ำหนักต่างๆภายในโครงสร้างของโครงข่ายยิ่งดีขึ้น หรือมีความเหมาะสมมากขึ้นนั่นเอง โดยการที่มีฟังก์ชันการวัดค่าความผิดพลาดนั้นทำให้สามารถปรับปรุงค่าถ่วงน้ำหนักได้ในลักษณะคล้ายๆ กับการค่อยๆ ไต่ระดับลงมาตามพื้นผิวของหุบเขา เพื่อหาตำแหน่งที่ต่ำที่สุดของหุบเขา หรือภาวะที่ค่าความผิดพลาดของโครงข่ายน้อยที่สุดนั่นเอง

ส่วนมากในการคำนวณค่าความผิดพลาดนั้น มักจะใช้ค่า Mean Squared Error (MSE) โดยสามารถคำนวณได้จากสมการ ดังต่อไปนี้

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(O_j - T_j)^2 \qquad \text{.....3.10}$$

โดยที่ N = จำนวนของชุดข้อมูลที่นำมาใช้ในการสอน

O_j = ค่าเอาต์พุตของโหนด j ที่ได้จากโครงข่าย

T_j = ค่าเอาต์พุตจริงของโหนด j จากข้อมูลในอดีต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.5 การพิจารณาจำนวนโหนดในชั้นอินพุตและชั้นกลาง

แม้ว่าจะมีความพยายามในการพัฒนาแบบจำลองพยากรณ์และอัลกอริทึมในการเรียนรู้ของโครงข่ายประสาทเทียมแล้ว แต่การออกแบบเพื่อให้ได้โครงสร้างของโครงข่ายที่เหมาะสมกับงานใดๆ นั้นก็ยังมีปัญหาอยู่บ้าง ในการหาโครงสร้างที่มีขนาดของโครงข่ายเล็กที่สุด (Parsimonious Network) ที่สามารถสร้างความผิดพลาดของเอาต์พุตสำหรับข้อมูลที่นำมาใช้ในการสอน และรวมถึงค่าความผิดพลาดของข้อมูลที่ไม่ใช่กรณีที่นำมาใช้สอนในระดับที่เหมาะสมและยอมรับได้ (Optimal Error) ซึ่งกฎโดยทั่วไปจะมีการพิจารณาให้ค่าถ่วงน้ำหนักทั้งหมดภายในโครงข่ายนั้นควรมีจำนวนน้อยกว่า 1/10 ของจำนวนชุดข้อมูลที่นำมาใช้ในการสอน (Training Record)

ในการเลือกจำนวนโหนดของชั้นอินพุตนั้น ถ้าใช้จำนวนอินพุตน้อยเกินไปจะมีผลทำให้แบบจำลองพยากรณ์ไม่ครอบคลุมลักษณะของข้อมูล (Inadequate Modeling) ในขณะที่ถ้าใช้จำนวนอินพุตมากเกินไปก็จะทำให้โครงสร้างของโครงข่ายมีความซับซ้อนมากเกินไป ทำให้เกิดผลกระทบทั้งในเรื่องของเวลาที่ใช้ในการสอนมากขึ้น และมิติของโครงข่ายที่ซับซ้อนขึ้นด้วย ซึ่งวิธีการหนึ่งที่ถูกนำมาเสนอเพื่อใช้พิจารณาจำนวนของโหนดในชั้นอินพุตของโครงข่าย คือการใช้หลักการของ Auto Correlation โดยการคำนวณค่าสัมประสิทธิ์ระหว่างข้อมูลอินพุตในช่วงเวลาที่เหลื่อมล้ำกัน (Lag) เพื่อหาระยะเวลาที่ข้อมูลอินพุตยังมีความสัมพันธ์กัน ว่ามีช่วงเวลาที่เหลื่อมล้ำกันอยู่กี่ช่วงในอนุกรมเวลานั่นเอง

ในส่วนของจำนวนโหนดในชั้นกลางนั้น หลักจากที่พิจารณาหาจำนวนโหนดในชั้นอินพุตแล้ว ก็อาจจะใช้การทดลองซ้ำๆ กับข้อมูลที่นำมาใช้ในการสอนโดยการเปลี่ยนจำนวนโหนดของชั้นกลางสำหรับการทดลองในแต่ละครั้งเพื่อหาจำนวนโหนดที่เหมาะสมที่สุด นอกจากนี้อาจพิจารณาได้จากกฎของ Baum-Hausster ซึ่งระบุวิธีการหาจำนวนของโหนดชั้นกลางของโครงข่ายไว้ดังสมการต่อไปนี้

$$N_{hidden} = \frac{N_{train} E_{tolerance}}{N_{pts} + N_{output}} \qquad \text{.....3.11}$$

โดยที่ N_{hidden} = จำนวนโหนดในชั้นกลาง

N_{train} = จำนวนของชุดข้อมูลที่ใช้ในการสอน

$E_{tolerance}$ = ความแปรปรวนของค่าความผิดพลาด (Error Tolerance) โดยทั่วไปจะพิจารณาที่ 0.01

N_{pts} = จำนวนข้อมูลในแต่ละชุดข้อมูลที่นำมาสอน

N_{output} = จำนวนของโหนดในชั้นเอาต์พุต

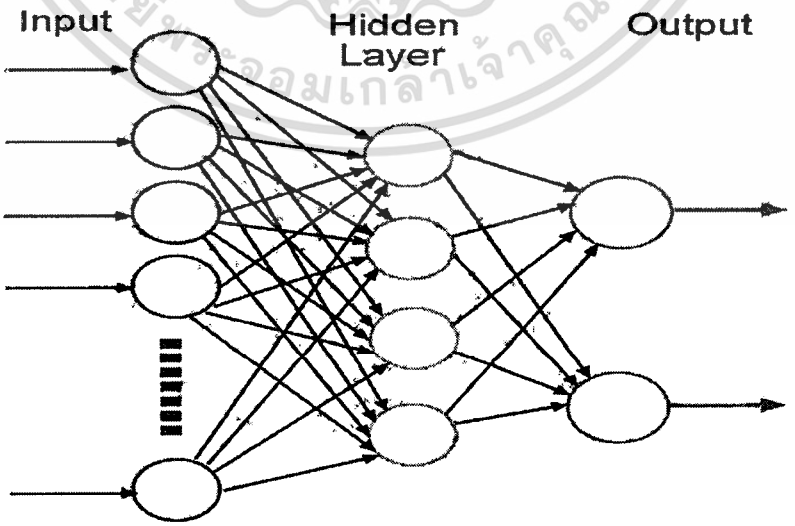
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 ประเภทของโครงข่ายประสาทเทียม

จากการพัฒนาอย่างต่อเนื่องในช่วงที่ผ่านมามีโครงข่ายประสาทเทียมหลายประเภท ที่ถูกนำมาประยุกต์ใช้กับงานด้านการพยากรณ์ ซึ่งโครงสร้างของโครงข่ายประสาทเทียมเหล่านี้ก็มีความแตกต่างกันในประเด็นต่างๆ เช่น ประเภทของโหนดที่นำมาใช้ในโครงข่าย ลักษณะโครงสร้างของโครงข่าย และอัลกอริทึมที่นำมาใช้ในการเรียนรู้ภายในโครงข่าย เป็นต้น โดยในส่วนต่อไปนี้จะอธิบายถึงรูปแบบโครงสร้างของโครงข่ายประสาทเทียมแต่ละประเภทที่แตกต่างกัน พร้อมทั้งผลกระทบในแง่ประสิทธิภาพของโครงข่ายอย่างสังเขป ซึ่งโครงสร้างหลักๆของโครงข่ายประสาทเทียมนั้นมีอยู่ 2 รูปแบบ คือ แบบที่มีการคำนวณส่งสัญญาณไปข้างหน้า (Feed Forward Neural Network) และแบบที่มีการป้อนย้อนกลับ (Recurrent Neural Network)

3.3.1 โครงสร้างแบบส่งสัญญาณไปข้างหน้า (Feed Forward Neural Network)

โครงสร้างแบบที่มีการคำนวณสัญญาณไปข้างหน้า หรือ Feed Forward Network นั้นเป็นโครงสร้างที่มีการจัดโครงข่ายในลักษณะที่เป็นลำดับชั้น (Layer) ดังแสดงตัวอย่างในรูปที่ 3.5 โดยภายในโครงข่ายนั้น จะประกอบด้วยกลุ่มของอินพุตที่มีหน้าที่ส่งสัญญาณอินพุตต่อไปให้กับโครงข่ายในส่วนที่เหลือ ซึ่งหลังจากชั้นแรกที่เป็นอินพุตนี้อาจจะมีชั้นกลางถัดไปอีกหนึ่ง หรือหลายๆชั้นก็ได้และชั้นสุดท้ายจะเป็นชั้นของเอาต์พุต ซึ่งเป็นค่าผลลัพธ์ของการคำนวณจากโหนดต่างๆ ภายในโครงข่ายทุกโหนด โดยในการคำนวณของโหนดต่างๆ ในแต่ละชั้นนั้นจะส่งค่าที่ได้ต่อไปให้กับโหนดที่อยู่ในชั้นถัดไปเท่านั้น



รูปที่ 3.5 แสดงโครงสร้างของโครงข่ายประสาทเทียมแบบส่งสัญญาณไปข้างหน้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

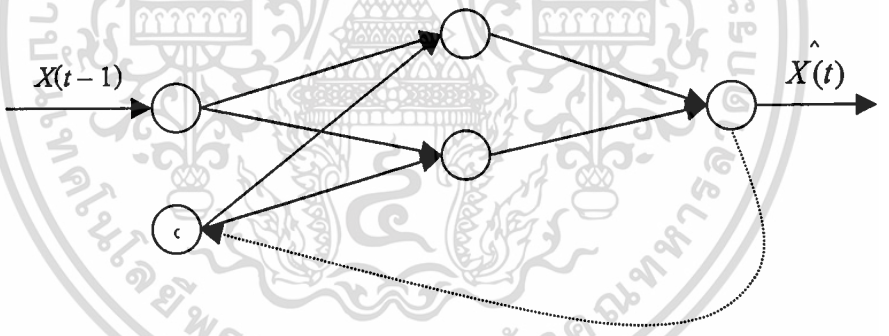
3.3.2 โครงสร้างแบบมีการป้อนย้อนกลับ (Recurrent Neural Network : RNN)

สามารถแบ่งได้เป็น 2 ลักษณะ ดังต่อไปนี้คือ

3.3.2.1 โครงสร้างแบบมีการป้อนย้อนกลับแบบง่าย (Simple Recurrent Network : SRN)

รูปแบบโครงสร้างของโครงข่ายแบบนี้นั้นจะใกล้เคียงกับลักษณะโครงสร้างแบบส่งสัญญาณไปข้างหน้า แต่จะมีส่วนที่เพิ่มขึ้นมาคือ มีการป้อนย้อนกลับ (Feedback Connection) จากบางชั้นของโครงข่ายมายังชั้นในลำดับก่อนหน้า เพื่อช่วยทำให้โครงข่ายสามารถเรียนรู้ และสามารถสร้างแบบจำลองพยากรณ์ที่มีความเหมาะสมกับข้อมูลในลักษณะที่เป็นอนุกรมเวลาได้มากขึ้น

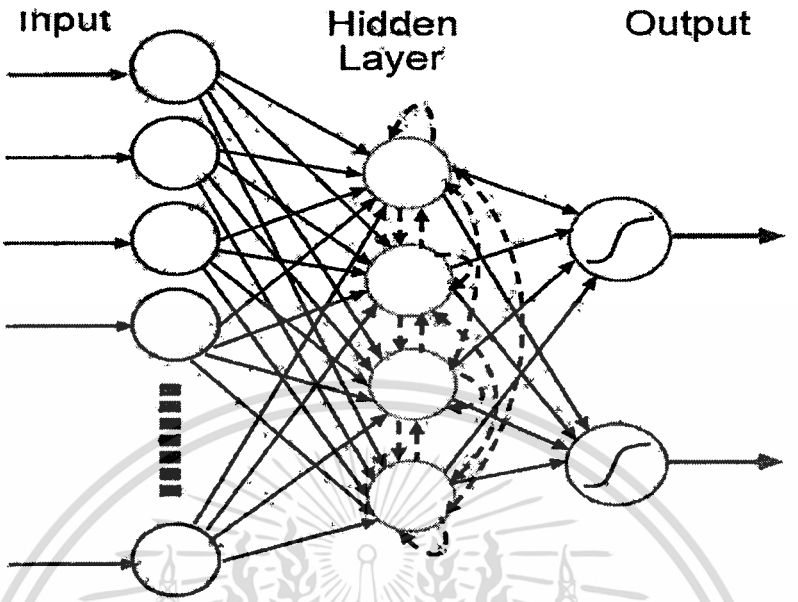
โครงสร้างแบบมีการป้อนกลับนี้เริ่มศึกษาครั้งแรกโดย Jordan ซึ่งแสดงโครงสร้างได้ดังรูปที่ 3.6 โดยจะเห็นได้ว่าส่วนที่เป็นอินพุตของโครงข่ายนี้จะรับค่าทั้งจากอินพุตภายนอกโครงข่าย (External Input) และจากอินพุตที่แสดงสถานะภายในของโหนดภายในโครงข่าย (State Unit) โดยข้อมูลในส่วนหลังนี้จะเป็นอินพุตที่รับข้อมูลมาจากการป้อนกลับค่าจากชั้นเอาต์พุตของโครงข่าย



รูปที่ 3.6 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบง่าย –Jordan Network

ในปัจจุบันได้มีการพัฒนาโครงสร้างอีกรูปแบบหนึ่ง คือ โครงสร้างแบบ Elman ดังแสดงไว้ดังรูปที่ 3.7 ซึ่งโครงสร้างนี้เป็นที่ยอมรับโดยทั่วไปว่าเป็นโครงสร้างแบบหนึ่งที่นำไปประยุกต์ใช้งานได้ง่าย และมีประสิทธิภาพ โดยในส่วนของชั้นอินพุตของโครงสร้างจะประกอบด้วยค่าที่ได้จากค่าเอาต์พุตจากชั้นกลางของการคำนวณในครั้งที่ผ่านมา โดยเก็บไว้ในส่วนของชั้น Context หรือ State Unit ที่ทำหน้าที่เสมือนชั้นความจำระยะสั้นของสถานะภายในโครงข่าย (Short Term Memory :STM) ซึ่งวิธีการง่ายๆในการเพิ่มความจำระยะสั้นให้กับโครงสร้างแบบ Elman คือการเพิ่มจำนวนโหนดในชั้นกลางของโครงข่ายนั่นเอง แต่ก็จะส่งผลให้การเรียนรู้มีความซับซ้อน และใช้เวลามากยิ่งขึ้นด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



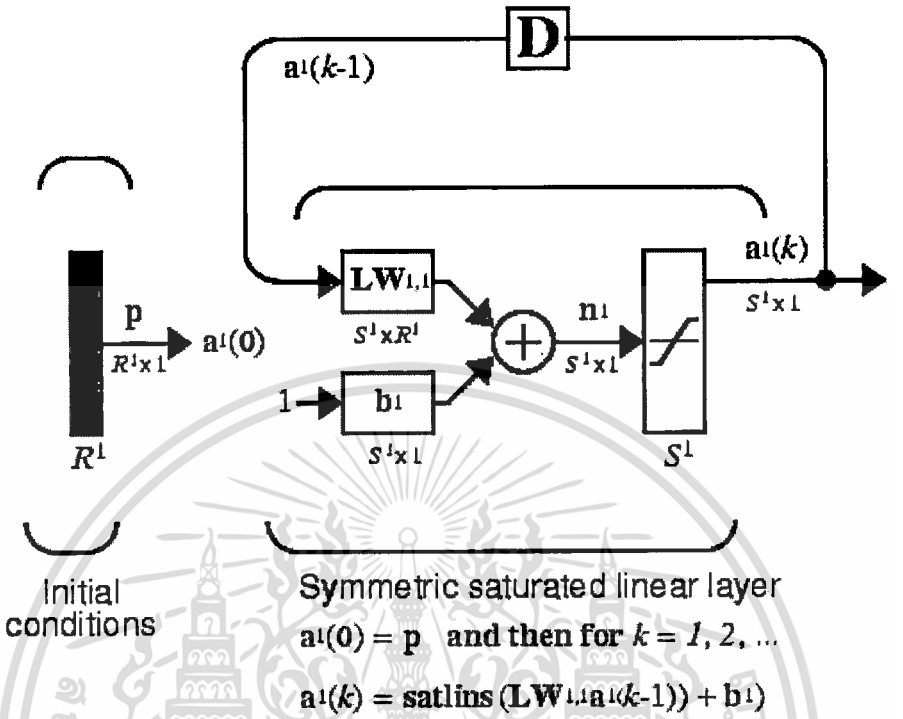
รูปที่ 3.7 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบง่าย –Elman Network

3.3.2.2 โครงสร้างแบบมีการป้อนย้อนกลับแบบสมบูรณ์ (Fully Recurrent Neural Network)

ความแตกต่างระหว่างโครงสร้างแบบมีการป้อนย้อนกลับแบบง่าย และโครงสร้างที่มีการป้อนย้อนกลับแบบสมบูรณ์ (Fully Recurrent Network) คือ โครงสร้างที่มีการป้อนย้อนกลับแบบสมบูรณ์นั้น จะมีการเชื่อมต่อในทุกรูปแบบทั้งแบบทางตรง (Direct loop) และทางอ้อม (Indirect loop) และค่าถ่วงน้ำหนักของการเชื่อมต่อระหว่างแต่ละโหนดในโครงข่ายนี้ สามารถมีการเรียนรู้ได้อย่างเป็นอิสระ ดังนั้นจึงทำให้การสอนโครงข่ายนั้นทำได้ยากและต้องใช้เวลานานยิ่งขึ้น แต่อย่างไรก็ตามในกรณีที่โครงสร้างแบบส่งสัญญาณไปข้างหน้านั้น ไม่สามารถให้ผลลัพธ์ตามที่ต้องการได้อย่างเหมาะสมแล้ว โครงสร้างที่มีการป้อนกลับแบบนี้ ก็เป็นอีกทางเลือกหนึ่งที่สามารถนำไปใช้ได้ ตัวอย่างของโครงสร้างแบบนี้ได้แก่ โครงสร้างของ Hopfield Network ดังในรูปที่ 3.8

สำหรับในโครงการพัฒนาระบบงานนี้ จะใช้หลักการของโครงข่ายประสาทเทียมแบบที่มีการป้อนย้อนกลับอย่างง่ายเพื่อใช้สร้างแบบจำลองพยากรณ์เพื่อใช้พยากรณ์การเปลี่ยนแปลงราคาหลักทรัพย์ โดยที่มีอินพุตตามที่กล่าวมาแล้วในบทที่ 2 และมีค่าเอาต์พุตเป็นค่าราคาหลักทรัพย์ในช่วงเวลาถัดไปในอนุกรมเวลา ซึ่งจะได้อธิบายขั้นตอนในการสร้างโครงข่ายในส่วนต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 แสดงโครงสร้างโครงข่ายประสาทเทียมแบบป้อนย้อนกลับแบบสมบูรณ์ –Hopfield Network

3.4 ขั้นตอนการพัฒนาโครงข่ายประสาทเทียมสำหรับระบบพยากรณ์ราคาหลักทรัพย์

ในการสร้างแบบจำลองพยากรณ์ เพื่อนำไปใช้ในการพยากรณ์ราคาหลักทรัพย์นั้น สามารถแบ่งกระบวนการออกได้เป็น 9 ขั้นตอนดังนี้

3.4.1 ขั้นตอนการกำหนดตัวแปรข้อมูล

เป็นการกำหนดตัวแปรข้อมูลที่ต้องการใช้สำหรับพยากรณ์ โดยจะนำตัวแปรหลายๆตัว มาใช้ในการพิจารณาตามที่กล่าวมาแล้วในบทที่ 2

3.4.2 ขั้นตอนการรวบรวมข้อมูล

เป็นการรวบรวมข้อมูลที่เกี่ยวข้องในลักษณะที่เป็นอนุกรมเวลา โดยจำนวนของข้อมูลต้องครอบคลุม และมีจำนวนมากพอสมควร ที่จะสอนให้โครงข่ายเกิดการเรียนรู้ได้อย่างมีประสิทธิภาพ

3.4.3 ขั้นตอนการกรองข้อมูล

เป็นขั้นตอนในการกลั่นกรองข้อมูลหรือประมวลผลข้อมูลเบื้องต้น เพื่อลดความแปรปรวนหรือความผิดปกติของข้อมูล และเพื่อให้สามารถนำข้อมูลไปใช้สร้างแบบจำลองพยากรณ์ได้อย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เหมาะสม หลังจากนั้นจะนำข้อมูลดังกล่าวไปแปลงค่าข้อมูลให้อยู่ในรูปแบบที่เหมาะสมกับอัลกอริทึมที่ใช้ภายในโครงข่าย

3.4.4 ขั้นตอนการการแบ่งข้อมูล

เป็นขั้นตอนในการแบ่งข้อมูลในแต่ละกลุ่มออกเป็น 2 ส่วน คือ ข้อมูลสำหรับใช้ในการสอนให้โครงข่ายเกิดการเรียนรู้ (Training Set) และข้อมูลสำหรับใช้ทดสอบความถูกต้องของแบบจำลองพยากรณ์ (Testing Set)

3.4.5 ขั้นตอนการสร้างโครงข่ายประสาทเทียม

เป็นขั้นตอนในการเลือกโครงสร้างที่เหมาะสมโดยการกำหนดชั้น (Layer) ภายในโครงข่าย และกำหนดจำนวนโหนดในแต่ละชั้นของโครงข่าย

3.4.6 ขั้นตอนการเลือกใช้อัลกอริทึม

เป็นขั้นตอนในการเลือกอัลกอริทึมเพื่อนำมาใช้ในการสอน ให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ ซึ่งในโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้เลือกใช้อัลกอริทึมแบล็คพร็อบพาเกชั่น

3.4.7 ขั้นตอนการสอนโครงข่าย

เป็นขั้นตอนในการสอนโครงข่ายให้เกิดการเรียนรู้ โดยนำข้อมูลสำหรับใช้ในการสอน (Training Set) มาใช้ในโครงข่ายตามอัลกอริทึมที่เลือกใช้ แล้วทำการปรับค่าถ่วงน้ำหนักที่เชื่อมอยู่ระหว่างโหนดให้มีความเหมาะสมมากขึ้น

3.4.8 ขั้นตอนการทดสอบแบบจำลองพยากรณ์ฯ

หลักจากที่โครงข่ายประสาทเทียมได้เรียนรู้ และสร้างแบบจำลองพยากรณ์ฯแล้วนั้น ขั้นตอนต่อไปจะนำข้อมูลสำหรับทดสอบ (Testing Set) มาใช้ทดสอบแบบจำลองพยากรณ์ฯดังกล่าว ซึ่งถ้าค่าข้อมูลที่ได้จากการทดสอบมีความผิดพลาดสูง ก็จะต้องกลับไปพิจารณาความเหมาะสมของข้อมูลที่ใช้ในการสอนโครงข่ายประสาทเทียม รวมถึงค่าตัวแปรอื่นๆที่ใช้ในการกำหนดให้โครงข่ายเกิดการเรียนรู้แล้วทำการปรับค่าเหล่านั้นให้เหมาะสม เพื่อทำให้แบบจำลองพยากรณ์ฯมีประสิทธิภาพ และมีความถูกต้องมากยิ่งขึ้น

3.4.9 ขั้นตอนการนำแบบจำลองพยากรณ์ฯไปใช้งาน

เป็นขั้นตอนสุดท้าย โดยเป็นการนำแบบจำลองพยากรณ์ฯจากโครงข่ายประสาทเทียมไปใช้งานจริงโดยการนำค่าข้อมูลอินพุตในช่วงที่เกิดขึ้นในปัจจุบันไปทดสอบ เพื่อพยากรณ์ราคาหลักทรัพย์ในอนาคตในช่วงเวลาถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การพัฒนาระบบพยากรณ์ราคาหลักทรัพย์

ในโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้ เป็นการนำเทคโนโลยีดาต้าไมนิ่งมาใช้ในลักษณะงานด้านการพยากรณ์ค่าข้อมูล โดยนำอัลกอริทึมแบล็คพร็อบพาเกชั่น มาใช้ในการสร้างแบบจำลองพยากรณ์ราคาหลักทรัพย์ ซึ่งจะแบ่งการทำงานของระบบออกเป็น 2 ระบบ คือ การสร้างแบบจำลองพยากรณ์ฯให้กับระบบ และการเรียกแบบจำลองพยากรณ์ฯ ที่สร้างไว้มาใช้ในงานเพื่อพยากรณ์ค่าของข้อมูล ซึ่งในบทนี้จะอธิบายถึงองค์ประกอบและการใช้งานระบบพยากรณ์ราคาหลักทรัพย์ที่สร้างขึ้น

4.1 องค์ประกอบของระบบพยากรณ์ราคาหลักทรัพย์

4.1.1 การสร้างแบบจำลองพยากรณ์ฯใหม่

4.1.1.1 ส่วนรับข้อมูล เป็นการเลือกไฟล์ข้อมูลที่จะนำมาใช้ในการสร้างและทดสอบแบบจำลองพยากรณ์ฯ ให้กับระบบโดยลักษณะของข้อมูลจะอยู่ในรูปของ Text File (มีสกุลเป็น *.txt) โดยที่แต่ละคอลัมน์ จะถูกแบ่งด้วยช่องว่าง และแต่ละชุดข้อมูลจะถูกแบ่งด้วยแถว

4.1.1.2 ส่วนประมวลผลข้อมูลเบื้องต้น เป็นการตรวจสอบข้อมูลที่จะนำมาใช้ในการสร้างและทดสอบแบบจำลองพยากรณ์ฯ ซึ่งสามารถแบ่งการทำงานได้เป็น 2 ขั้นตอนดังนี้

ขั้นตอนที่ 1 การจัดการกับข้อมูลที่ขาดหายไป (Missing Value)

สำหรับข้อมูลที่นำมาใช้ในโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้ ได้รับการสนับสนุนจากหลายแหล่งข้อมูล ดังนี้

- สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์แห่งประเทศไทย
- ตลาดหลักทรัพย์แห่งประเทศไทย
- ธนาคารแห่งประเทศไทย
- บริษัท ปตท. จำกัด(มหาชน)

ซึ่งลักษณะข้อมูลที่ได้รับดังกล่าวมีความสมบูรณ์และมีคุณภาพสูง เนื่องจากหน้าที่หลักหรือธุรกิจหลักของหน่วยงานดังกล่าวจำเป็นต้องรักษาข้อมูลนั้นๆให้มีเสถียรภาพ ถูกต้องแม่นยำและต้องปรับปรุงข้อมูลให้ทันสมัยอยู่ตลอดเวลา จะมีแต่เพียงข้อมูลราคาปิดของหลักทรัพย์ ณ บางช่วง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวลาที่ขาดหายไป เนื่องจาก ณ เวลาดังกล่าวบริษัทเจ้าของหลักทรัพย์ถูกตลาดหลักทรัพย์สั่งห้ามทำการซื้อขาย แต่ปัญหาดังกล่าวสามารถจัดการได้โดยการแทนค่าข้อมูล ณ ช่วงเวลาที่ขาดหายไปโดยการใช้ราคาปิดครั้งก่อนหน้า เพราะถือว่าปัญหาดังกล่าวไม่มีผลกระทบต่อการเปลี่ยนแปลงของราคาหลักทรัพย์ ณ ช่วงเวลานั้น เนื่องจากเมื่อตลาดหลักทรัพย์อนุญาตให้ทำการซื้อขายตามปกติแล้วราคาซื้อขายหลักทรัพย์นั้นๆก็จะใช้ราคาปิดก่อนหน้านั้นได้ทันที

ขั้นตอนที่ 2 การแปลงค่าข้อมูล (Transformation)

สำหรับขั้นตอนนี้เป็นการนำข้อมูลที่ได้จากขั้นตอนที่ 1 มาแปลงค่าให้ข้อมูลมีค่าอยู่ในช่วงระหว่าง 0 ถึง 1 โดยเลือกใช้วิธีการแปลงค่าข้อมูลแบบเชิงเส้นตามที่ได้กล่าวไว้ในบทที่ 2 เพื่อให้ข้อมูลสามารถนำไปใช้งานกับอัลกอริทึมแบล็คพร็อบพาเกชั่นได้ และให้สอดคล้องกับฟังก์ชันกระตุ้นที่นำมาใช้ในแบบจำลองพยากรณ์ฯ สำหรับในระบบนี้จะใช้ฟังก์ชันของซิกมอยด์ (Sigmoid Function)

4.1.1.3 ส่วนการสร้างแบบจำลองพยากรณ์ฯ สำหรับโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษฉบับนี้จะใช้อัลกอริทึมแบล็คพร็อบพาเกชั่น มาใช้ในการสอนให้ระบบเกิดการเรียนรู้และสร้างแบบจำลองพยากรณ์ฯขึ้น โดยผู้ใช้สามารถกำหนดตัวแปรบางตัวที่ใช้ควบคุมการเรียนรู้ได้ เช่น จำนวนโหนดในชั้นอินพุต (Input Node) จำนวนรอบของการเรียนรู้ อัตราการเรียนรู้ (Learning Rate) และค่าความผิดพลาดที่ยอมรับได้ แต่เนื่องจากระบบนี้ใช้ซอฟแวร์สำเร็จรูปในการประมวลผลข้อมูลทำให้ตัวแปรบางตัวถูกจำกัดเนื่องจากเงื่อนไขของเจ้าของซอฟแวร์เอง คือ จำนวนชุดข้อมูลสำหรับใช้ในการเรียนรู้และทดสอบไม่ให้เกินกว่า 500 ชุด จำนวนโหนดชั้นกลาง (Hidden Node) ไม่ให้เกินกว่า 1 ชั้นและใน 1 ชั้นให้มีจำนวนโหนดไม่เกินกว่า 10 โหนด ด้วยเงื่อนไขดังกล่าว ในระบบงานนี้ก็ยังสามารถสร้างแบบจำลองพยากรณ์ฯได้โดยประกอบด้วยขั้นตอน ดังนี้

ขั้นตอนที่ 1 การสอนให้ระบบเกิดการเรียนรู้ (Training Phase) เป็นการนำข้อมูลแต่ละชุดที่ใช้สำหรับสอนระบบมาทำการคำนวณตามอัลกอริทึมแบล็คพร็อบพาเกชั่น เพื่อปรับตัวแปรภายในโครงข่ายจนกระทั่งได้แบบจำลองพยากรณ์ฯที่เหมาะสม โดยระบบจะใช้ค่าความผิดพลาด (Mean Square Error) ตามที่ผู้ใช้กำหนดไว้เป็นเงื่อนไขในการจบการเรียนรู้ของแบบจำลองพยากรณ์ โดยที่จำนวนรอบในการเรียนรู้ยังไม่เกินกว่าค่ากำหนดไว้ หรืออาจกำหนดเงื่อนไขการจบการเรียนรู้โดยพิจารณาจากจำนวนรอบของการเรียนรู้ก็ได้

ขั้นตอนที่ 2 การทดสอบแบบจำลองพยากรณ์ฯที่ได้จากการสอน (Testing Phase) เป็นการประเมินแบบจำลองพยากรณ์ฯโดยใช้ข้อมูลส่วนที่เก็บไว้สำหรับทดสอบ (Test Set) โดยในการทำงานของขั้นตอนนี้จะเป็นการนำข้อมูลเข้าสู่แบบจำลองพยากรณ์ฯที่ได้จากการสอน และทำการคำนวณหาค่าเอาต์พุตของโครงข่ายเพื่อเปรียบเทียบกับค่าข้อมูลที่เกิดขึ้นจริง แล้วหาค่าความผิดพลาดเพื่อประเมิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประสิทธิภาพของระบบให้เป็นที่น่าพอใจและทำการบันทึกแบบจำลองพยากรณ์ฯไว้เพื่อนำไปใช้ในการพยากรณ์ในครั้งต่อๆไป

4.1.2 การใช้งานแบบจำลองพยากรณ์ฯ เป็นการเรียกแบบจำลองพยากรณ์ฯที่สร้างไว้แล้วมาใช้งาน โดยจะมีขั้นตอนหลักๆ ดังต่อไปนี้

4.1.2.1 การเลือกแบบจำลองพยากรณ์ฯ โดยเลือกไฟล์ที่เก็บไว้ในรูปแบบ QwikNet Network File ซึ่งเป็นไฟล์ที่ใช้กับซอฟแวร์สำเร็จรูป โดยคำนึงถึงความเหมาะสมกับข้อมูลที่จะนำมาทดสอบ

4.1.2.2 การนำข้อมูลเข้าระบบ เป็นส่วนที่รับข้อมูลอินพุตจากผู้ใช้ตามจำนวนของข้อมูลในชั้นอินพุต ที่ใช้สร้างแบบจำลองพยากรณ์ฯให้กับระบบ

4.1.2.3 การประมวลผล เป็นการคำนวณหาค่าเอาต์พุตโดยใช้ค่าตัวแปรของแบบจำลองพยากรณ์ฯ คำนวณร่วมกับค่าอินพุตที่ป้อนจากผู้ใช้งานตามลำดับชั้นของโครงสร้างเครือข่ายภายในระบบ

4.2 การใช้งานระบบพยากรณ์ราคาหลักทรัพย์

แสดงถึงขั้นตอนการใช้งานระบบพยากรณ์ราคาหลักทรัพย์ที่สร้างขึ้น สามารถแสดงได้ตามตัวอย่างดังรูปที่ 4.1 ถึง 4.13 ซึ่งมีวิธีการใช้งานสามารถสรุปได้ดังต่อไปนี้

4.2.1 ขั้นตอนการสร้างแบบจำลองพยากรณ์ฯใหม่

4.2.1.1 การนำเข้าข้อมูล ที่ใช้ในการสร้างแบบจำลองพยากรณ์ฯ โดยลักษณะของข้อมูลจะอยู่ในรูปของ Text File (มีสกุลเป็น *.txt) ตามที่กล่าวไปแล้ว

4.2.1.2 การกำหนดจำนวนอินพุตและเอาต์พุต ให้กับระบบพยากรณ์ราคาหลักทรัพย์

4.2.1.3 การกำหนดค่าตัวแปร เพื่อใช้ในการสร้างแบบจำลองพยากรณ์ฯ สามารถแบ่งออกเป็น 3 ส่วน ดังนี้

- ส่วนสำหรับการเรียนรู้ มีตัวแปรคือ อัตราการเรียนรู้ (Learning Rate)
- ส่วนสำหรับหยุดการเรียนรู้ มีตัวแปรคือ จำนวนรอบในการเรียนรู้ (Epochs) ค่าความผิดพลาดที่ยอมรับได้ (Root Mean Square Error)
- ส่วนสำหรับสร้างโครงข่าย มีตัวแปรคือ จำนวนชั้นและจำนวนโหนดชั้นกลาง (Hidden Node)

4.2.2.4 การสร้างแบบจำลองพยากรณ์ฯ โดยสอนให้ระบบเรียนรู้จากข้อมูลและตัวแปรต่างๆ ดังกล่าว

4.2.2.5 การทดสอบแบบจำลองพยากรณ์ฯ โดยนำข้อมูลสำหรับทดสอบ (Test Set) มาทดสอบแบบจำลองพยากรณ์ฯที่สร้างได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2.6 การวิเคราะห์ผลลัพธ์และบันทึกแบบจำลองพยากรณ์ฯ รวมถึงการบันทึกค่าถ่วงน้ำหนัก ที่ได้จากการสร้างแบบจำลองพยากรณ์ฯดังกล่าว เพื่อสามารถนำไปใช้งานต่อไป

4.2.2 ขั้นตอนการใช้งานแบบจำลองพยากรณ์ฯ เป็นขั้นตอนในการเรียกใช้แบบจำลองพยากรณ์ฯ ที่ได้สร้างไว้ ประกอบไปด้วยขั้นตอน ดังนี้

4.2.2.1 การเลือกแบบจำลองพยากรณ์ โดยคำนึงถึงข้อมูลที่จะนำมาทดสอบว่าตรงตามแบบจำลองพยากรณ์ฯที่สร้างไว้หรือไม่

4.2.2.2 การกำหนดจำนวนอินพุตและเอาต์พุต ให้กับข้อมูลดังกล่าว

4.2.2.3 การเรียกใช้ค่าถ่วงน้ำหนัก จากระบบพยากรณ์ราคาหลักทรัพย์

4.2.2.4 การคำนวณและหาผลลัพธ์ จากระบบพยากรณ์ราคาหลักทรัพย์

Date	Exche	Repu	FTSE10	DOW_J	HANG_	NIKKEI	KUALA	SINGA	TAIWA	PHILIPF	JAKAR	KOREA	NASDA	Desel	Bensi	Bensi	SET_In	SET5(	M_VALI	M_VALI	M_VALI	M_VALI	M_VALI	M_VALI	KBANK
1/4/2547	39.17	0.97	410.71	373.33	676.25	683.42	887.08	870.70	523.49	433.11	737.65	882.75	015.01	14.59	16.19	16.99	671.92	45.13	28,264	06,108	42,877	49,777	12,963	28,219	51.50
2/4/2547	39.14	0.97	465.61	470.60	731.76	815.95	887.34	880.45	545.54	450.91	750.65	883.69	057.17	14.59	16.19	16.99	693.12	46.45	42,255	69,735	50,763	33,707	64,668	54,248	55.50
5/4/2547	39.08	1.00	480.70	558.40	731.76	958.32	889.60	890.20	682.73	490.05	750.65	883.69	079.12	14.59	16.19	16.99	709.89	47.62	71,391	96,216	31,900	61,116	24,806	70,766	54.00
7/4/2547	39.10	1.00	468.70	480.10	920.05	019.62	882.34	884.43	646.74	516.81	774.40	909.93	050.24	14.59	16.19	16.99	698.82	46.69	78,195	31,170	26,453	09,077	29,987	94,388	53.50
8/4/2547	39.04	0.97	489.67	442.03	909.37	092.59	884.48	901.92	672.86	516.81	779.62	916.86	052.88	14.59	16.19	16.99	691.89	46.20	09,964	53,110	03,374	01,690	90,904	49,443	52.00
9/4/2547	39.11	0.97	489.67	442.03	909.37	897.51	883.09	901.92	520.36	516.81	779.62	905.44	052.88	14.59	16.19	16.99	691.39	46.22	70,861	62,141	66,750	45,883	16,869	26,455	53.00
12/4/2547	39.12	0.94	489.67	515.56	909.37	042.70	879.78	898.27	777.78	510.58	767.81	918.86	065.48	14.59	16.19	16.99	701.72	46.92	61,061	07,390	25,186	44,322	15,305	49,860	54.00
16/4/2547	39.43	0.94	537.28	451.97	458.38	624.56	861.81	853.99	816.20	498.11	776.57	698.68	995.74	14.59	16.19	16.99	712.20	47.83	30,861	99,860	19,628	32,499	16,893	35,023	54.50
19/4/2547	39.28	0.94	546.22	437.80	450.00	764.21	855.76	844.71	779.18	497.48	764.84	902.10	020.43	14.59	16.19	16.99	704.65	47.38	84,424	02,214	96,110	14,664	20,858	84,514	55.00
20/4/2547	39.33	0.97	569.02	314.50	394.37	952.26	861.33	850.42	799.97	525.96	810.86	918.90	978.63	14.59	16.19	16.99	713.95	48.10	96,747	88,757	98,958	98,466	68,412	76,893	55.00
21/4/2547	39.44	0.97	539.87	317.30	227.30	944.30	864.02	848.41	810.25	537.78	814.20	929.95	995.63	14.59	16.19	16.99	706.65	47.64	24,048	82,362	20,435	24,995	36,995	74,122	54.50
22/4/2547	39.56	0.97	571.83	461.20	167.70	980.10	871.63	863.71	732.09	548.28	804.43	924.01	032.91	14.59	16.19	16.99	690.96	46.65	51,868	93,615	42,826	50,859	90,639	40,860	54.50
23/4/2547	39.56	0.97	569.95	472.80	383.94	120.66	875.50	867.64	748.10	589.40	815.44	936.06	049.77	14.59	16.19	16.99	681.88	46.05	58,820	70,409	78,272	40,488	72,414	98,608	53.00
26/4/2547	39.54	1.00	571.85	444.70	132.68	163.89	870.12	849.38	710.70	571.90	811.75	919.74	036.77	14.59	16.19	16.99	667.61	45.02	85,115	72,939	99,430	58,034	50,202	03,774	50.00
27/4/2547	39.58	0.97	575.68	478.20	154.91	044.88	868.32	853.58	646.80	596.93	818.16	915.47	032.53	14.59	16.19	16.99	680.89	46.02	06,881	68,940	93,059	41,750	19,310	08,559	50.50
28/4/2547	39.70	0.97	524.50	342.60	165.31	004.29	861.57	847.76	574.75	610.31	817.93	901.83	989.54	14.59	16.19	16.99	672.34	45.38	02,524	12,932	50,392	71,404	22,360	90,940	50.50
29/4/2547	39.87	0.94	519.53	272.30	005.58	004.29	849.72	812.13	402.21	581.95	801.97	875.41	958.78	14.59	16.19	16.99	656.36	44.30	34,732	69,451	73,427	31,598	33,650	40,759	49.00
30/4/2547	39.95	0.94	489.69	225.60	942.96	761.79	838.21	842.03	117.81	555.01	783.41	862.84	920.15	14.59	16.19	16.99	648.15	43.69	63,995	96,982	13,905	99,566	76,222	57,575	48.50
4/5/2547	39.90	0.97	547.23	317.20	098.30	761.79	838.21	865.35	188.15	573.83	779.60	867.48	950.48	14.59	16.19	16.99	644.10	43.40	88,045	82,913	62,027	96,953	54,446	24,651	48.75
6/5/2547	39.74	0.97	516.17	241.30	010.31	571.34	842.77	866.61	909.79	551.69	746.12	837.68	937.74	14.59	16.19	16.99	634.01	42.84	23,172	29,057	09,910	44,777	86,979	46,227	49.00
7/5/2547	39.92	0.97	498.37	117.30	910.76	438.82	835.37	843.72	040.26	554.69	743.64	838.74	917.96	14.59	16.79	17.59	636.80	42.97	74,946	61,349	24,635	53,860	32,592	16,964	48.00
10/5/2547	40.34	0.97	395.16	990.02	485.50	884.70	813.44	791.78	825.05	554.69	707.22	790.68	896.07	14.59	16.79	17.59	605.62	40.81	35,972	05,474	17,888	21,737	00,523	27,172	45.25
11/5/2547	40.55	0.94	454.70	019.50	508.10	907.18	810.16	803.39	886.36	508.26	718.26	791.02	931.35	14.59	16.79	17.59	618.10	41.69	87,869	00,961	70,625	67,180	69,072	59,424	46.00
12/5/2547	40.46	0.94	412.93	045.20	528.18	153.58	818.72	802.17	958.79	540.96	744.29	817.09	925.59	14.59	16.79	17.59	622.01	41.97	32,086	74,899	25,685	71,820	48,647	59,699	46.25
13/5/2547	40.60	0.94	453.81	010.70	396.94	825.10	803.11	778.59	918.09	526.47	739.14	790.13	926.03	14.59	16.79	17.59	611.23	41.28	85,311	12,992	34,026	10,574	25,341	21,113	45.25
14/5/2547	40.78	0.94	441.79	012.90	276.86	849.63	793.97	754.96	777.32	521.61	722.71	768.46	904.25	14.59	16.79	17.59	609.72	41.14	58,757	63,305	21,311	01,109	37,493	53,146	45.50
17/5/2547	40.76	0.97	403.02	906.91	967.65	505.05	781.05	700.33	482.96	486.99	668.48	728.98	876.64	14.59	16.79	17.59	581.61	39.39	43,824	41,340	97,126	95,882	91,064	94,791	43.25
18/5/2547	40.83	1.00	414.41	968.51	072.39	711.09	783.27	722.80	557.68	476.08	676.15	741.99	897.82	14.59	16.79	17.59	582.51	39.49	21,331	77,488	02,225	41,475	12,319	16,911	43.25
19/5/2547	40.66	1.00	471.80	937.71	469.41	967.74	793.83	765.26	860.58	468.32	706.80	777.95	898.17	14.59	16.79	17.59	614.99	41.80	89,555	35,809	68,125	68,319	46,041	99,592	46.50
20/5/2547	40.67	1.00	428.71	937.64	339.62	862.04	788.87	741.57	815.33	472.25	706.80	767.79	896.59	14.59	16.79	17.59	599.88	40.74	72,302	31,377	70,668	24,011	18,816	06,398	46.00
21/5/2547	40.76	1.00	431.43	966.74	576.01	070.25	791.53	767.23	964.94	484.96	724.93	786.36	912.09	14.59	16.79	17.59	615.41	41.74	39,634	10,239	76,694	15,343	94,250	84,996	47.50
24/5/2547	40.69	0.97	428.87	958.43	662.97	101.64	794.39	779.79	942.08	481.91	732.58	799.64	922.98	14.59	16.79	17.59	608.90	41.21	15,537	15,904	09,558	74,293	80,481	15,380	46.50
25/5/2547	40.76	1.00	418.00	117.60	692.56	962.93	796.48	766.42	958.38	493.90	717.14	784.06	964.65	14.59	16.79	17.59	601.51	40.73	63,209	67,188	76,128	42,682	80,615	10,082	45.75
26/5/2547	40.69	1.00	438.29	109.90	692.56	152.09	799.25	776.70	027.27	507.65	718.01	784.06	976.15	14.59	16.79	17.59	609.60	41.40	36,597	48,496	53,707	98,574	48,965	92,199	46.75
27/5/2547	40.68	1.00	453.62	205.20	983.90	166.03	811.74	778.55	033.05	499.12	728.31	802.46	984.50	14.59	16.79	17.59	630.72	42.87	24,478	20,967	43,584	52,927	83,376	77,545	47.75
28/5/2547	40.68	1.00	430.69	188.40	116.87	309.57	808.44	787.18	137.26	517.20	733.99	816.51	986.74	14.59	16.79	17.59	638.59	43.41	06,036	41,364	94,737	77,593	57,900	39,696	48.25
31/5/2547	40.68	0.97	430.69	188.45	198.24	236.37	810.67	788.66	977.84	511.36	732.52	803.84	986.74	14.59	16.79	17.59	641.05	43.57	24,439	27,849	44,658	66,418	04,012	78,842	48.25

รูปที่ 4.1 แสดงตัวอย่างข้อมูลอินพุตเพื่อใช้ในการสร้างแบบจำลองพยากรณ์ฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

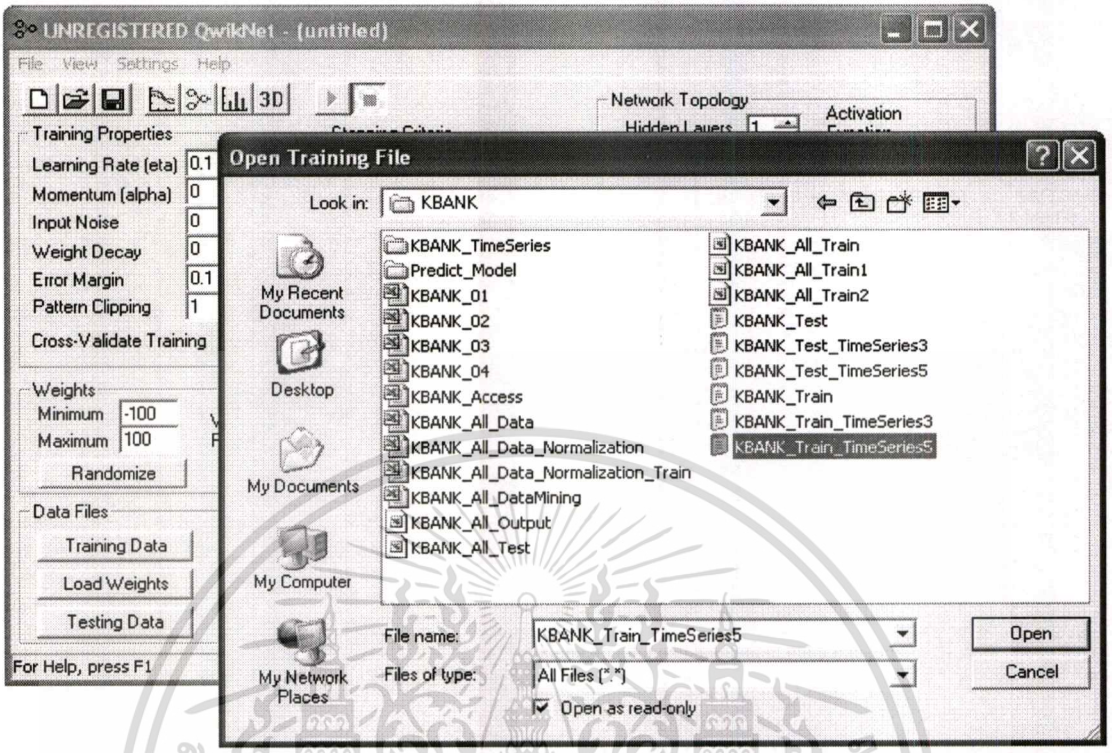
Date	Exchar	FTSE1(	DOW_.	HANG_	NIKKEI	KUALA	SINGAI	TAIWA	PHILIPF	JAKAR	KOREA	NASDA	Desel	Bensil$	Repurc	SET_In	SET50	M_VAL	M_VAL	M_VAL	M_VAL	M_VAL	M_VAL	KBANK
1/4/2547	0.0704	0.3687	0.7619	0.5503	0.5888	0.9385	0.8449	0.8577	0.6193	0.8325	0.8830	0.5163	0.8780	0.7778	0.0484	0.7691	0.7965	0.3042	0.3314	0.4274	0.4668	0.5553	0.4853	0.5188
2/4/2547	0.0673	0.3867	0.7859	0.5574	0.6080	0.9392	0.8574	0.8638	0.6436	0.8596	0.8850	0.5404	0.8780	0.7778	0.0484	0.8092	0.8339	0.4728	0.5591	0.6106	0.7449	1.0000	0.6714	0.5688
5/4/2547	0.0585	0.3917	0.8076	0.5574	0.6285	0.9456	0.8700	0.9021	0.6970	0.8596	0.8850	0.5530	0.8780	0.7778	0.0645	0.8409	0.8671	0.4558	0.5452	0.4306	0.6673	0.9046	0.4927	0.5500
7/4/2547	0.0606	0.3877	0.7883	0.5817	0.6374	0.9251	0.8626	0.8920	0.7335	0.9090	0.9410	0.5365	0.8780	0.7778	0.0645	0.8200	0.8407	0.4542	0.4140	0.3380	0.7944	0.6970	0.7638	0.5437
8/4/2547	0.0521	0.3946	0.7789	0.5803	0.6479	0.9312	0.8651	0.8993	0.7335	0.9198	0.9557	0.5380	0.8780	0.7778	0.0484	0.8065	0.8268	0.3434	0.3361	0.2998	0.4983	0.3787	0.3688	0.5250
9/4/2547	0.0622	0.3946	0.7789	0.5803	0.6198	0.9273	0.8651	0.8847	0.7335	0.9198	0.9314	0.5380	0.8780	0.7778	0.0484	0.8059	0.8274	0.2133	0.2096	0.1713	0.2187	0.0679	0.0660	0.5375
12/4/2547	0.0635	0.3946	0.7970	0.5803	0.6407	0.9179	0.8804	0.9286	0.7250	0.8953	0.9600	0.5452	0.8780	0.7778	0.0323	0.8255	0.8472	0.1805	0.1716	0.2006	0.2323	0.0946	0.1319	0.5500
16/4/2547	0.1070	0.4103	0.7813	0.5222	0.6092	0.8674	0.8234	0.9398	0.7080	0.9135	0.9174	0.5052	0.8780	0.7778	0.0323	0.8453	0.8730	0.3069	0.2617	0.2496	0.3526	0.5192	0.6398	0.5562
19/4/2547	0.0866	0.4132	0.7778	0.5211	0.6005	0.8504	0.8115	0.9289	0.7071	0.9307	0.9243	0.5194	0.8780	0.7778	0.0323	0.8310	0.8603	0.2833	0.2952	0.2577	0.3381	0.4175	0.3698	0.5625
20/4/2547	0.0937	0.4207	0.7474	0.5139	0.6277	0.8661	0.8188	0.9347	0.7460	0.9848	0.9601	0.4954	0.8780	0.7778	0.0484	0.8486	0.8807	0.2808	0.2772	0.2977	0.3245	0.2429	0.2748	0.5625
21/4/2547	0.1095	0.4111	0.7481	0.4924	0.6265	0.8736	0.8162	0.9376	0.7621	0.9918	0.9837	0.5052	0.8780	0.7778	0.0484	0.8348	0.8676	0.3758	0.3697	0.3687	0.4048	0.3209	0.3679	0.5562
22/4/2547	0.1260	0.4216	0.7836	0.4847	0.6317	0.8950	0.8359	0.9158	0.7765	0.9714	0.9710	0.5265	0.8780	0.7778	0.0484	0.8051	0.8396	0.3151	0.2987	0.2514	0.3835	0.3301	0.3797	0.5562
23/4/2547	0.1256	0.4210	0.7865	0.5126	0.6520	0.9059	0.8410	0.9203	0.8326	0.9944	0.9967	0.5362	0.8780	0.7778	0.0484	0.7879	0.8226	0.2648	0.2257	0.3405	0.4916	0.2434	0.3971	0.5375
26/4/2547	0.1236	0.4216	0.7795	0.4802	0.6582	0.8908	0.8175	0.9099	0.8087	0.9867	0.9619	0.5287	0.8780	0.7778	0.0645	0.7610	0.7934	0.2000	0.1586	0.2566	0.2157	0.2570	0.4930	0.5000
27/4/2547	0.1281	0.4229	0.7878	0.4830	0.6410	0.8857	0.8229	0.8921	0.8429	1.0000	0.9528	0.5263	0.8780	0.7778	0.0484	0.7861	0.8217	0.2140	0.1925	0.2464	0.1677	0.2148	0.3670	0.5063
28/4/2547	0.1454	0.4061	0.7543	0.4844	0.6352	0.8667	0.8154	0.8720	0.8611	0.9995	0.9237	0.5017	0.8780	0.7778	0.0484	0.7699	0.8036	0.1989	0.2018	0.2298	0.1910	0.2097	0.2334	0.5063
29/4/2547	0.1702	0.4044	0.7370	0.4638	0.6352	0.8334	0.7696	0.8239	0.8224	0.9863	0.8673	0.4840	0.8780	0.7778	0.0323	0.7397	0.7730	0.2477	0.1654	0.2685	0.3176	0.3044	0.7050	0.4875
30/4/2547	0.1816	0.3946	0.7255	0.4557	0.6001	0.8011	0.8080	0.7446	0.7857	0.9277	0.8405	0.4619	0.8780	0.7778	0.0323	0.7242	0.7557	0.2440	0.1755	0.5416	0.3787	0.3321	0.7770	0.4812
4/5/2547	0.1741	0.4135	0.7481	0.4757	0.6001	0.8011	0.8380	0.7642	0.8113	0.9198	0.8504	0.4793	0.8780	0.7778	0.0484	0.7165	0.7474	0.1163	0.0936	0.1980	0.1217	0.2989	0.4678	0.4844
6/5/2547	0.1517	0.4033	0.7293	0.4644	0.5726	0.8139	0.8396	0.6866	0.7811	0.8501	0.7869	0.4720	0.8780	0.7778	0.0484	0.6974	0.7316	0.1388	0.1231	0.3696	0.2100	0.4358	0.6367	0.4875
7/5/2547	0.1769	0.3975	0.6987	0.4515	0.5535	0.7931	0.8102	0.7230	0.7852	0.8450	0.7891	0.4606	0.8780	0.8889	0.0484	0.7027	0.7353	0.2268	0.1615	0.3660	0.2009	0.3267	0.7481	0.4750
10/5/2547	0.2367	0.3636	0.6673	0.3967	0.4734	0.7314	0.7434	0.6630	0.7852	0.7692	0.6866	0.4481	0.8780	0.8889	0.0484	0.6437	0.6740	0.2431	0.1768	0.7107	0.5278	0.2846	0.7431	0.4406
11/5/2547	0.2655	0.3832	0.6746	0.3996	0.4767	0.7222	0.7583	0.6801	0.7219	0.7922	0.6873	0.4683	0.8780	0.8889	0.0323	0.6673	0.6990	0.2624	0.2381	1.0000	0.4600	0.3104	0.7496	0.4500
12/5/2547	0.2534	0.3694	0.6809	0.4022	0.5123	0.7462	0.7567	0.7003	0.7665	0.8463	0.7429	0.4650	0.8780	0.8889	0.0323	0.6747	0.7069	0.2401	0.2531	0.3751	0.3625	0.3364	0.3297	0.4531
13/5/2547	0.2734	0.3829	0.6724	0.3853	0.4648	0.7024	0.7264	0.6889	0.7467	0.8356	0.6854	0.4653	0.8780	0.8889	0.0323	0.6543	0.6874	0.1472	0.1235	0.1929	0.1724	0.2418	0.3842	0.4406
14/5/2547	0.2986	0.3789	0.6730	0.3698	0.4684	0.6766	0.6960	0.6497	0.7401	0.8014	0.6392	0.4528	0.8780	0.8889	0.0323	0.6515	0.6834	0.1084	0.0871	0.2421	0.1724	0.1767	0.3305	0.4437
17/5/2547	0.2957	0.3662	0.6468	0.3299	0.4186	0.6403	0.6258	0.5677	0.6928	0.6866	0.5550	0.4370	0.8780	0.8889	0.0484	0.5983	0.6338	0.1624	0.1378	0.3991	0.3227	0.2153	0.4944	0.4156
18/5/2547	0.3049	0.3699	0.6620	0.3434	0.4483	0.6466	0.6547	0.5885	0.6779	0.7046	0.5828	0.4491	0.8780	0.8889	0.0645	0.6000	0.6366	0.2238	0.2005	0.3875	0.4209	0.2504	0.3871	0.4156
19/5/2547	0.2809	0.3888	0.6544	0.3946	0.4854	0.6763	0.7093	0.6729	0.6673	0.7683	0.6595	0.4493	0.8780	0.8889	0.0645	0.6615	0.7021	0.3193	0.3400	0.4715	0.3360	0.4406	0.4605	0.4563
20/5/2547	0.2626	0.3746	0.6544	0.3779	0.4702	0.6623	0.6788	0.6603	0.6727	0.7683	0.6378	0.4484	0.8780	0.8889	0.0645	0.6329	0.6721	0.2426	0.2535	0.3525	0.3193	0.1988	0.1928	0.4500
21/5/2547	0.2947	0.3755	0.6616	0.4084	0.5002	0.6698	0.7118	0.7020	0.6901	0.8061	0.6774	0.4573	0.8780	0.8889	0.0645	0.6622	0.7004	0.3145	0.3197	0.2773	0.2646	0.2696	0.2731	0.4688
24/5/2547	0.2847	0.3747	0.6595	0.4196	0.5048	0.6778	0.7280	0.6856	0.6859	0.8220	0.7057	0.4635	0.8780	0.8889	0.0484	0.6499	0.6854	0.2661	0.2655	0.1978	0.2844	0.1711	0.1483	0.4563
25/5/2547	0.2957	0.3711	0.6988	0.4234	0.4847	0.6837	0.7108	0.7002	0.7023	0.7898	0.6725	0.4874	0.8780	0.8889	0.0645	0.6360	0.6718	0.1740	0.1717	0.1579	0.1572	0.1335	0.1530	0.4469
26/5/2547	0.2857	0.3778	0.6969	0.4234	0.5121	0.6915	0.7240	0.7194	0.7210	0.7916	0.6725	0.4940	0.8780	0.8889	0.0645	0.6513	0.6906	0.1819	0.1990	0.1205	0.2641	0.2321	0.1108	0.4594
27/5/2547	0.2845	0.3828	0.7204	0.4610	0.5141	0.7266	0.7264	0.7210	0.7094	0.8131	0.7117	0.4988	0.8780	0.8889	0.0645	0.6912	0.7324	0.3700	0.3952	0.3095	0.5017	0.4737	0.3247	0.4719
28/5/2547	0.2845	0.3753	0.7163	0.4781	0.5348	0.7173	0.7375	0.7500	0.7341	0.8249	0.7417	0.5001	0.8780	0.8889	0.0645	0.7081	0.7477	0.4812	0.5476	0.4773	0.5427	0.5809	0.3121	0.4781
31/5/2547	0.2845	0.3753	0.7163	0.4886	0.5242	0.7236	0.7394	0.7056	0.7261	0.8218	0.7147	0.5001	0.8780	0.8889	0.0484	0.7107	0.7523	0.2855	0.2664	0.1954	0.4207	0.1671	0.1728	0.4781

รูปที่ 4.2 แสดงตัวอย่างข้อมูลอินพุตที่ผ่านการแปลงค่าข้อมูล

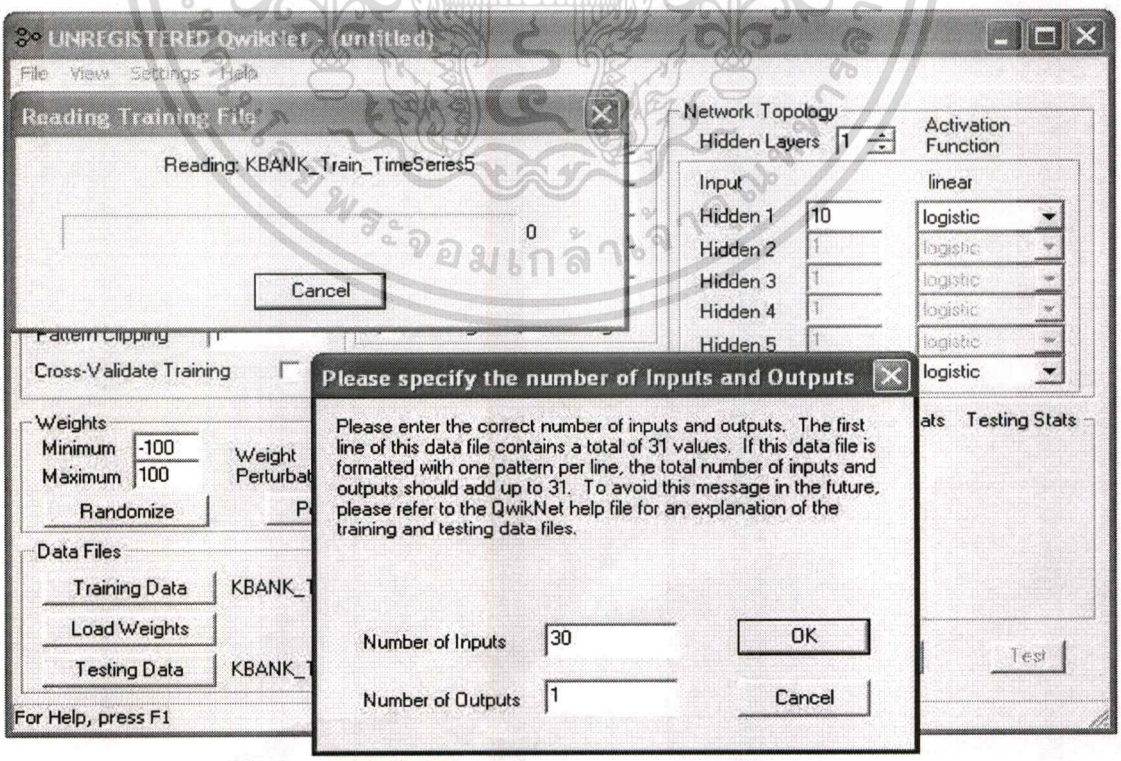
วันที่ t-10	วันที่ t-9	วันที่ t-8	วันที่ t-7	วันที่ t-6	วันที่ t-5	วันที่ t-4	วันที่ t-3	วันที่ t-2	วันที่ t-1	วันที่ t
0.1906	0.1875	0.2156	0.2062	0.1737	0.1906	0.1875	0.2156	0.2062	0.2094	0.2094
0.1875	0.2156	0.2062	0.1737	0.1906	0.1875	0.2156	0.2062	0.2094	0.2094	0.2000
0.2156	0.2062	0.1737	0.1906	0.1875	0.2156	0.2062	0.2094	0.2094	0.2000	0.1969
0.2062	0.1737	0.1906	0.1875	0.2156	0.2062	0.2094	0.2094	0.2000	0.1969	0.1838
0.1737	0.1906	0.1875	0.2156	0.2062	0.2094	0.2094	0.2000	0.1969	0.1838	0.1875
0.1906	0.1875	0.2156	0.2062	0.2094	0.2094	0.2000	0.1969	0.1838	0.1875	0.1825
0.1875	0.2156	0.2062	0.2094	0.2094	0.2000	0.1969	0.1838	0.1875	0.1825	0.1937
0.2156	0.2062	0.2094	0.2094	0.2000	0.1969	0.1838	0.1875	0.1825	0.1937	0.1875
0.2062	0.2094	0.2094	0.2000	0.1969	0.1838	0.1875	0.1825	0.1937	0.1875	0.2000
0.2094	0.2094	0.2000	0.1969	0.1838	0.1875	0.1825	0.1937	0.1875	0.2000	0.2031
0.2094	0.2000	0.1969	0.1838	0.1875	0.1825	0.1937	0.1875	0.2000	0.2031	0.2062
0.2000	0.1969	0.1838	0.1875	0.1825	0.1937	0.1875	0.2000	0.2031	0.2062	0.1969
0.1969	0.1838	0.1875	0.1825	0.1937	0.1875	0.2000	0.2031	0.2062	0.1969	0.2000
0.1838	0.1875	0.1825	0.1937	0.1875	0.2000	0.2031	0.2062	0.1969	0.2000	0.2000
0.1875	0.1825	0.1937	0.1875	0.2000	0.2031	0.2062	0.1969	0.2000	0.2000	0.2000
0.1825	0.1937	0.1875	0.2000	0.2031	0.2062	0.1969	0.2000	0.2000	0.2000	0.1937
0.1937	0.1875	0.2000	0.2031	0.2062	0.1969	0.2000	0.2000	0.2000	0.1937	0.1937
0.1875	0.2000	0.2031	0.2062	0.1969	0.2000	0.2000	0.2000	0.1937	0.1937	0.1937
0.2000	0.2031	0.2062	0.1969	0.2000	0.2000	0.2000	0.1937	0.1937	0.1937	0.1969
0.2031	0.2062	0.1969	0.2000	0.2000	0.2000	0.1937	0.1937	0.1937	0.1969	0.2000
0.2062	0.1969	0.2000	0.2000	0.2000	0.1937	0.1937	0.1937	0.1969	0.2000	0.2000
0.1969	0.2000	0.2000	0.2000	0.1937	0.1937	0.1937	0.1969	0.2000	0.2000	0.2125
0.2000	0.2000	0.2000	0.1937	0.1937	0.1937	0.1969	0.2000	0.2000	0.2125	0.2094
0.2000	0.2000	0.1937	0.1937	0.1937	0.1969	0.2000	0.2000	0.2125	0.2094	0.2375
0.2000	0.1937	0.1937	0.1937	0.1969	0.2000	0.2000	0.2125	0.2094	0.2375	0.2531
0.1937	0.1937	0.1937	0.1969	0.2000	0.2000	0.2125	0.2094	0.2375	0.2531	0.2469
0.1937	0.1937	0.1969	0.2000	0.2000	0.2125	0.2094	0.2375	0.2531	0.2469	0.2500
0.1937	0.1969	0.2000	0.2000	0.2125	0.2094	0.2375	0.2531	0.2469	0.2500	0.2500
0.1969	0.2000	0.2000	0.2125	0.2094	0.2375	0.2531	0.2469	0.2500	0.2500	0.2469
0.2000	0.2000	0.2125	0.2094	0.2375	0.2531	0.2469	0.2500	0.2500	0.2469	0.2500
0.2000	0.2125	0.2094	0.2375	0.2531	0.2469	0.2500	0.2500	0.2469	0.2500	0.2656
0.2125	0.2094	0.2375	0.2531	0.2469	0.2500	0.2500	0.2469	0.2500	0.2656	0.2594
0.2094	0.2375	0.2531	0.2469	0.2500	0.2500	0.2469	0.2500	0.2656	0.2594	0.2656
0.2375	0.2531	0.2469	0.2500	0.2500	0.2469	0.2500	0.2656	0.2594	0.2656	0.2594
0.2531	0.2469	0.2500	0.2500	0.2469	0.2500	0.2656	0.2594	0.2656	0.2594	0.2625
0.2469	0.2500	0.2500	0.2469	0.2500	0.2656	0.2594	0.2656	0.2594	0.2625	0.2688
0.2500	0.2500	0.2469	0.2500	0.2656	0.2594	0.2656	0.2594	0.2625	0.2688	0.2719

รูปที่ 4.3 แสดงตัวอย่างข้อมูลราคาปิดหลักทรัพย์อนุกรมเวลา 10 วันที่ผ่านการแปลงค่าข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

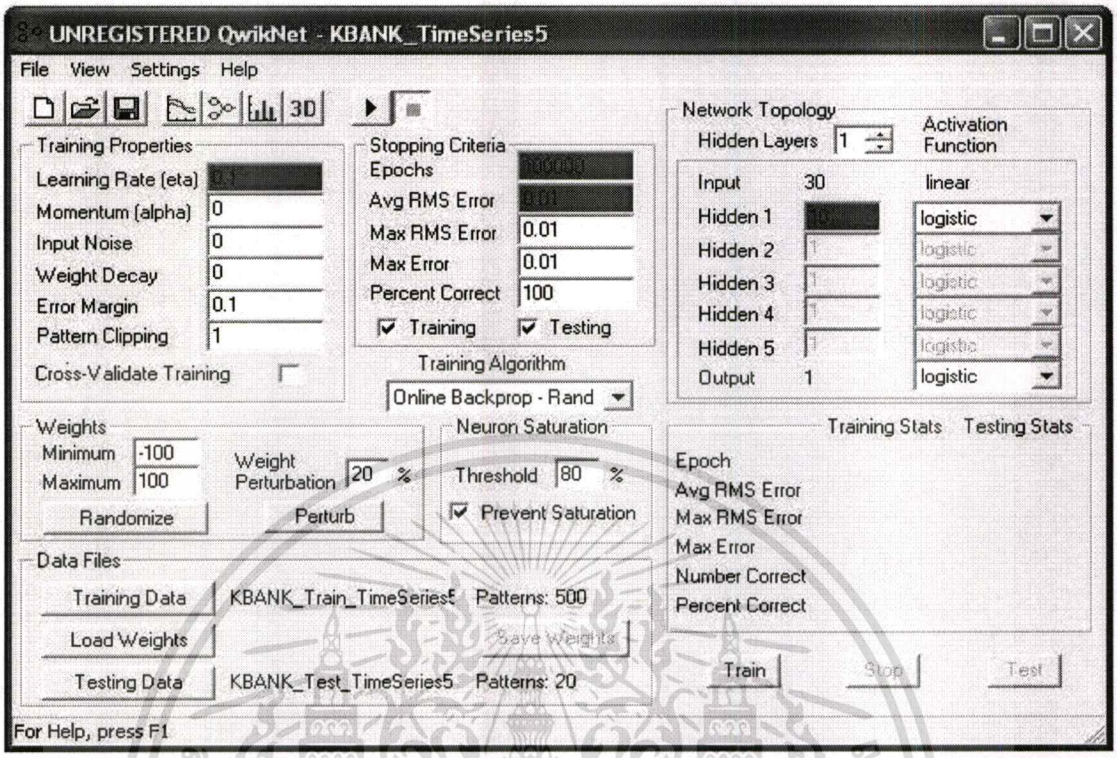


รูปที่ 4.4 แสดงหน้าจอการนำเข้าข้อมูลเพื่อสร้างแบบจำลองพยากรณ์ฯ

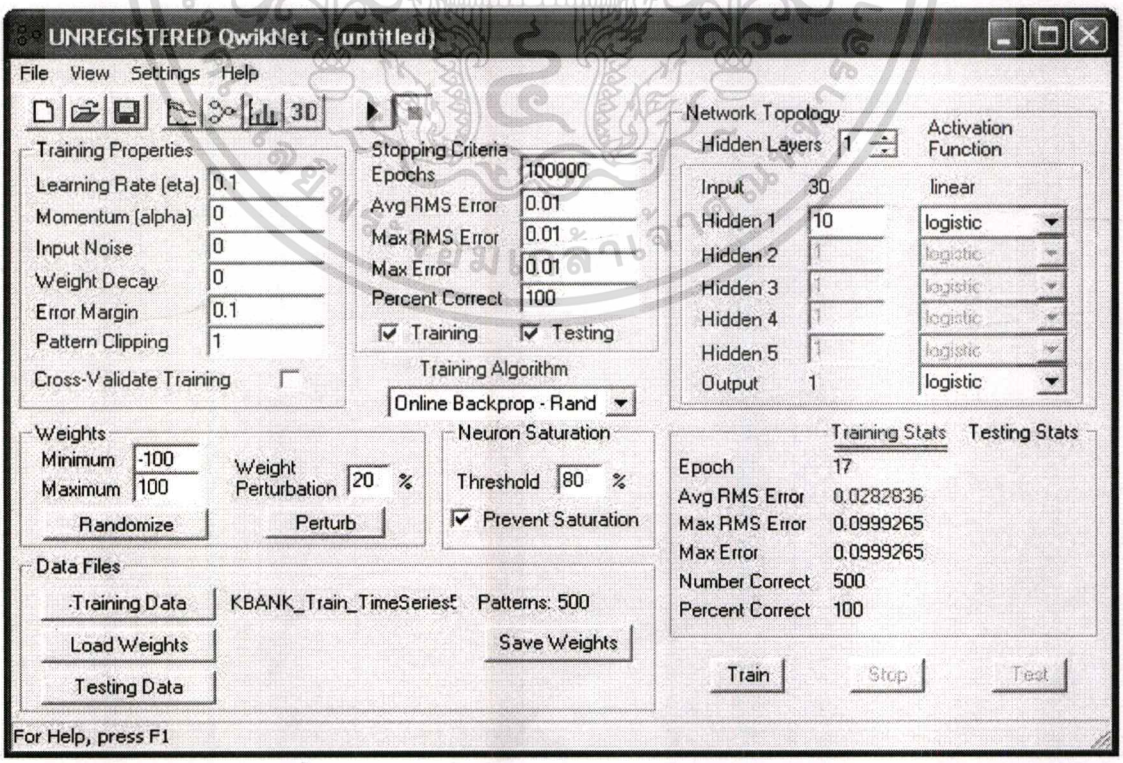


รูปที่ 4.5 แสดงหน้าจอสำหรับกำหนดจำนวนอินพุตและเอาต์พุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

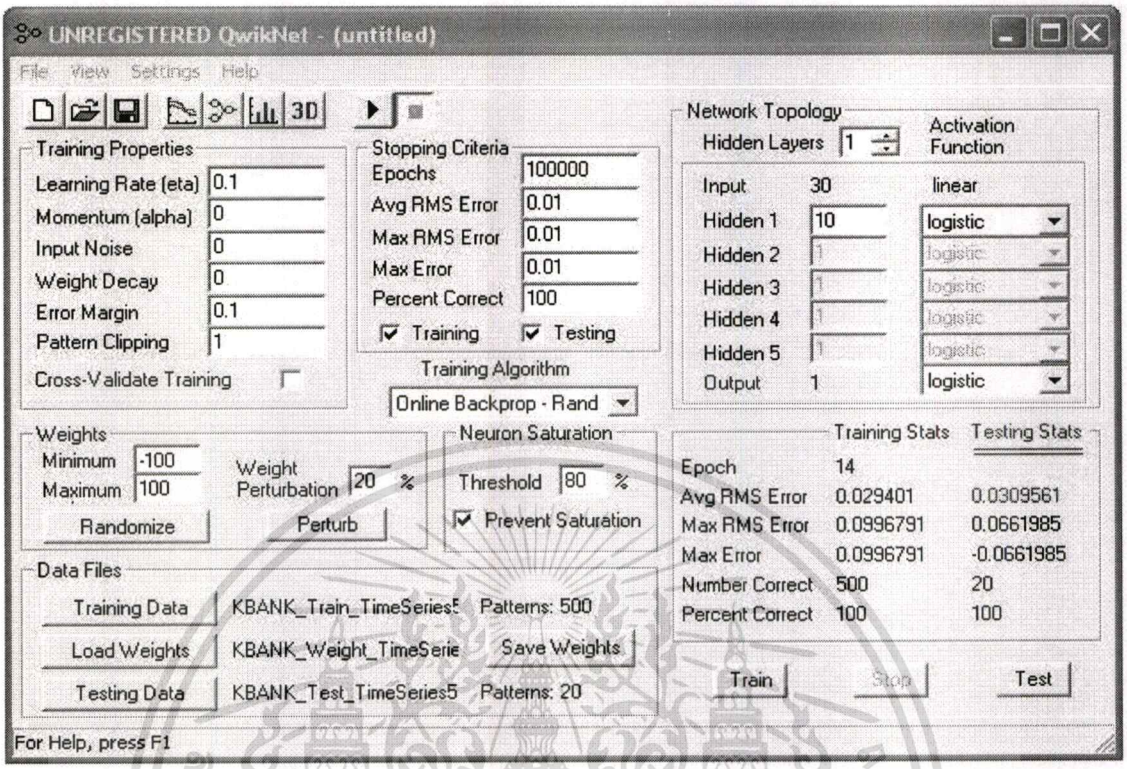


รูปที่ 4.6 แสดงหน้าจอการกำหนดค่าตัวแปรต่างๆเพื่อสร้างแบบจำลองพยากรณ์ฯ

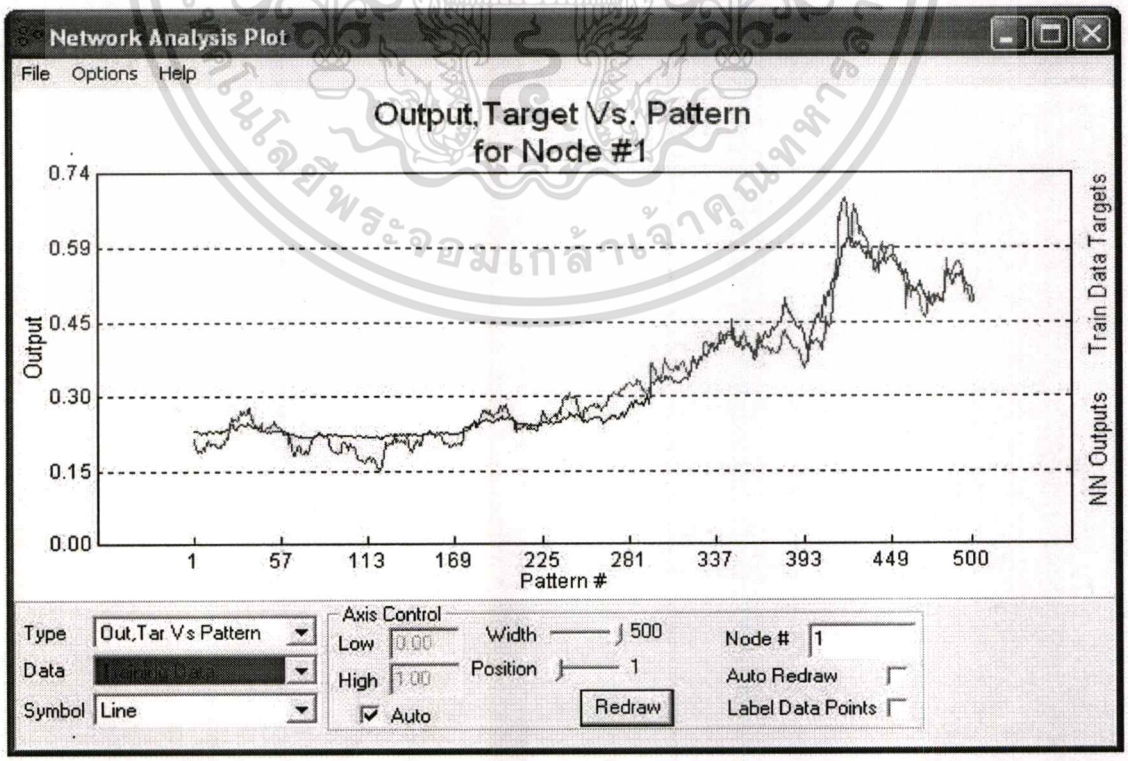


รูปที่ 4.7 แสดงหน้าจอการสร้างแบบจำลองพยากรณ์ฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

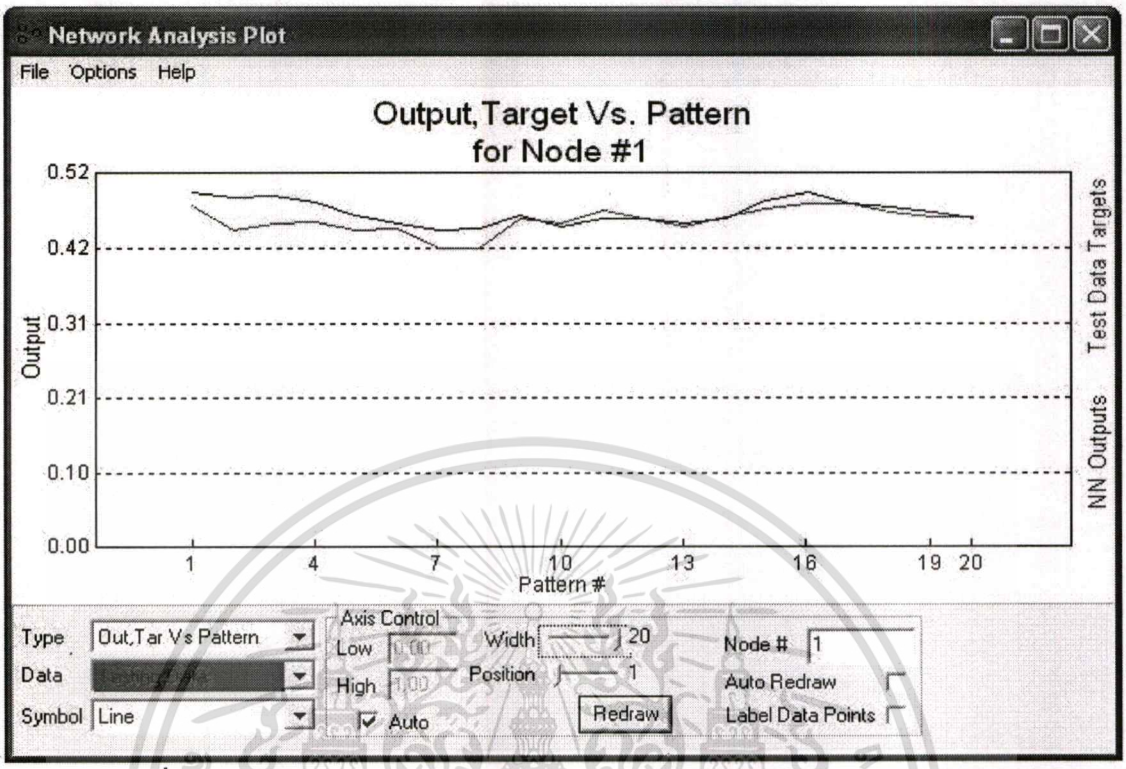


รูปที่ 4.8 แสดงหน้าจอการทดสอบแบบจำลองพยากรณ์ฯ

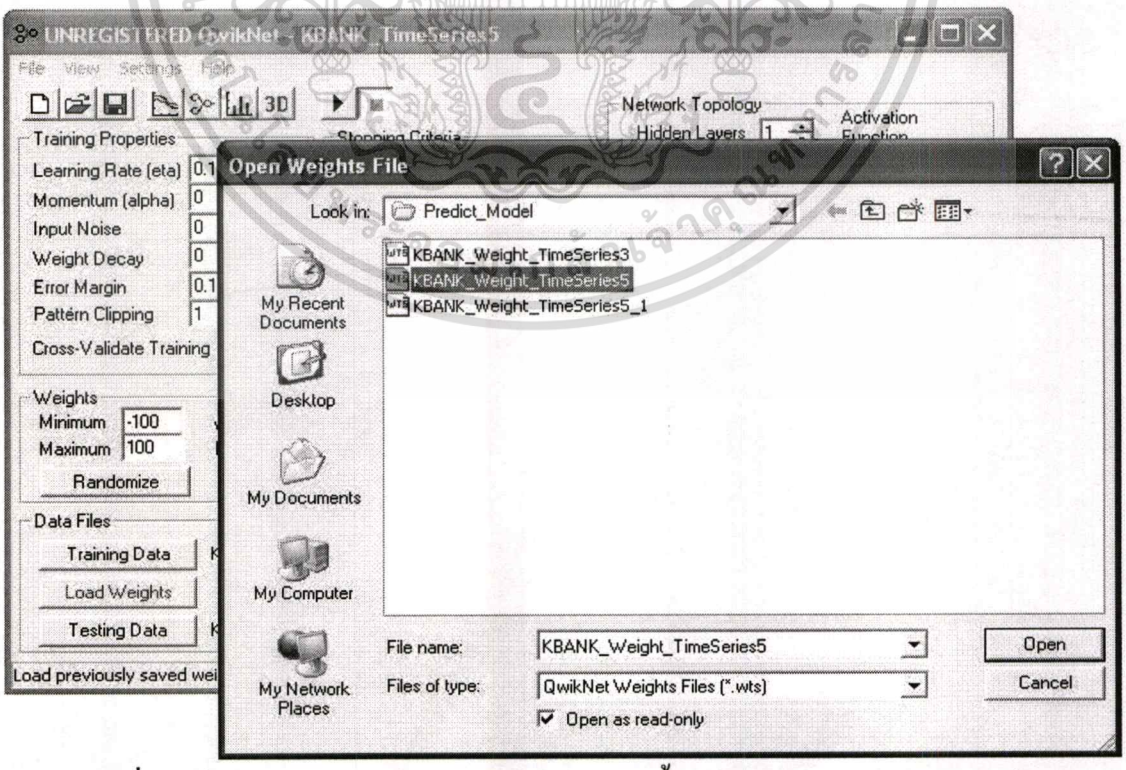


รูปที่ 4.9 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลเรียนรู้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

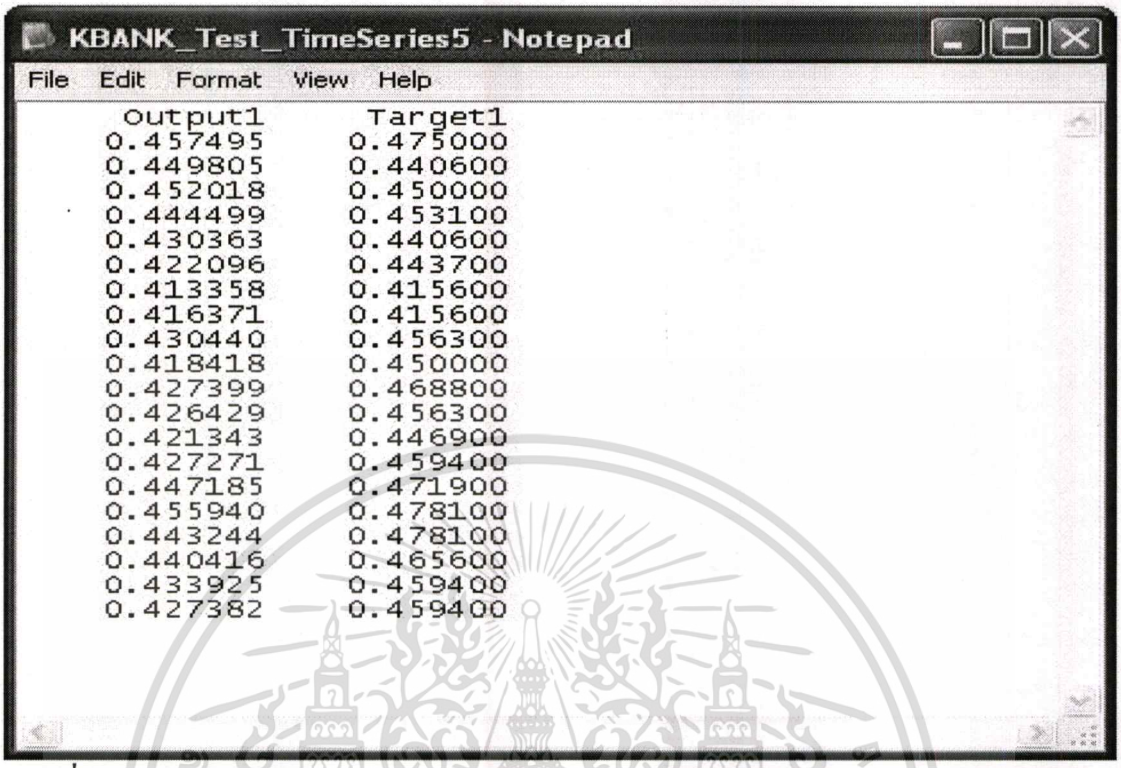


รูปที่ 4.10 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลทดสอบ



รูปที่ 4.11 แสดงตัวอย่างหน้าจอการเรียกใช้ค่าถ่วงน้ำหนักของแบบจำลองพยากรณ์ฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

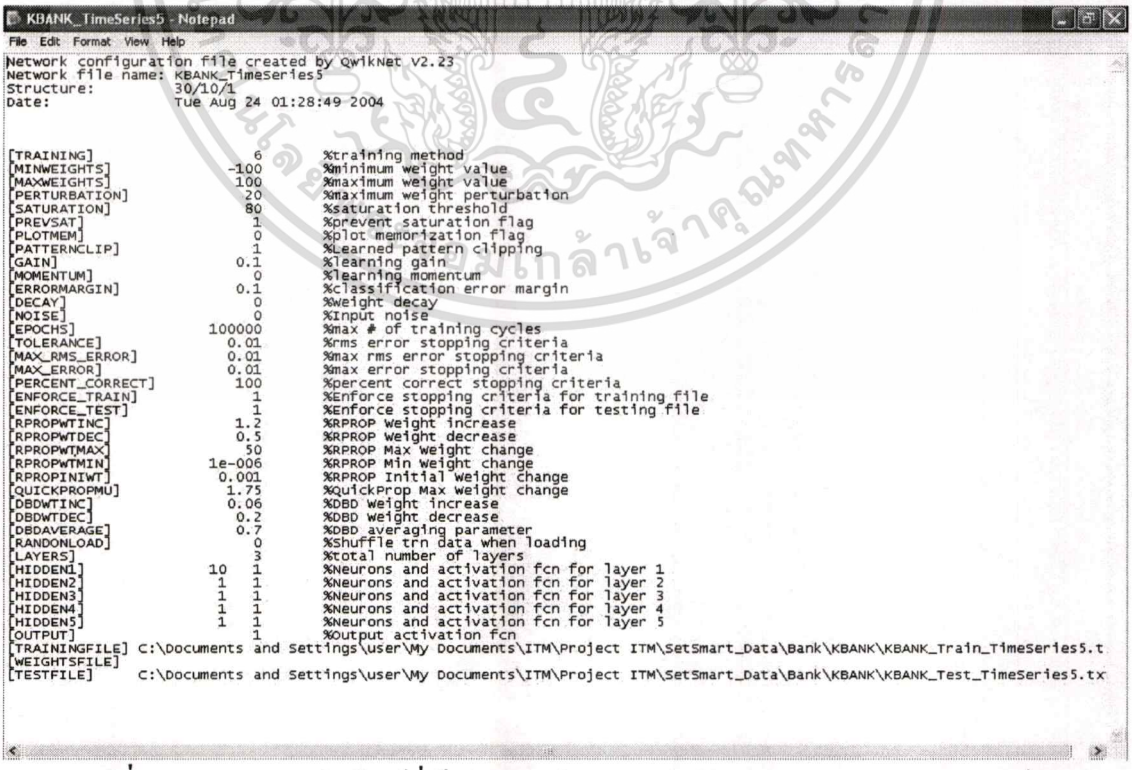


KBANK_Test_TimeSeries5 - Notepad

File Edit Format View Help

Output1	Target1
0.457495	0.475000
0.449805	0.440600
0.452018	0.450000
0.444499	0.453100
0.430363	0.440600
0.422096	0.443700
0.413358	0.415600
0.416371	0.415600
0.430440	0.456300
0.418418	0.450000
0.427399	0.468800
0.426429	0.456300
0.421343	0.446900
0.427271	0.459400
0.447185	0.471900
0.455940	0.478100
0.443244	0.478100
0.440416	0.465600
0.433925	0.459400
0.427382	0.459400

รูปที่ 4.12 แสดงตัวอย่างผลการทดสอบแบบจำลองพยากรณ์ฯกับข้อมูลทดสอบในแบบรายงาน



KBANK_TimeSeries5 - Notepad

File Edit Format View Help

```
Network configuration file created by QwikNet V2.23
Network file name: KBANK_TimeSeries5
Structure:         30/10/1
Date:              Tue Aug 24 01:28:49 2004


[TRAINING]                6       %training method
[MINWEIGHTS]           -100       %minimum weight value
[MAXWEIGHTS]            100       %maximum weight value
[PERTURBATION]           20       %maximum weight perturbation
[SATURATION]             80       %saturation threshold
[PREVSAT]                 1       %prevent saturation flag
[PLOTMEM]                 0       %plot memorization flag
[PATTERNCLIP]             1       %Learned pattern clipping
[GAIN]                  0.1       %learning gain
[MOMENTUM]                0       %learning momentum
[ERRORMARGIN]           0.1       %classification error margin
[DECAY]                   0       %weight decay
[NOISE]                   0       %Input noise
[EPOCHS]             100000       %max # of training cycles
[TOLERANCE]            0.01       %rms error stopping criteria
[MAX_RMS_ERROR]        0.01       %max rms error stopping criteria
[MAX_ERROR]            0.01       %max error stopping criteria
[PERCENT_CORRECT]       100       %percent correct stopping criteria
[ENFORCE_TRAIN]           1       %Enforce stopping criteria for training file
[ENFORCE_TEST]            1       %Enforce stopping criteria for testing file
[RPROPWTINC]            1.2       %RPROP Weight increase
[RPROPWTDEC]            0.5       %RPROP Weight decrease
[RPROPWTMAX]             50       %RPROP Max Weight change
[RPROPWTMIN]         1e-006       %RPROP Min Weight change
[RPROPINIWT]          0.001       %RPROP Initial Weight change
[QUICKPROPMU]          1.75       %QuickProp Max Weight change
[DBDWTINC]             0.06       %DBD Weight increase
[DBDWTDEC]              0.2       %DBD Weight decrease
[DBDAVERAGE]            0.7       %DBD averaging parameter
[RANDONLOAD]              0       %Shuffle trn data when loading
[LAYERS]                  3       %total number of layers
[HIDDEN1]            10   1       %Neurons and activation fcn for layer 1
[HIDDEN2]             1   1       %Neurons and activation fcn for layer 2
[HIDDEN3]             1   1       %Neurons and activation fcn for layer 3
[HIDDEN4]             1   1       %Neurons and activation fcn for layer 4
[HIDDEN5]             1   1       %Neurons and activation fcn for layer 5
[OUTPUT]                  1       %Output activation fcn
[TRAININGFILE] C:\Documents and Settings\user\My Documents\ITM\Project ITM\SetSmart_Data\Bank\KBANK\KBANK_Train_TimeSeries5.t
[WEIGHTSFILE]
[TESTFILE]     C:\Documents and Settings\user\My Documents\ITM\Project ITM\SetSmart_Data\Bank\KBANK\KBANK_Test_TimeSeries5.tx
```

รูปที่ 4.13 แสดงตัวอย่างไฟล์ที่เก็บรายละเอียด ตัวแปรต่างๆของแบบจำลองพยากรณ์ฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

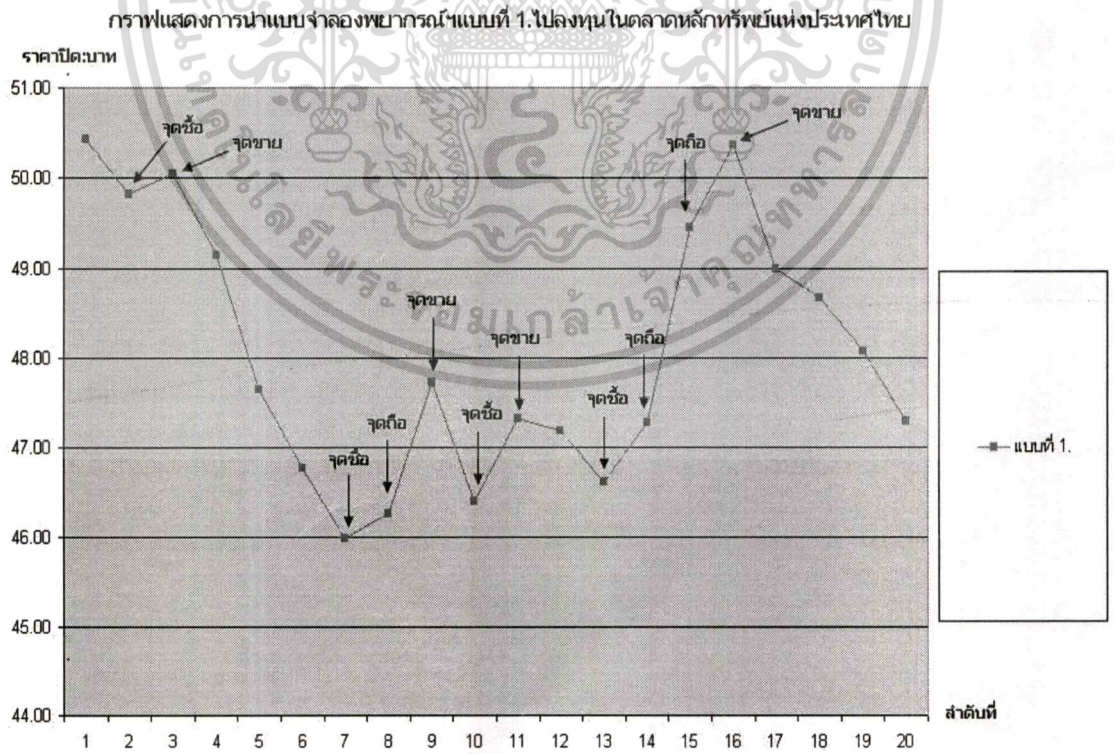
4.3 การใช้งานระบบพยากรณ์ราคาหลักทรัพย์ สำหรับลงทุนในตลาดหลักทรัพย์ฯ

เป็นขั้นตอนในการวิเคราะห์และนำระบบพยากรณ์ฯ ไปใช้งานซึ่งต้องอาศัยทักษะในการวิเคราะห์ข้อมูลในเชิงเทคนิค โดยนำผลลัพธ์คือราคาปิดหลักทรัพย์ในวันถัดไป ที่ได้จากแต่ละแบบพยากรณ์ฯไปสร้างเป็นกราฟ ซึ่งการตัดสินใจลงทุนจะอาศัยแนวโน้มของกราฟ และจากกราฟดังกล่าวสามารถแยกออกเป็นจุดเพื่อสนับสนุนการตัดสินใจลงทุนได้ ดังนี้

4.3.1 จุดซื้อ ให้ทำการซื้อหลักทรัพย์ตามราคาปิดในวันนี้ โดยพิจารณาจากราคาปิดหลักทรัพย์ในวันถัดไป ที่ระบบพยากรณ์ฯพยากรณ์ว่าราคาจะเพิ่มสูงขึ้น และในพอร์ตการลงทุนไม่มีหุ้นนั้นๆถือครองอยู่

4.3.2 จุดถือ ไม่ให้ทำการซื้อขายหลักทรัพย์ในวันนี้ โดยพิจารณาจากราคาปิดหลักทรัพย์ในวันถัดไป ที่ระบบพยากรณ์ฯพยากรณ์ว่าราคาจะเพิ่มสูงขึ้นอีก และในพอร์ตการลงทุนมีหุ้นนั้นๆถือครองอยู่

4.3.3 จุดขาย ให้ทำการขายหลักทรัพย์ตามราคาปิดในวันนี้ โดยพิจารณาจากราคาปิดหลักทรัพย์ในวันถัดไป ที่ระบบพยากรณ์ฯพยากรณ์ว่าราคาจะลดลง และในพอร์ตการลงทุนมีหุ้นนั้นๆถือครองอยู่



รูปที่ 4.14 กราฟแสดงแนวโน้มการเปลี่ยนแปลงของราคาหลักทรัพย์ จากแบบจำลองพยากรณ์ฯ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากที่นำราคาหลักทรัพย์ที่พยากรณ์ได้ มาสร้างเป็นกราฟและทำการลงทุนตามแนวโน้มหรือจุดเพื่อสนับสนุนการตัดสินใจดังกล่าว วิธีทดสอบความแบบจำลองพยากรณ์ฯจะนำวิธีการคำนวณหาอัตราผลตอบแทนจากการลงทุน โดยมีวิธีการคำนวณดังนี้

$$\text{อัตราผลตอบแทนจากการลงทุน} = \frac{\text{มูลค่าของพอร์ตการลงทุน–เงินลงทุน–ค่านายหน้า}}{\text{เงินลงทุน}} \times 100$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
การทดสอบระบบ และสรุป

5.1 การทดสอบระบบ

ในการทดสอบระบบพยากรณ์ราคาหลักทรัพย์นี้ จะนำเอาข้อมูลสำหรับทดสอบ (Test Set) มาทดสอบกับแบบจำลองพยากรณ์ฯต่างๆที่สร้างขึ้น ทั้งจากการใช้หลายๆตัวแปรตามที่ได้กล่าวไว้ในบทที่ 2 และจากข้อมูลราคาปิดของราคาหลักทรัพย์ที่นำมาจัดเรียงให้มีลักษณะเชิงอนุกรมเวลา เพื่อใช้เป็นอินพุตของระบบพยากรณ์ฯ โดยมีรายละเอียด ดังนี้

ตารางที่ 5.1 แสดงรายละเอียดข้อมูลอินพุตเพื่อใช้ในการสร้างระบบพยากรณ์ราคาหลักทรัพย์

ประเภทของข้อมูล	แหล่งที่มาของข้อมูล
• ราคาหลักทรัพย์ของ บริษัท ธนาคารกสิกรไทย จำกัด(มหาชน)	ตลาดหลักทรัพย์แห่งประเทศไทย
• ดัชนีราคาหลักทรัพย์ในประเทศ	สำนักงานคณะกรรมการกำกับหลักทรัพย์และตลาดหลักทรัพย์แห่งประเทศไทย
• ดัชนีราคาหลักทรัพย์ต่างประเทศ	
• มูลค่าการซื้อขายในตลาดหลักทรัพย์แห่งประเทศไทย	
• อัตราแลกเปลี่ยนเงินตราต่างประเทศไทย-สหรัฐเอมริกา	ธนาคารแห่งประเทศไทย
• อัตราดอกเบี้ยเงินฝากธนาคารแห่งประเทศไทย	
• ราคาน้ำมันสำเร็จรูปในประเทศไทย	บริษัท ปตท. จำกัด(มหาชน)

หลังจากนั้นจะนำข้อมูลอินพุตมาจัดเรียงโดยใช้วันที่เป็นดัชนีร่วมกัน จากนั้นจะทำการกำหนดค่าตัวแปรต่างๆเพื่อใช้ประกอบการสร้างแบบจำลองพยากรณ์ฯ แสดงได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.1.1 ข้อมูลที่ใช้ในการสร้างแบบจำลองพยากรณ์ฯ = 500 ชุดข้อมูล

5.1.2 ข้อมูลที่ใช้ในการทดสอบแบบจำลองพยากรณ์ฯ = 20 ชุดข้อมูล

5.1.3 จำนวนชั้นกลาง (Hidden Layer) = 1 ชั้น

5.1.4 จำนวนโหนดชั้นกลาง (Hidden Node) = 10 โหนด

5.1.5 อัตราการเรียนรู้ (Learning Rate) = 0.1

5.1.6 ค่าความผิดพลาด (Root Mean Square Error) = 0.01

5.1.7 จำนวนรอบที่ใช้ในการสร้างแบบจำลองพยากรณ์ฯ = 10,000 รอบ

จากการศึกษาโดยการปรับเปลี่ยนข้อมูลอินพุตที่อาศัยหลายๆตัวแปร ตามรายละเอียดข้างต้น กับข้อมูลอินพุตที่ใช้เพียงราคาปิดของหลักทรัพย์มาจัดเรียงให้อยู่ในรูปของอนุกรมเวลาในช่วงเวลาต่างๆกัน เช่น อนุกรมเวลาย้อนหลัง 5 วัน, 10 วัน และ 15 วันตามลำดับ สามารถแสดงผลการทดสอบได้ โดยนำเอาผลลัพธ์ที่แบบจำลองพยากรณ์ฯคำนวณได้ไปเปรียบเทียบกับข้อมูลจริงเพื่อวัดค่าเฉลี่ยความผิดพลาด และจะเปรียบเทียบอัตราผลตอบแทนจากการลงทุนโดยนำผลลัพธ์ที่ได้จากแต่ละแบบจำลองพยากรณ์ฯไปใช้สนับสนุนการตัดสินใจลงทุน แสดงผลการทดสอบได้ดังตารางต่อไปนี้

ตารางที่ 5.2 แสดงผลลัพธ์ที่ได้จากแบบจำลองพยากรณ์ต่างๆ

แบบจำลองพยากรณ์ราคาหลักทรัพย์	ค่าเฉลี่ยความผิดพลาด		อัตราผลตอบแทนจากการลงทุน
	เรียนรู้	ทดสอบ	
แบบที่1. สร้างจากข้อมูลหลายตัวแปรและราคาปิดหลักทรัพย์ลักษณะอนุกรมเวลา 5 วัน	0.039494	0.026635	16.88%
แบบที่2. สร้างจากราคาปิดหลักทรัพย์ลักษณะอนุกรมเวลา 5 วัน	0.026525	0.027168	00.03%
แบบที่3. สร้างจากราคาปิดหลักทรัพย์ลักษณะอนุกรมเวลา 10 วัน	0.022291	0.028704	-1.56%
แบบที่4. สร้างจากราคาปิดหลักทรัพย์ลักษณะอนุกรมเวลา 15 วัน	0.037494	0.046746	-3.59%

จากผลการทดสอบแต่ละแบบจำลองพยากรณ์ฯ สามารถนำเสนอการเปรียบเทียบค่าราคาปิดของหลักทรัพย์ที่เกิดขึ้นจริงกับผลลัพธ์ที่แต่ละแบบจำลองพยากรณ์ฯคำนวณได้ ดังตารางที่ 5.3

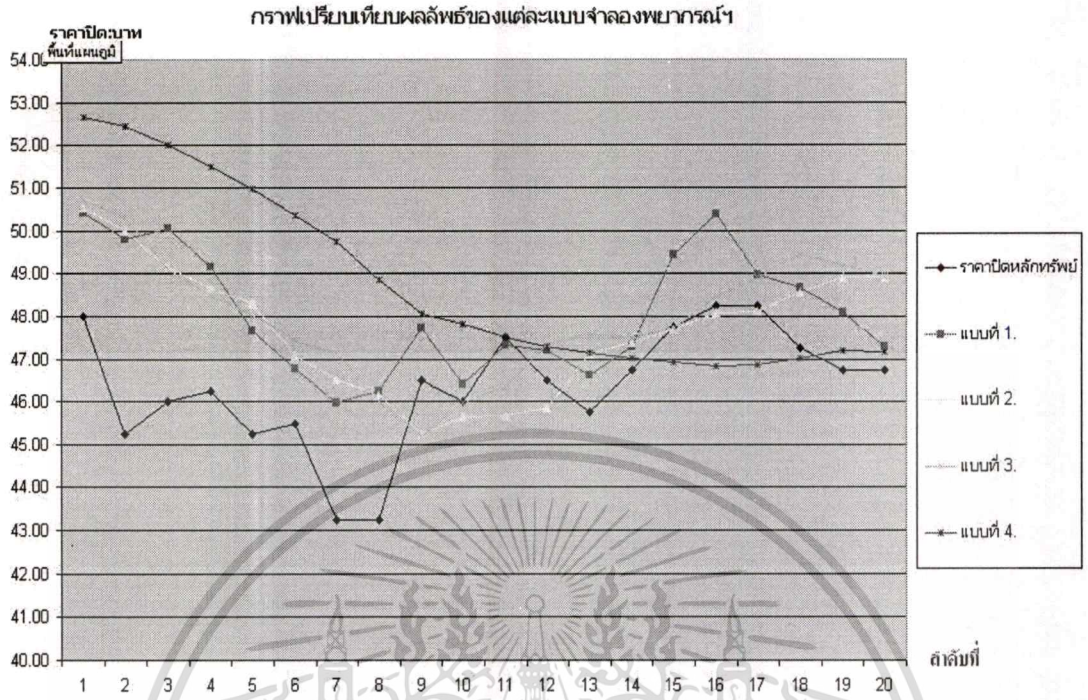
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 5.3 แสดงราคาปิดที่เกิดขึ้นจริง และราคาที่แต่ละแบบจำลองพยากรณ์ฯคำนวณได้

วันที่	ราคาปิดของหลักทรัพย์ บริษัท ธนาคารกสิกรไทย จำกัด(มหาชน)				
	เกิดขึ้นจริง	แบบที่ 1.	แบบที่ 2.	แบบที่ 3.	แบบที่ 4.
7 พ.ค.47	48.00	50.43	50.55	50.70	52.64
10 พ.ค.47	45.25	49.82	49.99	50.34	52.42
11 พ.ค.47	46.00	50.05	49.19	49.09	52.00
12 พ.ค.47	46.25	49.15	48.65	48.48	51.50
13 พ.ค.47	45.25	47.65	48.30	48.02	50.96
14 พ.ค.47	45.50	46.77	47.06	47.45	50.35
17 พ.ค.47	43.25	45.99	46.53	47.14	49.73
18 พ.ค.47	43.25	46.26	46.16	46.00	48.86
19 พ.ค.47	46.50	47.72	45.23	45.17	48.06
20 พ.ค.47	46.00	46.40	45.62	45.98	47.82
21 พ.ค.47	47.50	47.33	45.68	46.35	47.51
24 พ.ค.47	46.50	47.18	45.84	47.34	47.29
25 พ.ค.47	45.75	46.62	47.10	47.55	47.14
26 พ.ค.47	46.75	47.28	47.38	47.47	47.01
27 พ.ค.47	47.75	49.45	47.70	47.73	46.91
28 พ.ค.47	48.25	50.38	48.09	48.46	46.84
31 พ.ค.47	48.25	48.99	48.12	48.92	46.86
1 มิ.ย.47	47.25	48.68	48.54	49.43	47.03
3 มิ.ย.47	46.75	48.08	48.90	49.18	47.19
4 มิ.ย.47	46.75	47.29	48.91	48.94	47.16

จากผลลัพธ์ของตารางที่ 5.3 สามารถนำมาเปรียบเทียบและแสดงผลได้ดังรูปที่ 5.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.1 แสดงผลการเปรียบเทียบค่าข้อมูลจริงกับค่าข้อมูลที่ได้จากแต่ละแบบจำลองพยากรณ์ฯ

5.2 สรุปผลการทดสอบ

จากผลการพยากรณ์ราคาหลักทรัพย์โดยใช้อัลกอริทึมแบล็คพร็อบพาเกชั่น โดยที่นำข้อมูลหลายๆตัวแปร ประกอบกับข้อมูลราคาปิดของหลักทรัพย์ในอดีตที่มีลักษณะเป็นอนุกรมเวลามาใช้เป็นอินพุตของระบบพยากรณ์ฯดังกล่าว จะให้อัตราผลตอบแทนการลงทุนสูงกว่าหากเปรียบเทียบในเชิงคุณภาพ เนื่องจากแบบจำลองพยากรณ์ฯมีการเรียนรู้ที่ดีกว่า ทำให้สามารถปรับเปลี่ยนการเปลี่ยนแปลงของราคาหลักทรัพย์ในทิศทางตรงข้ามได้เร็วกว่า แบบจำลองพยากรณ์ฯที่นำข้อมูลราคาปิดหลักทรัพย์ที่มีลักษณะอนุกรมเวลามาใช้เป็นอินพุตเท่านั้น ที่เมื่อเกิดการเปลี่ยนแปลงของราคาหลักทรัพย์ในทิศทางตรงข้ามแล้วไม่สามารถปรับเปลี่ยนราคาที่พยากรณ์ได้ทันที ตามผลการทดสอบที่ปรากฏ แต่ถ้าหากเปรียบเทียบในเชิงปริมาณ ค่าเฉลี่ยความผิดพลาดจะมีค่าใกล้เคียงกันซึ่งถือว่าเป็นค่าที่ยอมรับได้

ดังนั้นระบบการพยากรณ์ราคาหลักทรัพย์นี้สามารถนำไปใช้เป็นเครื่องมือ หรือระบบเพื่อสนับสนุนการตัดสินใจลงทุนหลักทรัพย์ในตลาดหลักทรัพย์แห่งประเทศไทยได้จริง ตรงตามวัตถุประสงค์ของโครงการพัฒนาระบบและวิชาโครงการศึกษากรณีพิเศษนี้ ถึงกระนั้นก็ตามการกำหนดค่าตัวแปรต่างๆ เพื่อใช้ประกอบการสร้างแบบจำลองพยากรณ์ยังเป็นสิ่งที่ต้องคำนึงถึง เพื่อให้ระบบสามารถสร้างแบบจำลองพยากรณ์ที่มีประสิทธิภาพสูงขึ้นต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

กัลยา วานิชย์บัญชา. 2545. **การวิเคราะห์สถิติ: สำหรับการบริหารและวิจัย.**พิมพ์ครั้งที่ 6.กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.

ตลาดหลักทรัพย์แห่งประเทศไทย. 2542. **ตลาดหุ้นในประเทศไทย.**กรุงเทพฯ. [Online] เข้าถึงได้จาก : http://www.set.or.th/th/index.html.

Alex Berson and Stephen J. Smith. 1997. **Data Warehousing, Data Mining & OLAP.** McGraw Hill.

Cabana, Peter etac. 1997. **Discovery Data Mining: from Concept to Implementation.** New Jersey: Prentice Hall.

Han, J. and Kamber. 2001. **Data Mining: Concepts and Techniques.** San Diego: Morgan Kaufmann.

Michael J.A. Berry and Gordon Linoff. 1997. **Data Mining Techniques.** John Wiley & Sons Inc.

Richard J.Roiger and Michael W.Geatz. 2003. **Data Mining A Tutorial-Based Primer.** United States of America.

Robert Groth. 1997. **Data Mining.** Prentice Hall PTR.

Sholom M. Weiss and Nitin Indurkhya. 1998. **Predictive Data Mining.** Margan Kaufmann Publishers Inc.

Vasant Dhar and Roger Stein. 1997. **Seven Methods for Transforming Corporate Data Into Business Intelligence.** Prentice Hall.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อผู้เขียน	นายธีรพงษ์ วรวงศ์
วันเดือนปีเกิด	9 ตุลาคม พ.ศ.2519
สถานที่เกิด	จ.ชลบุรี
วุฒิการศึกษาระดับปริญญาตรี	บริหารธุรกิจบัณฑิต สาขาการเงิน
สถานที่สำเร็จการศึกษา	มหาวิทยาลัยกรุงเทพ
ปีที่สำเร็จการศึกษา	ปีการศึกษา 2542
ประวัติการทำงาน	2542 – ปัจจุบัน บริษัท อินชัวร์เนต จำกัด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้