

ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

ระบบการพัฒนาระบบดาต้าไมน์นิ่งโดยใช้โครงข่ายประสาทเทียม  
Development of Data Mining system using Neural Network



\*H002370\*

โดย

ปาณิสตา เย็นระสา

รหัสประจำตัว 47066107

อาจารย์ที่ปรึกษา

รศ.ดร.วรพจน์ กวีสุระเดช

วัน เดือน ปี.....	24 ก.พ. 2550
เลขทะเบียน.....	02370
เลขเรียกหนังสือ.....	วคท. 5458 2548
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."	

6117120 57  
11285 195x

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการพัฒนาระบบงาน  
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ  
ภาคฤดูร้อน ปีการศึกษา 2548  
คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	การพัฒนาระบบการค้าไม้เนื้อแข็งโดยใช้โครงข่ายประสาทเทียม
นักศึกษา	นางสาวปาณิสรา เขื่อนนะสา
อาจารย์ที่ปรึกษา	รศ.ดร. วรพงษ์ กรีสุระเดช
ระดับการศึกษา	วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2548

### บทคัดย่อ

ในปัจจุบันแต่ละองค์กรมีการเก็บข้อมูลที่ได้จากการดำเนินงานจำนวนมาก ซึ่งการนำข้อมูลเหล่านี้มาทำให้กลายเป็นข้อมูลที่มีประโยชน์ต่อองค์กรนั้น ถือได้ว่าเป็นสิ่งสำคัญ ดังนั้นจึงได้มีการนำเอาหลักการของดาต้าไมน์นิ่งมาประยุกต์ใช้ในการค้นหารูปแบบความสัมพันธ์ของข้อมูล และวิเคราะห์หาข้อมูลที่เป็นประโยชน์ เพื่อใช้เป็นแนวทางในการสนับสนุนการตัดสินใจให้ดียิ่งขึ้น

โครงการนี้เป็นการพัฒนาระบบโครงข่ายประสาทเทียม ซึ่งเป็นเทคนิคหนึ่งของดาต้าไมน์นิ่ง โดยใช้หลักการการจดจำและกระบวนการเรียนรู้แบบแบคพรอพพาเกชัน ระบบนี้จะช่วยในการวิเคราะห์และทำนายแนวโน้มของข้อมูลในอนาคต ซึ่งในการพัฒนาระบบใช้โปรแกรม Microsoft Visual Studio.net 2003 ในการสร้างโครงข่ายประสาทเทียม และใช้ Microsoft SQL Server 2000 เป็นฐานข้อมูล

<b>Title Security</b>	Development of Data Mining System using Neural Network
<b>Student</b>	Miss.Panisa Yennasa
<b>Advisor</b>	Assoc. Prof. Dr. Worapoj Kreesuradej
<b>Level of Study</b>	Master of Science in Information Technology
<b>Major</b>	Information Science
<b>Academic Year</b>	2005

## ABSTRACT

Nowadays, each organization tries to keep huge number of transactional data. These coming data can become information that useful for the organization, which means these data are very important. Thus, there are a lot of data mining theories to use in pattern recognition and seeking for data's relationship pattern. Also, analyze useful information in order to give the best direction for supporting better decision.

This project is made to develop the Neural Network system which is one of the data mining techniques. The Neural network system uses theories of recognition and back-propagation learning process. This system will help a lot in analyzing and predicting possible direction of data in the future. In order to develop this project, it necessary to use Microsoft Visual Studio.net 2003 for building the neural network and use Microsoft SQL server 2000 as a database.

## กิตติกรรมประกาศ

ในการพัฒนาโครงการพัฒนาระบบค้ำไม้หนึ่งโดยใช้โครงข่ายประสาทเทียม สามารถสำเร็จลุล่วงไปด้วยดี ข้าพเจ้าขอขอบพระคุณ ผศ.ดร. วรพจน์ กรีสุระเดช อาจารย์ที่ปรึกษา ที่กรุณาให้คำแนะนำ ข้อเสนอแนะ และเป็นที่ปรึกษาในการแก้ปัญหาต่างๆ รวมทั้งเป็นผู้ตรวจสอบความถูกต้องของ โครงการนี้

นอกจากนี้ข้าพเจ้าขอขอบพระคุณ บิดา มารดา ที่ได้ให้ความสนับสนุนทางด้านกำลังใจและทุนทรัพย์จนการทำโครงการนี้สำเร็จด้วยดี ขอขอบพระคุณอาจารย์ทุกท่านที่ให้ความรู้ทั้งในภาคทฤษฎีและภาคปฏิบัติ รวมทั้งรุ่นพี่ และเพื่อนๆ ทุกคนที่มีส่วนช่วยเหลือในด้านต่างๆ เกี่ยวกับโครงการไว้ ณ ที่นี้ด้วย

ปาณิสตา เย็นนะสา

# สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่	
1. บทนำ.....	1
1.1 ความเป็นมาและปัญหา.....	1
1.2 วัตถุประสงค์.....	1
1.3 ขอบเขตการพัฒนา.....	2
1.4 ขั้นตอนและวิธีการดำเนินงาน.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
2. ทฤษฎีค้ำไม้หนึ่ง.....	3
2.1 ความหมายของค้ำไม้หนึ่ง.....	3
2.2 กระบวนการทำงานของค้ำไม้หนึ่ง.....	3
2.3 วิธีการและเทคนิคที่ใช้ในการทำค้ำไม้หนึ่ง.....	6
2.4 ประโยชน์ของค้ำไม้หนึ่ง.....	8
2.5 ตัวอย่างการประยุกต์ใช้ค้ำไม้หนึ่ง.....	8
3. โครงข่ายประสาทเทียม.....	10
3.1 ความหมายของโครงข่ายประสาทเทียม.....	10
3.2 โครงสร้างของโครงข่ายประสาทเทียม.....	11
3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม.....	12
3.4 การทำงานของเซลล์ประสาทเทียม.....	13
3.5 การเรียนรู้ของโครงข่ายประสาทเทียม.....	13

## สารบัญ(ต่อ)

หน้า

4. การวิเคราะห์และออกแบบระบบ .....	19
4.1 รายละเอียดของระบบ .....	19
4.2 การจำลองระบบงานด้วย UML .....	21
4.3 รายละเอียดเกี่ยวกับตารางที่ใช้ในระบบ .....	30
5. การพัฒนาระบบ .....	33
5.1 ผังโครงสร้างของระบบ .....	33
5.2 อัลกอริทึมที่ใช้ในการพัฒนาระบบ .....	39
5.3 การใช้งานระบบ และตัวอย่างหน้าจอที่ได้จากการใช้งานระบบ .....	43
6. สรุปผลและข้อเสนอแนะ .....	50
6.1 การทดสอบระบบ .....	56
6.2 การสรุปผลการทดลอง .....	58
บรรณานุกรม.....	60
ประวัติผู้เขียน .....	61

## สารบัญตาราง

ตารางที่	หน้า
4.1 คำอธิบายยูสเคสไดอะแกรมของ Build Model.....	22
4.2 คำอธิบายยูสเคสไดอะแกรมของ Prediction.....	23
4.3 ตารางที่ใช้ในการเก็บโปรเจค .....	30
4.4 ตาราง ProjectDesc .....	30
4.5 ตาราง ConfigModel .....	30
4.6 ตาราง ProjectModel.....	31
4.7 ตาราง ColumnsDesc .....	31
4.8 ตาราง ColumnsDomain .....	31
4.9 ตาราง ColumnsMaxMin.....	32
6.1 แสดงผลจากการทดสอบที่ทำการปรับจำนวนรอบที่ใช้สร้างแบบจำลอง.....	56
6.2 แสดงผลจากการทดสอบที่ทำการปรับจำนวนโหนดในชั้นฮิดเดน.....	57

# สารบัญรูป

รูปที่	หน้า
2.1 กระบวนการค้นหาความรู้ในฐานข้อมูล .....	4
3.1 โครงสร้างเซลล์ประสาทของสมองมนุษย์.....	10
3.2 โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม .....	11
3.3 สถาปัตยกรรมของ Feedforward Network .....	12
3.4 สถาปัตยกรรมของ Feedback Network .....	12
3.5 การคำนวณของเซลล์ประสาทภายในโครงข่ายประสาทเทียม .....	13
3.6 กระบวนการเรียนรู้แบบแบคพรอพพาเกชัน.....	15
4.1 ยูสเคสไดอะแกรมของระบบการค้าไม้หนึ่ง โดยใช้โครงข่ายประสาทเทียม.....	22
4.2 ซีเควนซ์ไดอะแกรมของการเชื่อมต่อฐานข้อมูล.....	23
4.3 ซีเควนซ์ไดอะแกรมของการเลือกข้อมูล.....	24
4.4 ซีเควนซ์ไดอะแกรมของการแก้ไขข้อมูล.....	25
4.5 ซีเควนซ์ไดอะแกรมของการแปลงค่าข้อมูล .....	26
4.6 ซีเควนซ์ไดอะแกรมของการสร้างแบบจำลอง .....	27
4.7 ซีเควนซ์ไดอะแกรมของการบันทึกโปรเจค.....	28
5.1 แสดงผังโครงสร้างหลักของระบบการค้าไม้หนึ่ง โดยใช้โครงข่ายประสาทเทียม.....	33
5.2 แสดงผังโครงสร้างของการสร้างแบบจำลอง.....	34
5.3 แสดงผังงานโฟว์ชาร์ตการแก้ไขข้อมูล.....	35
5.4 แสดงผังงานโฟว์ชาร์ตการไม้หนึ่งข้อมูล.....	37
5.5 แสดงผังโครงสร้างของการพยากรณ์ข้อมูล .....	38
5.6 หน้าจอแรกที่ปรากฏเมื่อเริ่มใช้งาน .....	43
5.7 หน้าจอหลักในส่วนของการสร้างแบบจำลองใหม่ .....	44
5.8 หน้าจอการเชื่อมต่อฐานข้อมูล SQL Server 2000.....	45
5.9 หน้าจอการเลือกตาราง .....	45
5.10 หน้าจอการระบุคอล์มน์ .....	46
5.11 หน้าจอแสดงชนิดข้อมูลและข้อมูลที่ใช้ในการสร้างแบบจำลอง.....	46

## สารบัญรูป(ต่อ)

รูปที่	หน้า
5.12 หน้าจอการเลือกวิธีการแก้ไขข้อมูล.....	47
5.13 หน้าจอการแปลงค่าข้อมูล .....	48
5.14 หน้าจอแสดงผลการแปลงค่าข้อมูลให้อยู่ในช่วง 0 - 1.....	48
5.15 หน้าจอการสร้างแบบจำลอง และการทดสอบแบบจำลอง.....	49
5.16 หน้าจอแสดงผลการสร้างแบบจำลองในรูปแบบกราฟ.....	50
5.17 หน้าจอแสดงผลการทดสอบแบบจำลองในรูปแบบกราฟ.....	50
5.18 หน้าจอแสดงผลการทดสอบแบบจำลองในรูปแบบรายงาน.....	51
5.19 หน้าจอการบันทึกแบบจำลองใหม่ .....	51
5.20 หน้าจอแสดง โครงสร้างของแบบจำลอง พร้อมกับข้อมูลอินพุตและข้อมูลเอาต์พุต.....	52
5.21 หน้าจอการระบุคอลัมน์ที่นำมาใช้พยากรณ์ข้อมูล .....	53
5.22 หน้าจอการแก้ไขข้อมูลที่นำมาใช้พยากรณ์ข้อมูล.....	53
5.23 หน้าจอการแปลงค่าข้อมูลให้อยู่ในช่วง 0 - 1.....	54
5.24 หน้าจอแสดงผลการแปลงค่าข้อมูลให้อยู่ในช่วง 0 - 1.....	54
5.25 หน้าจอแสดงผลลัพธ์ที่ได้จากการพยากรณ์ข้อมูล .....	55
6.1 หน้าจอแสดงผลการสร้างแบบจำลองที่ทำการเรียนรู้จำนวน 10,000รอบ.....	57
6.2 หน้าจอแสดงผลการทดสอบแบบจำลองที่ทำการเรียนรู้จำนวน 10,000รอบ .....	58

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและปัญหา

ปัจจุบันการดำเนินงานต่างๆ ในแต่ละองค์กรมีการใช้และเก็บข้อมูลเป็นจำนวนมาก และเพื่อไม่ให้ข้อมูลที่เก็บเหล่านั้นสูญเปล่า จึงได้มีการนำเอาหลักการของดาต้าไมน์นิ่งเข้ามาประยุกต์ใช้ เพื่อช่วยวิเคราะห์ข้อมูลที่มีปริมาณมากให้กลายเป็นข้อมูลที่มีประโยชน์ต่อองค์กร และจากการที่ดาต้าไมน์นิ่งถูกนำไปใช้งานและวัตถุประสงค์ที่หลากหลาย ทำให้มีการพัฒนาเทคนิคต่างๆ ที่เหมาะเพื่อใช้กับงานแต่ละประเภท ซึ่งวิธีการที่ต่างกันก็จะมีเหมาะสมกับวัตถุประสงค์ในงานด้านต่างๆ กันด้วย

การสร้างแบบจำลองพยากรณ์ (Predictive Modeling) เป็นเทคนิคหนึ่งของดาต้าไมน์นิ่งที่ใช้สำหรับงานทางด้านการพยากรณ์ข้อมูล ซึ่งสามารถนำมาประยุกต์ใช้กับหลักการและวิธีการของโครงข่ายประสาทเทียม (Neural Network) เพื่อพัฒนาระบบสำหรับวิเคราะห์และทำนายแนวโน้มของข้อมูลในอนาคต โดยสามารถนำมาใช้เป็นแนวทางในการสนับสนุนการตัดสินใจให้ดียิ่งขึ้น

### 1.2 วัตถุประสงค์

1. เพื่อศึกษาหลักการและวิธีการของดาต้าไมน์นิ่ง รวมทั้งเทคนิคการสร้างแบบจำลองพยากรณ์ (Predictive Modeling) ซึ่งเป็นเทคนิคหนึ่งที่ใช้ในการทำดาต้าไมน์นิ่ง
2. เพื่อศึกษาหลักการการทำงานของโครงข่ายประสาทเทียม และกระบวนการเรียนรู้แบบแบคพรอพพาเกชัน (Backpropagation)
3. เพื่อศึกษาวิธีการของพัฒนาระบบงานที่มีประสิทธิภาพ ในการสร้างระบบเพื่อการพยากรณ์ข้อมูล
4. เพื่อศึกษาวิธีการทำงานของ Microsoft SQL Server 2000 ในส่วนของฐานข้อมูล
5. เพื่อพัฒนาระบบดาต้าไมน์นิ่ง โดยใช้โครงข่ายประสาทเทียม สำหรับใช้ในการสร้างแบบจำลอง และการพยากรณ์ข้อมูล

### 1.3 ขอบเขตของการพัฒนา

1. พัฒนาระบบงานตามหลักการของคาค้าไมน์นึ่ง และใช้เทคนิคการสร้างแบบจำลองพยากรณ์ ซึ่งระบบงานที่ได้มีความสามารถในการสร้างแบบจำลองโดยใช้หลักการของโครงข่ายประสาทเทียมแบบแบคพรอพพากชัน และการพยากรณ์ข้อมูลที่ให้ค่าเป็นตัวเลข
2. พัฒนาระบบงานโดยใช้โปรแกรม Microsoft Visual Studio.net 2003 ร่วมกับใช้ Microsoft SQL Server 2000 เป็นฐานข้อมูล

### 1.4 ขั้นตอนและวิธีการดำเนินงาน

1. ศึกษาหลักการและวิธีการในการทำคาค้าไมน์นึ่ง รวมทั้งเทคนิคการสร้างแบบจำลองพยากรณ์
2. ศึกษาหลักการการทำงานของโครงข่ายประสาทเทียม และกระบวนการเรียนรู้แบบแบคพรอพพากชัน
3. ศึกษาวิธีการของการพัฒนาระบบงานที่ใช้สำหรับการพยากรณ์ข้อมูล
4. ศึกษาวิธีการทำงานของ SQL Server 2000 ในส่วนของฐานข้อมูล
5. ออกแบบและพัฒนาระบบงานตามหลักการที่ได้ศึกษา
6. ทดสอบระบบงานและตรวจสอบข้อผิดพลาดต่างๆ เพื่อทำการปรับปรุงและแก้ไข แล้วพัฒนาให้สมบูรณ์
7. สรุปผลการทดสอบ

### 1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อเข้าใจหลักการและวิธีการของคาค้าไมน์นึ่ง รวมทั้งเทคนิคการสร้างแบบจำลองพยากรณ์
2. เพื่อเข้าใจหลักการการทำงานของโครงข่ายประสาทเทียมแบบแบคพรอพพากชัน
3. สามารถนำหลักการและวิธีการที่ได้ศึกษามาประยุกต์ใช้ในการออกแบบและพัฒนา ระบบงานจริง
4. ระบบงานที่มีความสามารถในการสร้างแบบจำลองและการพยากรณ์ข้อมูล เพื่อใช้เป็นแนวทางในการสนับสนุนการตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

### ทฤษฎีดาต้าไมน์นิ่ง

ในปัจจุบันการจัดการทางด้านข้อมูลถือได้ว่าเป็นสิ่งสำคัญ จึงได้มีการนำเอาหลักการดาต้าไมน์นิ่งเข้ามาประยุกต์ใช้ เนื่องจากสามารถวิเคราะห์ข้อมูลหลายๆ รูปแบบที่มีปริมาณมากให้กลายเป็นข้อมูลที่มีประโยชน์ เพื่อช่วยในการสนับสนุนการตัดสินใจต่างๆ ที่เป็นประโยชน์ และยังสามารถคาดการณ์สิ่งที่จะเกิดขึ้นในอนาคต ได้อีกด้วย

#### 2.1 ความหมายของดาต้าไมน์นิ่ง

ดาต้าไมน์นิ่ง (Data Mining) เป็นเทคนิคที่ใช้ในการค้นหาข้อมูลเชิงวิเคราะห์ขั้นสูง สำหรับใช้จัดการกับข้อมูลที่อยู่ในฐานข้อมูลที่มีอยู่จำนวนมาก ซึ่งประกอบไปด้วยการค้นหารูปแบบหรือโครงสร้างของข้อมูล แยกแยะกลุ่มข้อมูล และเลือกเพิ่มข้อมูลที่มีคุณค่า เป็นต้น โดยมีวัตถุประสงค์หลักเพื่อใช้อธิบายข้อมูลในอดีต และทำนายข้อมูลในอนาคต

การทำดาต้าไมน์นิ่งเพื่อการอธิบาย เป็นการค้นหารูปแบบที่น่าสนใจจากกลุ่มข้อมูล รูปแบบนี้มักจะเป็นความสัมพันธ์หรือลักษณะที่เชื่อมโยงกันของข้อมูล การทำดาต้าไมน์นิ่งแบบนี้จะไม่ได้กำหนดล่วงหน้าว่าจะค้นหารูปแบบหรือแบบจำลองของอะไร แต่ให้ค้นหาทุกรูปแบบที่น่าสนใจจากข้อมูล

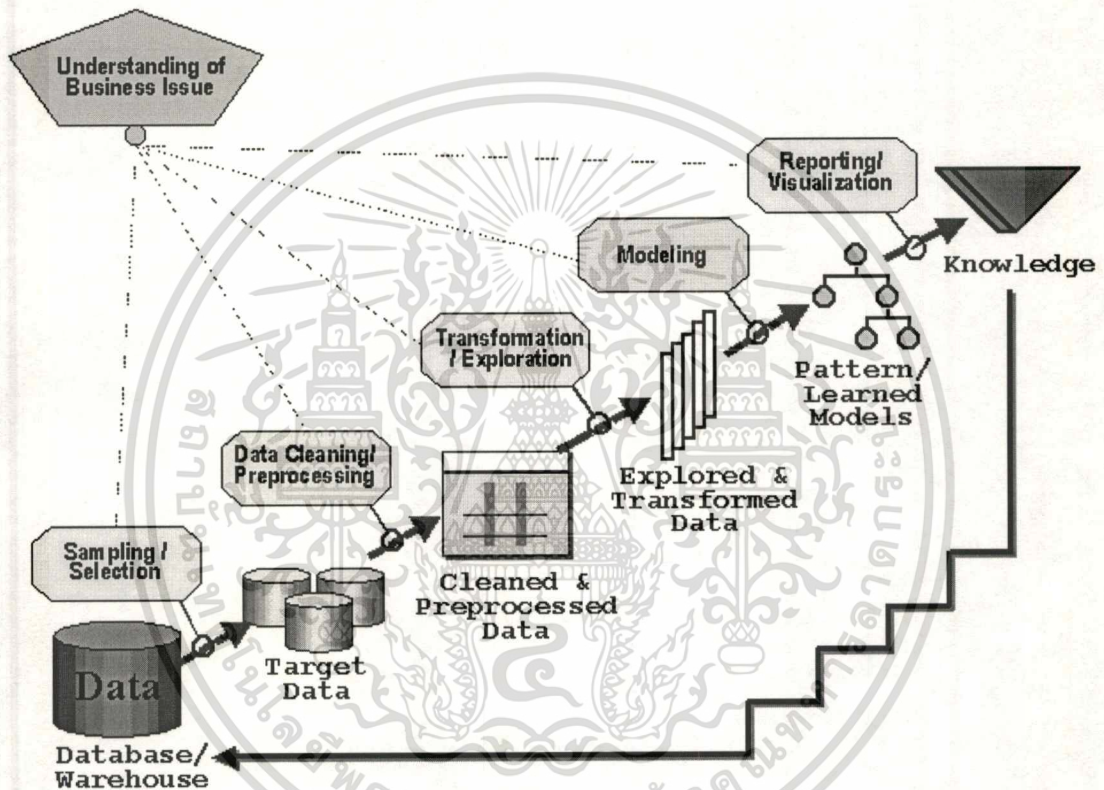
การทำดาต้าไมน์นิ่งเพื่อการทำนาย เป็นการนำความรู้ที่เรียนรู้มาจากข้อมูลที่มีอยู่เพื่อประโยชน์ในการทำนายข้อมูลใหม่ที่จะเกิดขึ้นในอนาคต เช่น จากข้อมูลลูกค้าของแผนกสินเชื่อของธนาคารที่ได้มีการจัดลำดับชั้นของลูกค้าไว้แล้วว่าใครเป็นลูกค้าชั้นดี ใครเป็นลูกค้าในระดับปานกลาง และใครเป็นลูกค้าที่มักจะผิดนัดชำระหนี้ โดยสามารถเรียนรู้จากข้อมูลเหล่านี้และค้นหาแบบจำลองที่สามารถใช้อธิบายลักษณะของลูกค้าชั้นดี ลูกค้าระดับปานกลาง และลูกค้าที่ไม่เป็นที่ต้องการ จากแบบจำลองที่ได้นี้สามารถนำไปใช้ทำนายลูกค้าใหม่ที่มาขอสินเชื่อได้ว่าน่าจะเป็นลูกค้าประเภทใด

#### 2.2 กระบวนการทำงานของดาต้าไมน์นิ่ง

ดาต้าไมน์นิ่งมีความหมายแตกต่างกันใน 2 แง่มุม คือ มุมมองทางวิชาการ และมุมมองเชิงธุรกิจ ในมุมมองเชิงวิชาการนั้นจะอ้างถึงกระบวนการทั้งหมดในการทำดาต้าไมน์นิ่งว่า การค้นหาเอกสารนี้เป็นเอกสารที่ส่งวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความรู้ในฐานข้อมูล (Knowledge discovery in database : KDD) และใช้คำว่าค้ำไม้หนึ่งแทนขั้นตอนหนึ่งของกระบวนการที่เกี่ยวข้องกับการค้นหารูปแบบ ความสัมพันธ์ของข้อมูลเท่านั้น ส่วนในแง่มุมมองเชิงธุรกิจจะใช้คำว่าค้ำไม้หนึ่งแทนความหมายของขั้นตอนทั้งหมด

จากที่ได้กล่าวแล้วว่าค้ำไม้หนึ่งในมุมมองเชิงวิชาการนั้น เป็นขั้นตอนหนึ่งที่สำคัญในกระบวนการค้นหาความรู้ในฐานข้อมูล ซึ่งประกอบด้วยขั้นตอนต่างๆ ดังนี้



รูปที่ 2.1 กระบวนการค้นหาความรู้ในฐานข้อมูล

### 2.2.1 การกำหนดวัตถุประสงค์ทางธุรกิจ (Business Objective Determination)

การกำหนดวัตถุประสงค์ทางธุรกิจ คือการกำหนดปัญหาและวัตถุประสงค์ทางธุรกิจให้ชัดเจน ซึ่งในส่วนนี้จะประกอบด้วยวิเคราะห์ทางธุรกิจเบื้องต้นว่ามีข้อมูลโดยอยู่บ้างและต้องการข่าวสารอะไรจากแหล่งข้อมูล และเป็นส่วนที่กำหนดว่าเมื่อไรที่จะใช้ค้ำไม้หนึ่งในการแก้ปัญหา

การกำหนดวัตถุประสงค์ทางธุรกิจนั้นเป็นงานที่ต้องทำเป็นสิ่งแรก เพราะเป็นตัวกำหนดถึงผลลัพธ์หรือเทคนิคที่ต้องนำมาใช้ในการทำค้ำไม้หนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 2.2.2 การเตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูล เป็นการจัดการข้อมูลให้สามารถนำเข้าสู่อัลกอริทึมของค้ำาไมน์นึ่งได้ ประกอบด้วยขั้นตอนย่อย 3 ขั้นตอน คือ

1. การเลือกข้อมูล (Data Selection) คือการคัดเลือกข้อมูลที่มีประโยชน์ต่อการใช้งาน หรือตัดข้อมูลที่ไม่ได้ส่งผลกระทบต่อผลลัพธ์ออก เป็นการเริ่มต้นของการเตรียมข้อมูล การเลือกข้อมูลนั้น จะแตกต่างกันไปตามวัตถุประสงค์ของแต่ละธุรกิจที่ได้กำหนดไว้ตั้งแต่ต้น และการเลือกข้อมูลก็ยังคงกำหนดโดยลักษณะงานที่จะถูกนำมาใช้อีกด้วย ตัวแปรที่ถูกเลือกสำหรับทำค้ำาไมน์นึ่งนั้น เรียกว่า Active Variable เพราะว่าตัวแปรนั้นจะถูกใช้สร้างความแตกต่างของกลุ่มย่อยต่างๆ และสามารถนำมาทำนายผลได้

ตัวแปรของข้อมูลที่ถูกเลือกมาแต่ละตัวนั้นจะต้องถูกทำความเข้าใจว่าหมายความว่าอะไร ประกอบด้วยอะไร และจะต้องมีคำอธิบายอย่างชัดเจนเกี่ยวกับชนิดของข้อมูล ค่าที่เป็นไปได้ แหล่งกำเนิดของข้อมูล รูปแบบของข้อมูล และลักษณะอื่น ๆ

2. การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing) เป็นกระบวนการทำงานที่เกี่ยวข้องกับข้อมูลในเบื้องต้น โดยข้อมูลที่ถูกเลือกมานั้นอาจยังไม่อยู่ในรูปแบบที่นำไปใช้ได้เลยจึงต้องมีการประมวลผลเบื้องต้น เพื่อให้ข้อมูลมีคุณภาพ ซึ่งสามารถทำได้หลายวิธีขึ้นอยู่กับข้อมูลที่เลือกมา ข้อมูลบางประเภทอาจทำแค่บางวิธี หรืออาจต้องทำทุกวิธี โดยแบ่งวิธีได้ 3 ประเภท คือ

- การแก้ไขข้อมูล (Data Cleaning) แบ่งเป็นการแก้ปัญหา 2 แบบ คือการแก้ปัญหาค่าหรือข้อมูลที่เป็นค่าว่าง (Missing Values) คือการที่ข้อมูลขาดหายไปในช่วงของข้อมูลที่นำมาใช้ และการแก้ปัญหาค่าหรือข้อมูลที่มีความคลาดเคลื่อน (Noisy Data) คือการที่ข้อมูลแตกต่างไปจากค่าข้อมูลที่เป็นไปได้ หรือที่คาดการณไว้

- การรวมข้อมูล (Data Integration) เป็นการนำเอาข้อมูลจากแหล่งข้อมูลต่างๆ ที่มีความหมายเหมือนกันมารวมกัน เพื่อให้จำนวนข้อมูลลดลงและไม่เปลืองทรัพยากรในการทำงาน

- การลดข้อมูล (Data Reduction) เป็นการลดข้อมูลเพื่อประหยัดเวลาและทรัพยากร โดยมีเทคนิคการลดข้อมูล 2 เทคนิคหลัก คือ เทคนิคการลดมิติ (Dimensionality Reduction) เป็นการลดตัวแปรที่ไม่ส่งผลกระทบต่อการทำงานของค้ำาไมน์นึ่ง และเทคนิคการลดขนาดของข้อมูล (Size Reduction) เป็นการลดขนาดของข้อมูล โดยที่ใช้วิธีการสุ่มตัวอย่าง (Sampling) เลือกเฉพาะข้อมูลที่จะนำมาใช้เป็นกลุ่มตัวอย่าง

3. การแปลงค่าข้อมูล (Data Transformation) เป็นการแปลงข้อมูลที่ไม่เหมาะสมหรือยังไม่อยู่ในรูปแบบที่จะนำไปใช้งานได้ให้อยู่ในรูปแบบที่พร้อมสำหรับการทำค้ำาไมน์นึ่ง โดยเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับขอบเขตข้อมูลให้อยู่ในช่วงที่เหมาะสมต่อการนำไปใช้งานกับแบบจำลองการวิเคราะห์ ซึ่ง จะปรับให้อยู่ในลักษณะที่ตรงกับอัลกอริทึมของค้ำไมน์นิ่ง

### 2.2.3 การไมน์นิ่งข้อมูล (Data Mining)

การไมน์นิ่งข้อมูล เป็นการเลือกอัลกอริทึมหรือแบบจำลองที่จะใช้เพื่อให้ได้ผลลัพธ์ตาม ความต้องการของวัตถุประสงค์หรือปัญหาที่ต้องการแก้ ซึ่งในแต่ละปัญหาอาจจะมีหลายเทคนิคใน การทำเพื่อให้ได้ผลลัพธ์ที่ต้องการจึงต้องมีการเลือกอย่างเหมาะสมด้วย

### 2.2.4 การวิเคราะห์ผลลัพธ์ (Analysis of Results)

การวิเคราะห์ผลลัพธ์ เป็นการวิเคราะห์ผลการประมวลผล ซึ่งจะทำการแปลความหมายและ ประเมินผลลัพธ์ที่ได้จากขั้นตอนการทำค้ำไมน์นิ่ง ว่าสามารถนำไปใช้บรรลุมิติประตามสงค์ที่ ต้องการหรือไม่ รวมทั้งเป็นการประเมินถึงความถูกต้อง เพราะบางครั้งผลที่ได้จากการทำค้ำไมน์ นิ่งอาจมีข้อผิดพลาดได้

### 2.2.5 การนำความรู้ที่ได้ไปประยุกต์ใช้ (Assimilation of Knowledge)

การนำความรู้ที่ได้ไปประยุกต์ใช้ เป็นการรวบรวมความเข้าใจทางธุรกิจ ซึ่งได้มาจาก ขั้นตอนการวิเคราะห์ผลลัพธ์ เพื่อนำมาประยุกต์ใช้กับงานต่างๆ โดยขั้นตอนนี้มีสิ่งที่จะต้องคำนึงถึง หลักๆ อยู่ 2 ประการ คือ

- การนำเสนอแนวความคิดทางธุรกิจที่ค้นพบใหม่
- ค้นหาแนวทางที่จะนำความรู้ที่ได้ค้นพบใหม่นี้ไปใช้ให้เกิดประโยชน์สูงสุด

## 2.3 วิธีการและเทคนิคที่ใช้ในการทำค้ำไมน์นิ่ง

การทำค้ำไมน์นิ่งนั้นมีเทคนิคต่างๆ มากมาย โดยในแต่ละปัญหาที่อาจจะมีการ ในการ แก้ปัญหาได้หลายวิธี ซึ่งสามารถแบ่งเทคนิคได้เป็น 4 ประเภทใหญ่ๆ คือ

### 2.3.1 การสร้างแบบจำลองพยากรณ์ (Predictive Modeling)

การสร้างแบบจำลองพยากรณ์ เป็นการนำข้อมูลที่มีอยู่เป็นจำนวนมากมาใช้ในการกำหนด ลักษณะและรูปแบบที่สำคัญของข้อมูล ซึ่งข้อมูลที่จะนำมาใช้ในการวิเคราะห์จะต้องเป็นข้อมูลที่ ถูกต้องเพื่อให้การพยากรณ์มีความแม่นยำ โดยในการสร้างแบบจำลองแบ่งออกเป็น 2 ขั้นตอน คือ

- ช่วงการเรียนรู้ (Training Phase) คือ การนำข้อมูลส่วนหนึ่งมาสอนให้ระบบเรียนรู้ เพื่อสร้างแบบจำลองขึ้นมาโดยใช้ข้อมูลในอดีต
- ช่วงการทดสอบ (Testing Phase) คือ การนำข้อมูลส่วนที่เหลือจากข้อมูลการสอน ระบบมาเป็นข้อมูลที่ใช้ทดสอบ เพื่อตรวจสอบความน่าเชื่อถือและประสิทธิภาพของแบบจำลองที่ สร้างขึ้นมาว่ามีความถูกต้องเหมาะสมหรือไม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสร้างแบบจำลองพยากรณ์ แบ่งเป็น 2 ลักษณะ ดังนี้

- การจัดกลุ่ม (Classification) เป็นการจัดกลุ่มให้กับแต่ละข้อมูลในฐานข้อมูล โดยมี ภาระมูลค่าหรือลักษณะที่เป็นไปได้ของข้อมูล ภายในแต่ละกลุ่ม
- การพยากรณ์ค่า (Prediction) เป็นการพยากรณ์ค่าที่เป็นไปได้ หรือการกระจายของค่าที่เป็นไปได้

### 2.3.2 การแบ่งกลุ่มข้อมูล (Database Segmentation)

การแบ่งกลุ่มข้อมูล เป็นการนำข้อมูลมาแบ่งเป็นกลุ่ม โดยที่ไม่รู้ล่วงหน้าว่าจะมีทั้งหมดกี่กลุ่ม ซึ่งการจัดกลุ่มได้พิจารณาจากคุณสมบัติในหลายๆ มิติของข้อมูล ถ้าข้อมูลมีลักษณะคล้ายคลึง เป็นกลุ่มเดียวกัน ได้ก็จะรวมเข้าด้วยกัน เพื่อลดขนาดข้อมูลและง่ายต่อการวิเคราะห์ และเมื่อได้แต่ละกลุ่มออกมาแล้วจะต้องแปลความหมายของแต่ละกลุ่มด้วย โดยการวิเคราะห์ถึงลักษณะที่เหมือนกันหรือความแตกต่างกันของข้อมูลในแต่ละกลุ่ม

### 2.3.3 การวิเคราะห์ความสัมพันธ์ (Link Analysis)

การวิเคราะห์ความสัมพันธ์ เป็นการค้นหาความสัมพันธ์กันของข้อมูลที่มีอยู่ เพื่อนำไปใช้ในการวิเคราะห์ หรือทำนายปรากฏการณ์ต่างๆ โดยเป็นเทคนิคที่นิยมมากในการวิเคราะห์หาความสัมพันธ์ระหว่างลูกค้ากับสินค้าหรือบริการ ซึ่งลักษณะของการหาความสัมพันธ์แบ่งได้เป็น 3 ลักษณะ ดังนี้

- การหาความสัมพันธ์ระหว่างข้อมูล (Association Discovery) เป็นการวิเคราะห์ข้อมูลที่เกิดขึ้นพร้อมกันภายในกลุ่มข้อมูลเดียวกัน
- การหาความสัมพันธ์ในลักษณะที่เป็นลำดับของข้อมูล (Sequential Pattern Discovery) เป็นการศึกษาความสัมพันธ์ระหว่างข้อมูล โดยเทียบกับเวลา ซึ่งเป็นการศึกษาพฤติกรรมในระยะยาว
- การหาความสัมพันธ์ของข้อมูลกับช่วงเวลาใดๆ (Similar Time Sequence Discovery) เป็นการศึกษาพฤติกรรมของข้อมูลที่เกิดขึ้นทั้งหมดหรือเกิดขึ้นในช่วงเวลาเดียวกัน เพื่อหาความสัมพันธ์ระหว่างกลุ่มข้อมูล

### 2.3.4 (Deviation Detection)

เป็นการตรวจหาค่าที่แตกต่างไปจากค่ามาตรฐานหรือค่าที่คาดไว้ว่าต่างไปมากน้อยเพียงใด ซึ่งมีการแสดงผลอยู่ในลักษณะที่สามารถทำความเข้าใจและแปลความหมายได้ง่าย โดยทั่วไปมักใช้วิธีการทางสถิติ (Statistic) หรือการแสดงให้เห็นภาพ (Visualization)

## 2.4 ประโยชน์ของดาต้าไมน์นึ่ง

ดาต้าไมน์นึ่งถูกนำมาใช้ในการสร้างมูลค่าเพิ่มให้กับข้อมูลที่มีอยู่ ซึ่งประโยชน์ที่ได้รับจากการทำดาต้าไมน์นึ่ง สามารถสรุปได้ดังนี้

- ทำให้เกิดความรู้ที่สามารถนำมาใช้หรือประกอบการตัดสินใจได้ เนื่องจากดาต้าไมน์นึ่งใช้เทคนิคที่ซับซ้อน และมีลักษณะเป็นปัญญาประดิษฐ์ (Artificial Intelligence) ในการสร้างแบบจำลองที่อิงกับข้อมูล ซึ่งรวบรวมจากแหล่งต่างๆ ความรู้ที่ได้จึงช่วยให้สามารถคาดการณ์อนาคตได้ และช่วยให้ตัดสินใจได้อย่างมีประสิทธิภาพมากขึ้น
- ค้นพบในสิ่งที่คาดไม่ถึง เนื่องจากกรณีที่มีตัวแบบและการตรวจสอบที่หลากหลาย ทำให้ได้ความรู้ที่เป็นประโยชน์
- ใช้ในการหาข้อผิดพลาดของการปฏิบัติงาน หรือการให้บริการได้
- ช่วยประหยัดค่าใช้จ่าย โดยการทำให้ขั้นตอนการทำงานมีประสิทธิภาพมากขึ้น

## 2.5 ตัวอย่างการประยุกต์ใช้ดาต้าไมน์นึ่ง

เนื่องจากในปัจจุบันมีการนำหลักการและเทคนิคของดาต้าไมน์นึ่งมาใช้กันอย่างแพร่หลาย ดังนั้นจึงมีการค้นคว้าวิจัย และพัฒนาเพื่อประยุกต์ใช้กับงานในหลายๆ ด้าน โดยเฉพาะทางด้านธุรกิจ ซึ่งดาต้าไมน์นึ่งสามารถสร้างคุณค่าให้กับหลายๆ ภาคธุรกิจ ตัวอย่างการประยุกต์ใช้งานที่น่าสนใจ มีดังนี้

### 2.5.1 Direct Marketing

ดาต้าไมน์นึ่งมีบทบาทสำคัญในการทำการตลาดของธุรกิจขายตรง เพราะความสามารถในการทำนายพฤติกรรมลูกค้า เพื่อพิจารณาว่าลูกค้าลักษณะใดน่าจะมีการตอบรับที่ดีในการประชาสัมพันธ์สินค้าด้วยไปรษณีย์ หรือทำนายกลุ่มลูกค้าที่มีแนวโน้มจะซื้อสินค้าชนิดหนึ่ง ซึ่งจะทำให้การประชาสัมพันธ์สินค้าประเภทหนึ่งๆ ต่อกลุ่มลูกค้าเป้าหมายได้ผลมากขึ้น นอกจากนี้ยังช่วยลดค่าใช้จ่ายจากการไม่ต้องส่งใบโฆษณาไปยังลูกค้าทุกราย แต่จะส่งเฉพาะลูกค้าที่มีแนวโน้มจะตอบรับเท่านั้น

### 2.5.2 Customer Acquisition

สำหรับการทำการตลาดของบริษัทขนาดเล็ก การดึงลูกค้าเข้ามาก็เป็นวิธีการหนึ่งที่ใช้กัน โดยดาต้าไมน์นึ่งถือได้ว่าเป็นเครื่องมือสำคัญในการจัดกลุ่มลูกค้าที่เป็นลูกค้าชั้นดีของบริษัท จัดกลุ่มของตลาด โดยในสมัยก่อนต้องอาศัยนักสถิติที่มีความชำนาญ หรือจ้างให้บริษัทที่เชี่ยวชาญเฉพาะทางทำ ซึ่งจะทำให้เกิดความล่าช้าในการเริ่มโครงการใหม่ๆ และไม่ทันต่อสภาพแวดล้อมที่มีการแข่งขันสูง อีกทั้งมีโครงการใหม่ๆ ออกสู่ตลาดได้น้อยกว่าความต้องการของธุรกิจ การนำดาต้า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม่นิ่งมาพัฒนาเพื่อเป็นเครื่องมือให้นักการตลาดผลักดันให้เกิด โครงการใหม่ๆ ที่เหมาะสมกับ กลุ่มลูกค้าได้มากขึ้น และยังทำให้การวิเคราะห์ผลตอบรับจากลูกค้าทำได้อย่างมีประสิทธิภาพ

### 2.5.3 Customer Retention

ในการทำธุรกิจ การรักษาฐานลูกค้าไว้ก็เป็นสิ่งที่สำคัญยิ่งกว่าการหาลูกค้ารายใหม่ เพราะ ลูกค้ารายเดิมมีโอกาสที่จะสร้างรายได้ให้กับบริษัทได้มากกว่า ประกอบกับต้นทุนในการหาลูกค้า รายใหม่จะสูงกว่า ซึ่งเมื่อมีการนำค้ำไม่นิ่งมาใช้จะทำให้สามารถวางกลยุทธ์การตลาด เพื่อที่จะ รักษาลูกค้าไว้ในช่วงที่ยังเป็นลูกค้าอยู่ ซึ่งจะได้ผลดีกว่าการดึงลูกค้าตอนที่ยกเลิกบริการไปแล้ว

### 2.5.4 Cross Selling

การเพิ่มมูลค่าของลูกค้าเป็นงานการตลาดที่สำคัญประการหนึ่ง การเพิ่มมูลค่าของลูกค้า แตกต่างจากการเพิ่มส่วนแบ่งทางการตลาดตรงที่การเพิ่มส่วนแบ่งตลาดได้จากการเพิ่มจำนวนลูกค้า แต่การเพิ่มมูลค่าของลูกค้าคือการเพิ่มปริมาณการซื้อสินค้าของลูกค้าแต่ละราย ซึ่งในอดีตใช้วิธีการเคา ว่าในขณะที่ลูกค้าซื้อสินค้าชนิดหนึ่งๆ ควรจะนำเสนอสินค้าใดเพิ่ม แต่เมื่อมีการนำค้ำไม่นิ่งมาใช้ ทำให้นักการตลาดทราบว่าลูกค้ารายใดน่าจะซื้อสินค้าชนิดใหม่ของบริษัทด้วย หรือควรจะ นำเสนอสินค้าชนิดใดไปพร้อมๆ กัน

### 2.5.5 Trend Analysis

การทำความเข้าใจแนวโน้มของตลาด มีความสำคัญในการลดต้นทุน และการนำสินค้าออกสู่ตลาดในช่วงที่เหมาะสม สถาบันการเงินต่างๆ ต้องการวิธีการที่จะรู้ถึงรูปแบบของการเปลี่ยนแปลงปริมาณเงินฝากและถอนอย่างรวดเร็ว ส่วนผู้ค้าปลีกก็ต้องการที่จะรู้ว่าสินค้าชนิดใดจะขายได้เมื่อวางขายด้วยกัน

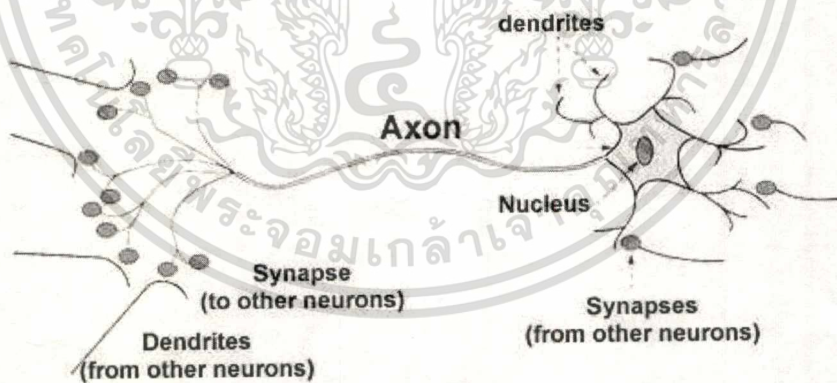
## บทที่ 3

### โครงข่ายประสาทเทียม

การสร้างแบบจำลองพยากรณ์ เป็นวิธีการหนึ่งที่ใช้ในการทำค้ำไมน์นิ่ง ในวิธีการนี้จะมีรูปแบบการทำงานที่คล้ายกับประสบการณ์การเรียนรู้ของมนุษย์ โดยจะใช้การสังเกตในการกำหนดลักษณะที่สำคัญและสิ่งที่แตกต่างกันเฉพาะตัวของสิ่งต่างๆ ซึ่งเทคนิคหนึ่งที่ใช้สำหรับการสร้างแบบจำลองพยากรณ์ก็คือ โครงข่ายประสาทเทียม (Artificial Neural Network)

#### 3.1 ความหมายของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียม เป็นโครงข่ายที่ได้พยายามเลียนแบบความสามารถในการวิเคราะห์และประมวลผลของระบบประสาทในสมองของมนุษย์ขึ้นมา โดยอาศัยการเรียนรู้จากข้อมูลที่มีอยู่ เพื่อหารูปแบบของข้อมูลเหล่านั้น แล้วใช้รูปแบบเหล่านั้นมาทำการวิเคราะห์และประมวลผล เพื่อหาผลลัพธ์ของปัญหาต่างๆ



รูปที่ 3.1 โครงสร้างเซลล์ประสาทของสมองมนุษย์

เทคนิคนี้ได้มาจากการศึกษาเครือข่ายงานไฟฟ้าชีวภาพ (bioelectric network) ในสมอง ซึ่งประกอบด้วย เซลล์ประสาทหรือนิวรอน (neurons) และจุดประสานประสาท (synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า เดนไดรต์ (Dendrite) ซึ่งเป็นอินพุต (Input) และปลายในการส่งกระแสประสาทเรียกว่า แอกซอน (Axon) ซึ่งเป็นเหมือนเอาต์พุต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

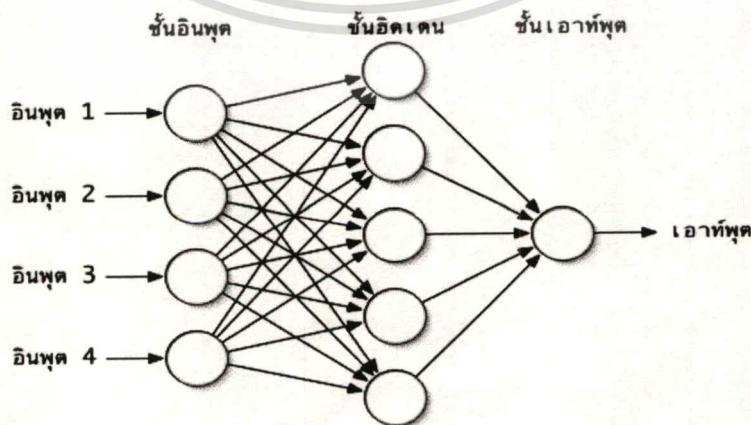
(Output) ของเซลล์ เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้า ภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรต์เข้าสู่นิวเคลียสซึ่งจะเป็นตัวตัดสินใจว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอกซอนของมัน

ดังนั้นการพยายามเลียนแบบธรรมชาติสมองในโครงข่ายประสาทเทียม จึงทำโดยสร้างระบบเครือข่ายที่ประกอบด้วยหน่วยย่อยจำนวนมาก แต่ละหน่วยแทนเซลล์ประสาทหนึ่งเซลล์ หน่วยย่อยเหล่านี้เชื่อมต่อกัน โดยมีรูปแบบการเชื่อมต่อได้หลายวิธี แล้วแต่ว่าจะเป็นแบบใด สัญญาณที่ส่งไปแทนด้วยค่าตัวเลขเป็นน้ำหนัก (Weights) นั่นคือถ้าน้ำหนักมีค่ามากก็แปลว่าสัญญาณแรงมาก

### 3.2 โครงสร้างของโครงข่ายประสาทเทียม

ภายในโครงข่ายประสาทเทียมจะพบเซลล์ประสาทเทียมจัดเรียงกันเป็นระดับชั้น ดังนี้

1. ระดับชั้นอินพุต (Input Layer) ระดับชั้นนี้จะรับสัญญาณจากภายนอก โดยไม่ได้ดำเนินการกับสัญญาณเอง เพียงแต่ทำหน้าที่ส่งสัญญาณไปยังระดับชั้นของเซลล์ประสาทที่อยู่ถัดมา จำนวนของเซลล์ประสาทในชั้นนี้จะขึ้นอยู่กับปัญหาของจำนวน และชนิดของข้อมูลที่เข้ามา
2. ระดับชั้นฮิดเดน (Hidden Layer) ระดับชั้นนี้จะไม่ได้ทำการติดต่อกับโลกภายนอก โดยตรง เป็นชั้นที่เซลล์ประสาทจะทำการคิดคำนวณกับข้อมูลที่ได้รับเข้ามาจากชั้นอินพุต จำนวนของเซลล์ประสาทชั้นนี้จะบ่งบอกถึงความซับซ้อนของการทำงานในโครงข่าย
3. ระดับชั้นเอาต์พุต (Output Layer) ระดับชั้นนี้จะส่งคืนสัญญาณผลลัพธ์ไปยังภายนอก หลังจากที่โครงข่ายได้ดำเนินการกับข้อมูลที่เข้ามาแล้ว



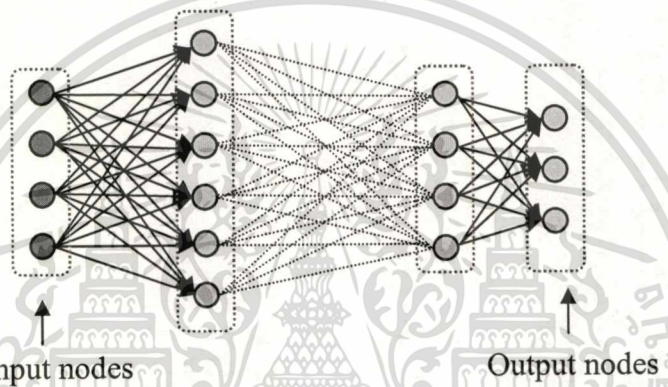
รูปที่ 3.2 โครงสร้างพื้นฐานของโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ ซึ่งได้รับการเผยแพร่เพื่อการศึกษาเท่านั้น เมื่ออนุญาตเห็นแก่ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.3 สถาปัตยกรรมของโครงข่ายประสาทเทียม

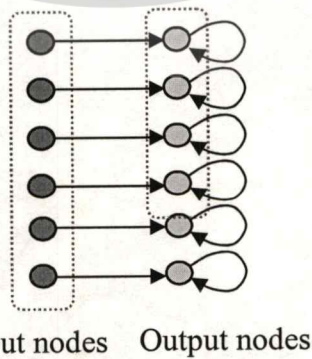
โครงข่ายประสาทเทียม ประกอบด้วยหน่วยประมวลผลย่อยๆ ที่มีการเชื่อมต่อกัน โดยลักษณะของการเชื่อมต่อระหว่างแต่ละหน่วยจากชั้นอินพุต ชั้นฮิดเดน จนถึงชั้นเอาต์พุต เพื่อทำการส่งต่อข้อมูลกัน มี 2 ลักษณะหลักๆ คือ

1. Feedforward Network ลักษณะการไหลของข้อมูลของการเชื่อมต่อแบบนี้ คือ ข้อมูลที่ประมวลผลจะถูกส่งไปในทิศทางเดียวจากชั้นอินพุตส่งต่อมาเรื่อยๆ จนถึงชั้นเอาต์พุต โดยไม่มีการย้อนกลับของข้อมูล และในระดับชั้นเดียวกันก็ไม่มีการเชื่อมต่อกัน



รูปที่ 3.3 สถาปัตยกรรมของ Feedforward Network

2. Feedback Network ลักษณะการไหลของข้อมูลของการเชื่อมต่อแบบนี้ คือ ข้อมูลที่ประมวลผลจะสามารถไหลย้อนกลับได้ หรืออาจจะมีการไหลของข้อมูลเป็นแบบ recursive ได้ จนกระทั่งได้คำตอบออกมา บางทีเรียกลักษณะแบบนี้ว่า Recurrent Network



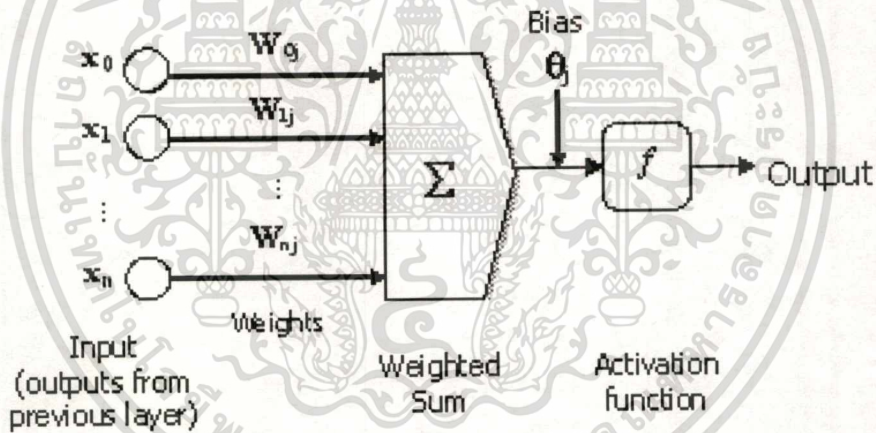
รูปที่ 3.4 สถาปัตยกรรมของ Feedback Network

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 3.4 การทำงานของเซลล์ประสาทเทียม

เซลล์ประสาทเทียมเป็นหน่วยประมวลผล ซึ่งมีการทำงาน ดังนี้

1. ชั้นอินพุตรับค่าสัญญาณเป็นข้อมูลเข้าจากภายนอก และแพร่สัญญาณไปยังชั้นฮิดเดน โดยข้อมูลเข้าจะต้องถูกทำการปรับค่าให้อยู่ในช่วงที่เหมาะสม โดยทั่วไปค่าจะอยู่ระหว่าง 0.0 ถึง 1.0
2. ทำการคูณค่าสัญญาณที่ได้รับเป็นข้อมูลเข้ากับค่าน้ำหนักของเส้นเชื่อม เพื่อนำไปคำนวณหาผลรวมสุทธิข้อมูลเข้าทั้งหมด (Total Net Input)
3. นำผลรวมสุทธิข้อมูลเข้าที่ได้มาผ่านฟังก์ชันกระตุ้น (Activation Function) เพื่อจำกัดขอบเขตของการขยายตัว โดยทั่วไปค่าที่ได้ออกมาจะอยู่ระหว่าง 0.0 ถึง 1.0 หรือ -1.0 ถึง 1.0 ผลลัพธ์ที่ได้เป็นสัญญาณข้อมูลออก ซึ่งเป็นการทำงานทั้งในชั้นฮิดเดน และชั้นเอาต์พุต
4. ชั้นเอาต์พุตส่งคืนสัญญาณผลลัพธ์ไปยังภายนอก



รูปที่ 3.5 การคำนวณของเซลล์ประสาทภายในโครงข่ายประสาทเทียม

### 3.5 การเรียนรู้ของโครงข่ายประสาทเทียม

โครงข่ายประสาทเทียมมีลักษณะการเชื่อมต่อภายในเหมือนกับเซลล์ประสาทในสมองของมนุษย์ โดยรับข้อมูลเข้าไปในหลายๆ เซลล์ แล้วนำข้อมูลเหล่านั้นมาผ่านกระบวนการเรียนรู้ เพื่อปรับเปลี่ยนข้อมูลให้ได้ตามต้องการ แล้วส่งออกไปยังเซลล์ประสาทอื่นๆ ในโครงข่าย ดังนั้นกระบวนการเรียนรู้จึงเป็นการนำค่าเอาต์พุตที่ได้จากโครงข่ายไปเปรียบเทียบกับค่าเป้าหมายที่ทราบอยู่แล้ว เพื่อหาค่าความผิดพลาดและทำการปรับปรุงค่าน้ำหนักสำหรับแต่ละการเชื่อมต่อ เพื่อให้

ประสิทธิภาพในการทำงานของโครงข่ายคีขึ้น ซึ่งสามารถแบ่งวิธีการเรียนรู้ของโครงข่ายประสาทเทียมได้ 2 แบบ คือ

1. การเรียนรู้แบบมีการชี้นำ (Supervised Learning) การเรียนรู้จะมีการกำหนดเซตของการฝึกหัดให้กับเครือข่าย ซึ่งเซตนี้จะประกอบด้วยอินพุต เอาท์พุตที่ต้องการ โดยทำการจับคู่การสอน (Training Pair) เมื่อทำการป้อนอินพุตให้กับเครือข่ายแล้ว เครือข่ายจะทำการประมวลผลจนได้ค่าเอาท์พุตและค่าน้ำหนักออกมาชุดหนึ่ง จากนั้นทำการคำนวณค่าวัดความผิดพลาด ถ้าค่าวัดความผิดพลาดยังมีค่าสูงจะต้องมีการปรับค่าน้ำหนักจนกว่าจะได้ค่าที่สามารถยอมรับได้จึงจะหยุดการสอนในเครือข่าย

2. การเรียนรู้แบบไม่มีการชี้นำ (Unsupervised Learning) การเรียนรู้การสอนโดยที่ไม่ต้องมีการจับคู่สอน เนื่องจากเมื่อใส่ค่าอินพุตชุดสู่เครือข่ายแล้ว เครือข่ายจะพยายามจัดกลุ่มอินพุตเซตที่มีลักษณะเดียวกันให้เอาท์พุตออกมาจากเครือข่ายที่เดียวกัน คือมีการปรับตัวเองภายใน (Self-Organization) การเรียนรู้โดยวิธีนี้จะป้อนอินพุตเข้าสู่เครือข่าย ภายในเครือข่ายจะมีเอาท์พุตโหนดอยู่หลายโหนด โดยแต่ละโหนดแทนกลุ่มของข้อมูลที่มีคุณสมบัติเหมือนกัน เมื่อป้อนอินพุตเข้าสู่เครือข่าย เครือข่ายจะทำการคำนวณหาความสัมพันธ์ที่มีภายในเซตของอินพุต โดยอาศัยค่าน้ำหนักเป็นตัวแยกความแตกต่างของอินพุตไปไว้ยังโหนดเอาท์พุตของเครือข่าย การเรียนรู้ คือใช้ค่าน้ำหนักเป็นตัวแยกความแตกต่างของอินพุต โดยวิธีนี้ไม่สามารถทำการระบุเอาท์พุตที่ถูกต้องได้ว่าโหนดใดเป็นข้อมูลของกลุ่มใด ซึ่งผู้ใช้งานต้องเป็นคนกำหนดเอง

### 3.5.1 การเรียนรู้แบบแบคพรอพพาเกชัน

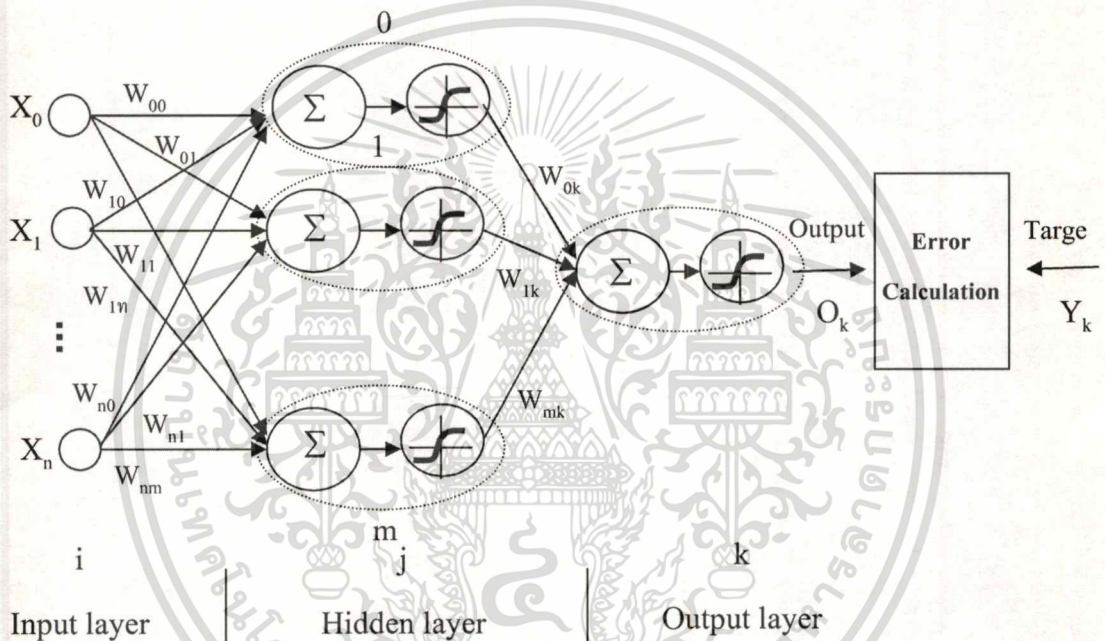
แบคพรอพพาเกชันเป็นอัลกอริทึมหนึ่งของโครงข่ายประสาทเทียม โดยมีลักษณะการเชื่อมต่อแบบ Feedforward Network และมีการเรียนรู้แบบมีการชี้นำ ซึ่งจะเรียนรู้จากการประมวลผลกลุ่มตัวอย่างที่เป็นข้อมูลในอดีต โดยที่ค่าเอาท์พุตที่ได้จากโครงข่ายในแต่ละครั้งของชุดข้อมูลที่ป้อนให้กับโครงข่ายนั้นจะถูกนำมาเปรียบเทียบกับค่าเอาท์พุตที่เกิดขึ้นจริงในอดีต เพื่อหาค่าความผิดพลาดในแต่ละครั้ง แล้วนำค่าความผิดพลาดที่ได้นั้นมาใช้ในการคำนวณเพื่อปรับค่าน้ำหนักระหว่างแต่ละเซลล์ในโครงข่ายต่อไป โดยที่การปรับค่าน้ำหนักจะมีการคำนวณในทิศทางย้อนกลับจากทางชั้นเอาท์พุตผ่านชั้นฮิดเดน จนปรับค่ามาถึงชั้นแรกที่ต่อจากชั้นอินพุต จากนั้นจึงคำนวณค่าเอาท์พุตใหม่ที่ได้ เพื่อเปรียบเทียบกับค่าวัดความผิดพลาดใหม่ โดยในกระบวนการเรียนรู้จะมีการคำนวณเช่นนี้เรื่อยไป จนกระทั่งค่าวัดความผิดพลาดอยู่ในระดับที่สามารถยอมรับได้

กระบวนการในการเรียนรู้ของโครงข่ายแบบแบคพรอพพาเกชันจะขึ้นกับค่าน้ำหนัก และฟังก์ชันกระตุ้น ซึ่งเป็นฟังก์ชันกระตุ้นการทำงานทั้งในระดับชั้นฮิดเดนและชั้นเอาท์พุต โดย

สามารถมีได้หลายรูปแบบ เช่น Threshold Function, Hard Limit, Hyperbolic Tan และ Sigmoid Function ซึ่งกระบวนการเรียนรู้มีขั้นตอนการทำงาน ดังนี้

1. เลือกอินพุตและเอาต์พุต เพื่อนำมาใช้ในการเรียนรู้ของโครงข่าย โดยที่รูปแบบจะอยู่ในรูปของ Vector Pairs  $\{(I_1, Y_1), (I_2, Y_2), \dots, (I_k, Y_k)\}$  ซึ่งแต่ละ  $I$  คือ อินพุตเวกเตอร์ และแต่ละ  $Y$  คือ เอาต์พุตเวกเตอร์ โดย  $I = (x_1, x_2, \dots, x_n; Y)$

2. สุ่มค่าน้ำหนักในแต่ละเส้นทางของชั้นฮิดเดน และชั้นเอาต์พุต โดยที่ช่วงของค่าน้ำหนักอยู่ระหว่าง 0.0 ถึง 1.0 รวมทั้งกำหนดค่าอัตราการเรียนรู้ (learning rate) ซึ่งมีค่าระหว่าง 0.0 ถึง 1.0



รูปที่ 3.6 กระบวนการเรียนรู้แบบแบคพรอพพาเกชัน

3. คำนวณหาผลรวมสุทธิข้อมูลเข้าที่ชั้นฮิดเดน ( $net_j$ )

$$net_j = \sum W_{ij} x_i + \theta_j$$

$W_{ij}$  คือ ค่าน้ำหนักที่ชั้นฮิดเดน จากเซลล์  $i$  ไปยังเซลล์  $j$

$\theta_j$  คือ ค่าเทอมไบแอส (bias) ที่ชั้นฮิดเดน

$x_i$  คือ ค่าอินพุตของชั้นฮิดเดน

4. คำนวณหาค่าเอาต์พุตที่ชั้นฮิดเดน ซึ่งจะเป็นค่าอินพุตของชั้นเอาต์พุต โดยหาจากฟังก์ชันกระตุ้น ในที่นี้ใช้ Sigmoid Function

$$O_j = (1 + e^{-net_j})^{-1}$$

$O_j$  คือ ค่าเอาต์พุตของชั้นฮิดเดน และเป็นค่าอินพุตของชั้นเอาต์พุต

5. คำนวณหาผลรวมสุทธิข้อมูลเข้าที่ชั้นเอาต์พุต ( $net_k$ )

$$net_k = \sum W_{jk} O_j + \theta_k$$

$W_{jk}$  คือ ค่าน้ำหนักที่ชั้นเอาต์พุตจากเซลล์  $j$  ไปเซลล์  $k$

$\theta_k$  คือ ค่าเทอมไบแอส (bias) ที่ชั้นเอาต์พุต

6. คำนวณหาค่าเอาต์พุตที่ชั้นเอาต์พุต โดยใช้ Sigmoid Function

$$O_k = (1 + e^{-net_k})^{-1}$$

$O_k$  คือ ค่าเอาต์พุตของชั้นเอาต์พุต

7. คำนวณหาค่าความผิดพลาดที่เกิดขึ้นที่ชั้นเอาต์พุต ( $Err_k$ )

$$Err_k = O_k(1 - O_k)(Y_k - O_k)$$

$Y_k$  คือ ค่าเอาต์พุตจริง

$O_k$  คือ ค่าเอาต์พุตจากการคำนวณ

8. คำนวณหาค่าความผิดพลาดที่เกิดขึ้นที่ชั้นฮิดเดน ( $Err_j$ )

$$Err_j = O_j(1 - O_j) \sum Err_k W_{jk}$$

$\sum Err_k W_{jk}$  คือ ผลรวมของค่าความผิดพลาดที่ชั้นเอาต์พุตคูณกับค่าน้ำหนักที่อยู่บนชั้น

นั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ห้องสมุดคณะเทคโนโลยีการเกษตร สงข.

### 9. คำวนหาการปรับแต่งค่าน้ำหนักที่ชั้นเอาท์พุท

$$\Delta W_{jk} = (I) \text{Err}_k O_j$$

$$W_{jk}(t+1) = W_{jk}(t) + \Delta W_{jk}$$

$W_{jk}(t+1)$  คือ ค่าน้ำหนักใหม่ที่ถูกปรับค่าที่ชั้นเอาท์พุท

$W_{jk}(t)$  คือ ค่าน้ำหนักเดิมที่ชั้นเอาท์พุท

$I$  คือ อัตราการเรียนรู้

$O_j$  คือ ค่าอินพุทของชั้นเอาท์พุท

และคำวนหาการปรับแต่งค่าเทอมไบแอสที่ชั้นเอาท์พุท

$$\Delta \theta_k = (I) \text{Err}_k$$

$$\theta_k(t+1) = \theta_k(t) + \Delta \theta_k$$

$\theta_k(t+1)$  คือ ค่าเทอมไบแอสใหม่ที่ถูกปรับค่าที่ชั้นเอาท์พุท

$\theta_k(t)$  คือ ค่าเทอมไบแอสเดิมที่ชั้นเอาท์พุท

### 10. คำวนหาการปรับแต่งค่าน้ำหนักที่ชั้นฮิดเดน

$$\Delta W_{ij} = (I) \text{Err}_j x_i$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}$$

$W_{ij}(t+1)$  คือ ค่าน้ำหนักใหม่ที่ถูกปรับค่าที่ชั้นฮิดเดน

$W_{ij}(t)$  คือ ค่าน้ำหนักเดิมที่ชั้นฮิดเดน

และคำวนหาการปรับแต่งค่าเทอมไบแอสที่ชั้นฮิดเดน

$$\Delta \theta_j = (I) \text{Err}_j$$

$$\theta_j(t+1) = \theta_j(t) + \Delta \theta_j$$

$\theta_j(t+1)$  คือ ค่าเทอมไบแอสใหม่ที่ถูกปรับค่าที่ชั้นฮิดเดน

$\theta_j(t)$  คือ ค่าเทอมไบแอสเดิมที่ชั้นฮิดเดน

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์สำหรับการใช้ในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

11. วงจรจะถูกทำซ้ำที่ขั้นที่ 3 โดยพิจารณารูปแบบการเรียนรู้ใหม่ จนกระทั่งรูปแบบทั้งหมดได้ถูกเรียนรู้โดยโครงข่ายหมดแล้ว

12. จากนั้นคำนวณหาค่าวัดความผิดพลาด โดยการใช้ Sum Square Error

$$SSE = \sum (Y_k - O_k)^2$$

13. ขั้นที่ 3 ถึง 12 จะถูกทำซ้ำ จนกระทั่งค่าวัดความผิดพลาดน้อยกว่าระดับที่กำหนดไว้ โครงข่ายก็จะหยุดการเรียนรู้

### 3.5.2 ข้อได้เปรียบและข้อจำกัดของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

#### 1. ข้อได้เปรียบของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

- สามารถเรียนรู้รูปแบบความสัมพันธ์ของข้อมูลได้หลายๆ รูปแบบ
- มีความสามารถในการจดจำรูปแบบของปัญหาได้มาก (Pattern Mapping) โดยที่การเรียนรู้ต้องการรูปแบบตัวอย่างที่จะเรียนรู้
- มีความยืดหยุ่นในการเรียนรู้ โดยสามารถเลือกทางเลือกต่างๆ เช่น สามารถทำการเลือกจำนวนชั้นของระดับชั้นฮิดเดน สามารถกำหนดค่าสัมประสิทธิ์การเรียนรู้ได้ และการแทนรูปแบบของข้อมูล ทำให้สามารถแก้ปัญหาต่าง ๆ ได้

- สามารถนำมาสร้างระบบการตัดสินใจที่ซับซ้อนได้

#### 2. ข้อจำกัดของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

- ต้องใช้เวลามากในการสอนโครงข่ายให้เกิดการเรียนรู้ (Convergence time) ตัวอย่างเช่น สำหรับระบบงานจริงอาจมีชุดข้อมูลตัวอย่าง 1,000 ชุด หรือบางกรณีอาจมีชุดข้อมูลตัวอย่างมากกว่านี้ ซึ่งอาจทำให้ใช้เวลาในการประมวลผลนานเป็นวันหรือมากกว่านั้น เพื่อให้ได้การเรียนรู้ที่สมบูรณ์

- มีโอกาสเกิดความล้มเหลวในการเรียนรู้ของโครงข่าย แม้จะมีการเพิ่มข้อมูลเข้าไป ก็ไม่สามารถทำการปรับโครงข่ายได้ แต่อาจทำการแก้ไขได้โดยการเพิ่มจำนวนชั้น หรือทำการเปลี่ยนชุดของข้อมูลที่นำมาใช้ในการเรียนรู้

- คุณภาพของโครงข่ายขึ้นอยู่กับคุณภาพของข้อมูล โดยต้องพิจารณาถึงข้อมูลว่าเป็นข้อมูลที่ดีหรือไม่ มีข้อมูลที่ครอบคลุมขอบเขตของปัญหาทั้งหมดหรือไม่ เพราะหาข้อมูลที่นำมาใช้ในการเรียนรู้เป็นข้อมูลที่มีคุณภาพไม่ดีพอ ไม่ครอบคลุมกับปัญหาทั้งหมด ก็จะทำให้คุณภาพของโครงข่ายไม่ดีตามไปด้วย

## บทที่ 4

### การวิเคราะห์และออกแบบระบบ

ในการวิเคราะห์และออกแบบระบบค้าไม้หนึ่งโดยใช้โครงข่ายประสาทเทียม สามารถแบ่งการทำงานของระบบได้เป็น 2 ส่วน คือ การสร้างแบบจำลองใหม่ให้กับระบบ และการเลือกแบบจำลองที่มีอยู่มาใช้งานเพื่อการพยากรณ์ข้อมูล ซึ่งในบทนี้จะอธิบายถึงรายละเอียดของระบบการจำลองระบบงานด้วย UML และรายละเอียดเกี่ยวกับตารางที่ใช้ในระบบ

#### 4.1 รายละเอียดของระบบ

ระบบค้าไม้หนึ่งโดยใช้โครงข่ายประสาทเทียม สามารถแบ่งการทำงานของระบบได้เป็น 2 ส่วนหลักๆ คือ

##### 4.1.1 ส่วนการสร้างแบบจำลองใหม่ (Build Model)

ในการสร้างแบบจำลองใหม่ เป็นการสร้างแบบจำลองโดยใช้โครงข่ายประสาทเทียม ซึ่งมีกระบวนการหลักๆ ดังนี้

1. ส่วนการเชื่อมต่อฐานข้อมูล (Connection Database) เป็นการติดต่อฐานข้อมูล SQL Server 2000 ซึ่งเป็นฐานข้อมูลที่ใช้ในระบบ

2. ส่วนการเลือกข้อมูล (Data Selection) เป็นการเลือกชุดข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลองให้กับระบบ

ซึ่งหลังจากเลือกข้อมูลเสร็จแล้วนั้น ระบบจะทำการสร้างตารางขึ้นมาใหม่ โดยทำการคัดลอกข้อมูลตามที่ใช้เลือกจากตารางต้นฉบับ เพื่อใช้ตารางนี้ในการทำกระบวนการขั้นต่างๆต่อไป โดยที่ตารางต้นฉบับยังคงเหมือนเดิม

3. ส่วนการประมวลผลข้อมูลเบื้องต้น เป็นการตรวจสอบข้อมูลที่นำมาใช้สร้างแบบจำลอง ซึ่งในระบบนี้ใช้วิธีการแก้ไขข้อมูล (Data Cleaning) โดยการจัดการกับข้อมูลที่เป็นค่าว่าง (Missing Value) ซึ่งสามารถแบ่งข้อมูลเป็น 2 ประเภท คือ

- ข้อมูลที่ไม่ได้เป็นตัวเลข (Categorical Data) มีวิธีการจัดการ 2 รูปแบบ คือ

1. การลบเรคคอร์ดที่มีค่าว่าง
2. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ข้อมูลที่เป็นตัวเลข (Numeric Data) มีวิธีการจัดการ 5 รูปแบบ คือ

1. การลบเรคคอร์ดที่มีค่าว่าง
2. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด
3. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่ามากที่สุด
4. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่าน้อยที่สุด
5. การแทนค่าที่ว่าง ด้วยข้อมูลที่เป็นค่าเฉลี่ย

4. ส่วนการแปลงค่าข้อมูล (Data Transformation) เป็นการแปลงค่าข้อมูลที่จะนำมาใช้ในแบบจำลองโครงข่ายประสาทเทียม ให้อยู่ในขอบเขตที่เหมาะสมตามลักษณะของฟังก์ชันกระตุ้นที่นำมาใช้ในแบบจำลอง สำหรับในระบบนี้จะใช้ฟังก์ชันของซิกมอยด์ (Sigmoid Function) ในการคำนวณ ดังนั้นจึงต้องทำการแปลงค่าของข้อมูลให้อยู่ในช่วง 0 ถึง 1

5. ส่วนการสร้างแบบจำลอง (Data Mining) เป็นการใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน ในการสอนให้ระบบเกิดการเรียนรู้และมีการสร้างแบบจำลองขึ้น โดยสามารถกำหนดตัวแปรบางตัวที่ใช้ควบคุมการเรียนรู้ได้ เช่น จำนวนโหนดในชั้นอินพุต (Numbers of Input) จำนวนโหนดในชั้นฮิดเดน (Numbers of Hidden) จำนวนโหนดในชั้นเอาต์พุต (Number of Output) จำนวนรอบของการเรียนรู้ และค่าความผิดพลาดที่ยอมรับได้ เพื่อช่วยให้ระบบมีการเรียนรู้ที่มีประสิทธิภาพมากยิ่งขึ้น

สำหรับการสร้างแบบจำลองของระบบนั้นประกอบด้วยขั้นตอน ดังต่อไปนี้

**ขั้นตอนที่ 1** การสอนให้ระบบเกิดการเรียนรู้ (Training) เป็นการนำข้อมูลแต่ละชุดที่ใช้สอนระบบมาทำการคำนวณตามอัลกอริทึมแบบแบคพรอพพาเกชัน เพื่อปรับตัวแปรภายในโครงข่าย จนกระทั่งได้แบบจำลองที่เหมาะสม โดยระบบจะใช้ค่าความผิดพลาด (Sum Square Error) ตามที่ผู้ใช้กำหนดเป็นเงื่อนไขในการจบการเรียนรู้ของแบบจำลอง โดยที่จำนวนรอบในการเรียนรู้ยังไม่เกินค่าที่กำหนดไว้ หรืออาจจบการเรียนรู้โดยพิจารณาจากจำนวนรอบของการเรียนรู้

**ขั้นตอนที่ 2** การทดสอบแบบจำลองที่ได้จากการเรียนรู้ (Testing) เป็นการประเมินแบบจำลอง โดยใช้ข้อมูลส่วนที่เก็บไว้เพื่อทดสอบ ซึ่งในการทำงานขั้นตอนนี้จะเป็นการนำข้อมูลเข้าสู่แบบจำลองที่ได้จากการเรียนรู้ แล้วทำการคำนวณหาเอาต์พุตของโครงข่ายเพื่อเปรียบเทียบกับค่าจริง แล้วหาค่าความผิดพลาดเพื่อประเมินประสิทธิภาพของแบบจำลองที่ได้

หลังจากการทดสอบแล้ว ถ้าประสิทธิภาพของแบบจำลองเป็นที่น่าพอใจ ผู้ใช้สามารถบันทึกแบบจำลองนี้เก็บไว้เพื่อนำไปใช้ในการพยากรณ์ข้อมูลต่อไป โดยในการบันทึกนั้นจะประกอบด้วยแบบจำลองที่สร้างโดยใช้โครงข่ายประสาทเทียม และ โครงสร้างของข้อมูลที่ใช้สร้างแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.1.2 ส่วนการเรียกใช้งานแบบจำลองที่มีอยู่เพื่อพยากรณ์ข้อมูล

การเรียกใช้แบบจำลองที่มีอยู่มาใช้งานในการพยากรณ์ข้อมูล มีกระบวนการหลักๆ ดังนี้

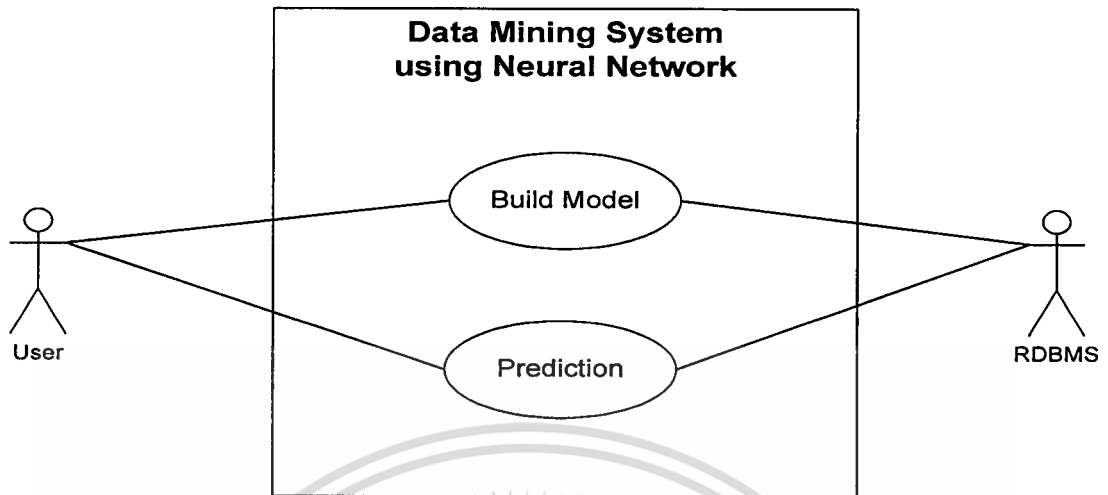
1. ส่วนการเลือกแบบจำลอง เป็นการเลือกแบบจำลองที่มีอยู่ โดยเลือกตามความเหมาะสมและสอดคล้องกันกับข้อมูลที่จะนำมาพยากรณ์
2. ส่วนการรับข้อมูล เป็นการรับข้อมูลที่นำมาใช้ในการพยากรณ์ โดยที่จำนวนคอลัมน์ ต้องเป็นไปตามจำนวนของข้อมูลในชั้นอินพุตที่ใช้สร้างแบบจำลองให้กับระบบ ซึ่งหลังจากรับข้อมูลเสร็จแล้วนั้น ระบบจะทำการสร้างตารางขึ้นมาใหม่โดยทำการคัดลอกข้อมูลตามที่ใช้เลือกจากตารางต้นฉบับ เพื่อใช้ตารางนี้ในการทำกระบวนการขั้นต่างๆ ต่อไป โดยที่ตารางต้นฉบับยังคงเหมือนเดิม
3. ส่วนการประมวลผลข้อมูลเบื้องต้น เป็นการตรวจสอบข้อมูลที่นำมาใช้ในการพยากรณ์ ซึ่งในระบบนี้ใช้วิธีการแก้ไขข้อมูล โดยการจัดการกับข้อมูลที่เป็นค่าว่าง ซึ่งมีวิธีการจัดการด้วยการลบเรคคอร์ดที่มีค่าว่าง
4. ส่วนการแปลงค่าข้อมูล เป็นการแปลงค่าข้อมูลที่นำมาใช้ในการพยากรณ์ ให้อยู่ในขอบเขตที่เหมาะสมตามลักษณะฟังก์ชันกระตุ้นของแบบจำลองที่เลือกมาใช้พยากรณ์
5. ส่วนการประมวลผล เป็นการคำนวณหาค่าเอาต์พุตโดยใช้ค่าตัวแปรของแบบจำลองคำนวณร่วมกับค่าอินพุตที่ได้จากส่วนการรับข้อมูล โดยได้ผ่านส่วนการประมวลผลข้อมูลเบื้องต้น และส่วนการแปลงค่าข้อมูลเรียบร้อยแล้ว

#### 4.2 การจำลองระบบงานด้วย UML

สำหรับระบบคาด้าไมน์นิ่ง โดยใช้โครงข่ายประสาทเทียม ได้ใช้ยูสเคสไดอะแกรม (Use case Diagram) และซีควเอนซ์ไดอะแกรม (Sequence Diagram) ในการจำลองระบบงาน

##### 4.2.1 ยูสเคสไดอะแกรม

ยูสเคสไดอะแกรมของระบบคาด้าไมน์นิ่ง โดยใช้โครงข่ายประสาทเทียม ประกอบด้วย 2 ยูสเคส แสดงดังรูปที่ 4.1



รูปที่ 4.1 ยูสเคสไดอะแกรมของระบบการค้าไม้หนึ่งโดยใช้โครงข่ายประสาทเทียม

ตารางที่ 4.1 คำอธิบายยูสเคสไดอะแกรมของ Build Model

ยูสเคส	Build Model
วัตถุประสงค์	เพื่อสร้างแบบจำลองใหม่ให้กับระบบ
เมื่อทำงานสำเร็จ	ได้แบบจำลองใหม่ให้กับระบบ ซึ่งแบบจำลองใหม่ที่ได้ประกอบด้วยแบบจำลองที่สร้างโดยใช้โครงข่ายประสาทเทียม และ โครงสร้างของข้อมูลที่ใช้สร้างแบบจำลอง
Actor ที่เกี่ยวข้อง	User
อินพุต	ชุดข้อมูล และเงื่อนไขในการสร้างแบบจำลอง
เอาต์พุต	แบบจำลองพร้อมกับ โครงสร้างของข้อมูลที่ใช้สร้างแบบจำลอง, ผลลัพธ์ของการสร้างแบบจำลองในรูปแบบของกราฟและรายงาน, ผลลัพธ์ของการทดสอบแบบจำลองในรูปแบบของกราฟและรายงาน
รายละเอียด	<ol style="list-style-type: none"> <li>1. ผู้ใช้เลือกชุดข้อมูล และเงื่อนไขในการสร้างแบบจำลองเข้าสู่ระบบ</li> <li>2. ระบบจะทำการตรวจสอบข้อมูล และเงื่อนไขที่ใช้ในการสร้างแบบจำลอง จากนั้นทำการสร้างแบบจำลองตามเงื่อนไข</li> <li>3. ระบบทำการแสดงผลลัพธ์ในการสร้างแบบจำลอง</li> </ol>

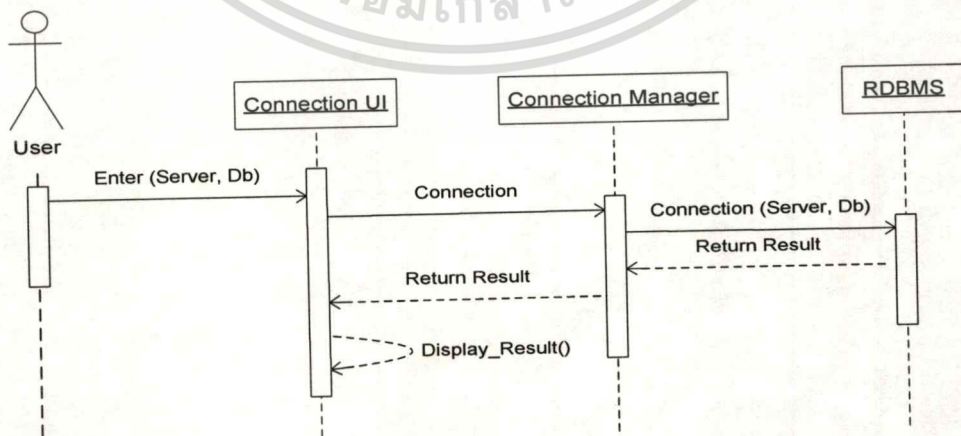
ตารางที่ 4.2 คำอธิบายยูสเคสโคอะเกมของ Prediction

ยูสเคส	Prediction
วัตถุประสงค์	เพื่อพยากรณ์ข้อมูลโดยใช้แบบจำลองที่มีอยู่
เมื่อทำงานสำเร็จ	ผลลัพธ์ในการพยากรณ์ข้อมูล
Actor ที่เกี่ยวข้อง	User
อินพุต	ชุดข้อมูล และแบบจำลองที่ใช้ในการพยากรณ์ข้อมูล
เอาต์พุต	ผลลัพธ์ที่ได้จากการพยากรณ์ข้อมูล
รายละเอียด	<ol style="list-style-type: none"> <li>1. ผู้ใช้เลือกแบบจำลอง และชุดข้อมูลที่ใช้ในการพยากรณ์ข้อมูล</li> <li>2. ระบบทำการตรวจสอบโครงสร้างข้อมูลของชุดข้อมูลที่ใช้พยากรณ์กับโครงสร้างข้อมูลที่ใช้สร้างแบบจำลอง ถ้าสอดคล้องกันระบบจะทำการพยากรณ์ข้อมูลตามเงื่อนไขของแบบจำลองที่เลือก</li> <li>3. ระบบทำการแสดงผลลัพธ์ของการพยากรณ์ข้อมูล</li> </ol>

#### 4.2.2 ซีเควนซ์โคอะเกม

จากการออกแบบแผนภาพยูสเคสสามารถนำมาอธิบายขั้นตอนการปฏิบัติงานหรือขั้นตอนการติดต่อกับระบบในแต่ละงานได้โดยใช้ซีเควนซ์โคอะเกม ซึ่งในโคอะเกมนี้สามารถอธิบายการทำงานเป็นขั้นตอนเพิ่มเติมได้ ดังต่อไปนี้

##### 1. Build Model Use Case



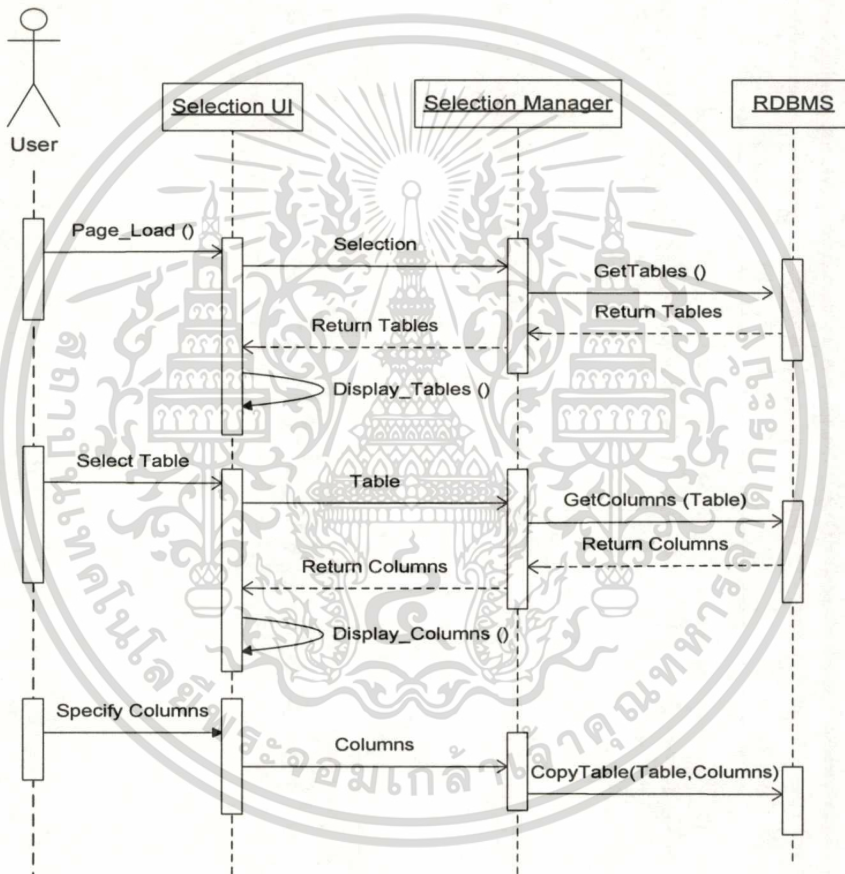
เอกสารนี้เป็นเอกสารที่สงวนไว้รูปที่ 4.2 ซีเควนซ์โคอะเกมของการเชื่อมต่อฐานข้อมูล  
 ไม่ว่าการณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Actor: User

Description: ผู้ใช้ทำการเชื่อมต่อฐานข้อมูล

Basic Flow:

1. ผู้ใช้ป้อนชื่อเซิร์ฟเวอร์และชื่อฐานข้อมูลที่ต้องการเชื่อมต่อ
2. ระบบทำการเชื่อมต่อฐานข้อมูลผ่าน RDBMS
3. ระบบแสดงผลการเชื่อมต่อฐานข้อมูล



รูปที่ 4.3 ซีควেনซ์ไดอะแกรมของการเลือกข้อมูล

Actor: User

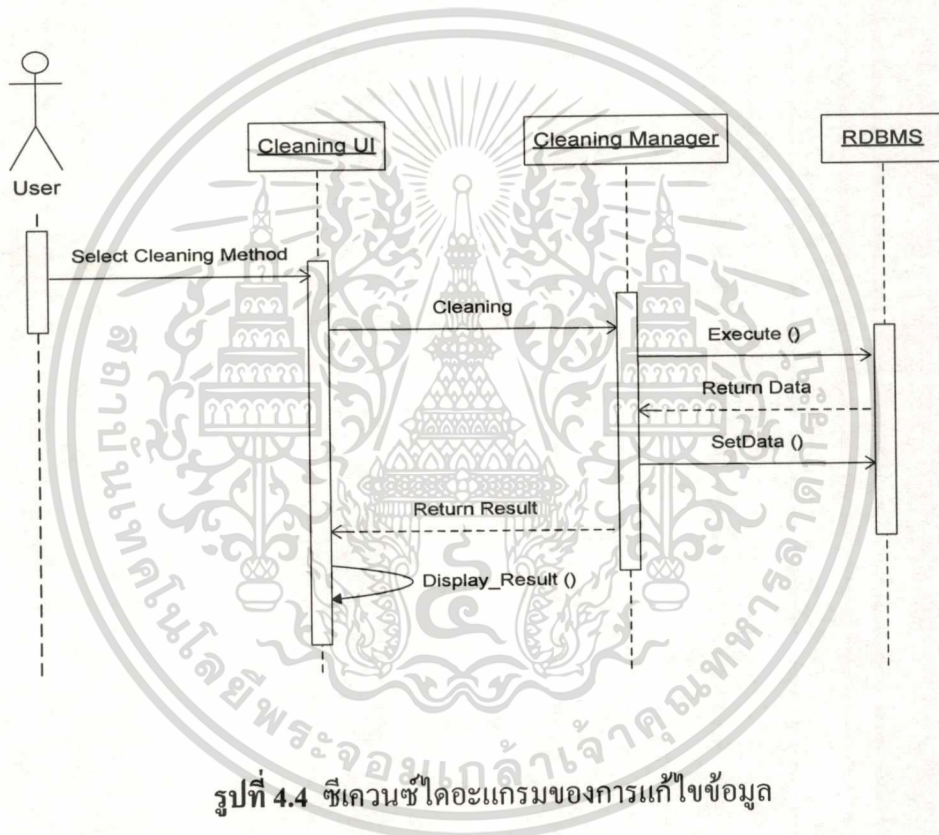
Description: ผู้ใช้ทำการเลือกข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลอง

Basic Flow:

1. ผู้ใช้ต้องการทำการเลือกข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ระบบแสดงตารางที่อยู่ในฐานข้อมูลที่เลือกทั้งหมด
3. ผู้ใช้เลือกตาราง
4. ระบบแสดงคอลัมน์ที่อยู่ในตารางที่เลือก
5. ผู้ใช้ระบุคอลัมน์ที่จะใช้ในการสร้างแบบจำลอง
6. ระบบสร้างตารางใหม่โดยการคัดลอกข้อมูลตามทีผู้ใช้เลือกจากตารางต้นฉบับ



รูปที่ 4.4 ซีควเอนซ์ไดอะแกรมของการแก้ไขข้อมูล

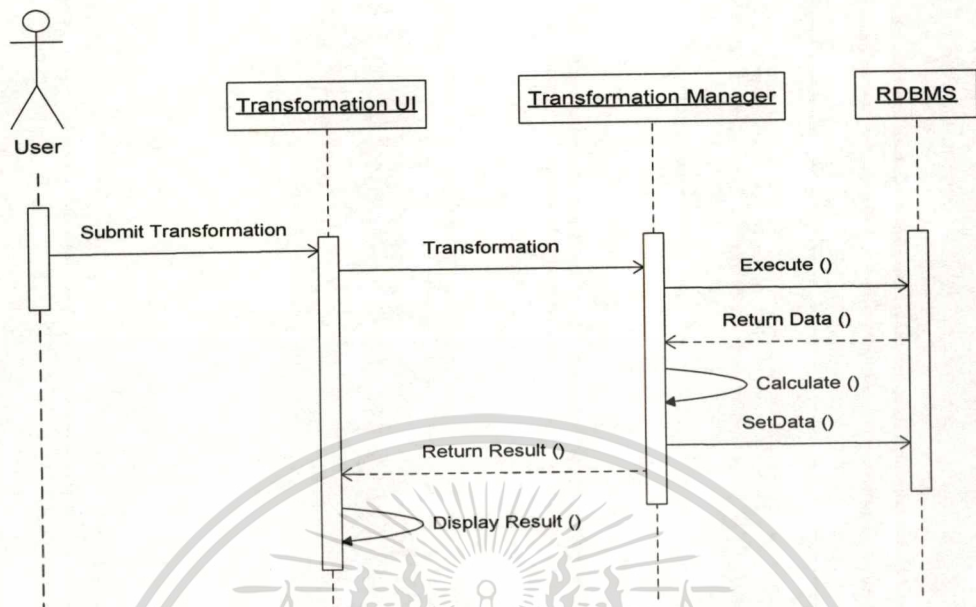
Actor: User

Description: ผู้ใช้ทำการแก้ไขข้อมูล ด้วยการเลือกวิธีการแก้ไขข้อมูล

Basic Flow:

1. ผู้ใช้เลือกวิธีการแก้ไขข้อมูล
2. ระบบจัดการกับข้อมูลที่เป็นค่าว่างตามวิธีที่ผู้ใช้เลือก โดยทำการลบหรือแก้ไขข้อมูลในเรคคอร์ดที่มีค่าว่าง
3. ระบบแสดงผลการแก้ไขข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



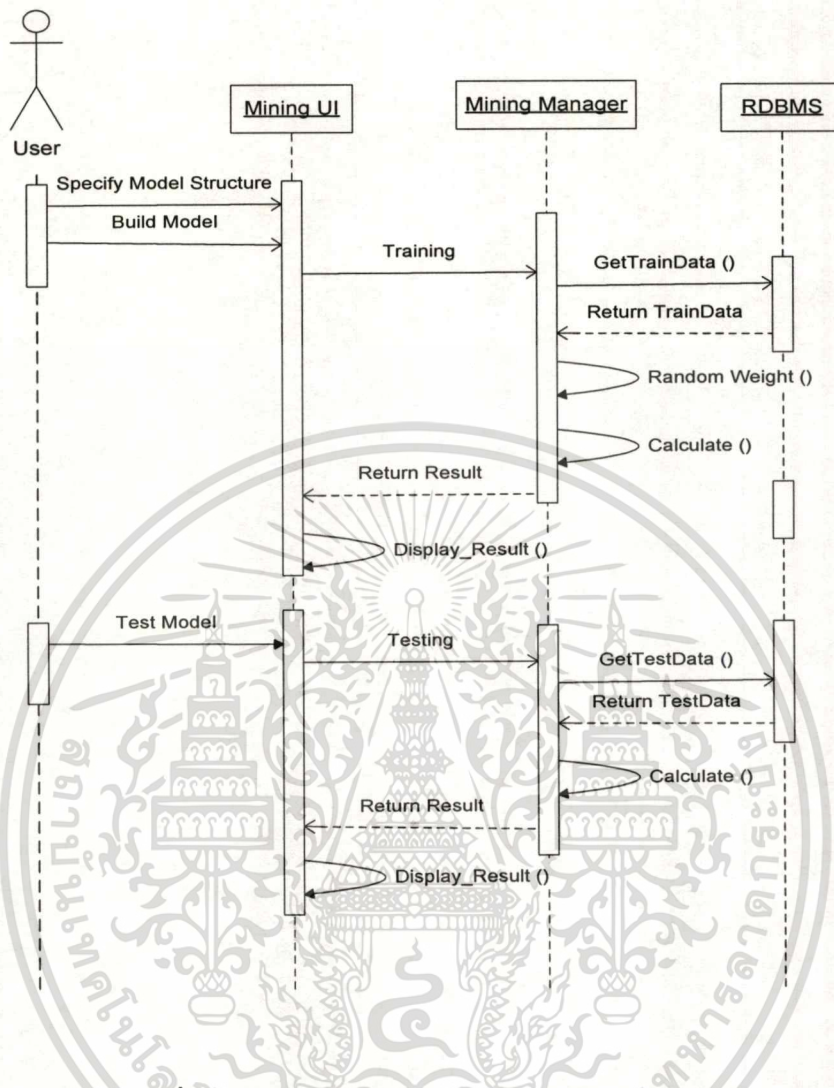
รูปที่ 4.5 ซีเควนซ์ไดอะแกรมของการแปลงค่าข้อมูล

Actor: User

Description: ผู้ใช้ทำการแปลงค่าข้อมูลให้อยู่ในช่วงที่เหมาะสม

Basic Flow:

1. ผู้ใช้ยืนยันการแปลงค่าข้อมูล
2. ระบบทำการคำนวณค่าให้อยู่ในช่วงที่เหมาะสมตามสูตร
3. ระบบทำการแปลงค่าข้อมูล โดยทำการแก้ไขข้อมูลในทุกเรคคอร์ด
4. ระบบแสดงผลการแปลงค่าข้อมูล



รูปที่ 4.6 ซีเควนซ์ไดอะแกรมของการสร้างแบบจำลอง

Actor: User

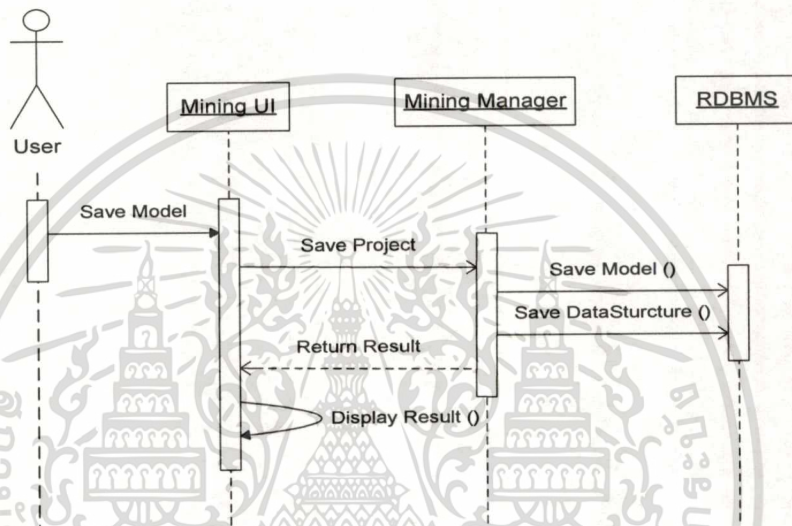
Description: ผู้ใช้ทำการสร้างแบบจำลอง

Basic Flow:

1. ผู้ใช้กำหนดโครงสร้างของแบบจำลอง
2. ระบบทำการสร้างแบบจำลอง โดยดึงข้อมูลสำหรับการเรียนรู้จากตารางที่คัดลอกไว้ และทำการสร้างค่าน้ำหนักขึ้นมาด้วยการสุ่ม แล้วคำนวณตามอัลกอริทึม
3. เมื่อทำการเรียนรู้เสร็จสมบูรณ์แล้ว ระบบทำการปรับค่าน้ำหนักตามที่ได้จากการเรียนรู้ จากนั้นแสดงผลของการสร้างแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. ทำการทดสอบแบบจำลอง โดยดึงข้อมูลสำหรับการทดสอบจากตารางที่คัดลอกไว้
5. ระบบทำการคำนวณค่า จากข้อมูลและค่านำหนัก
6. ระบบแสดงผลของการทดสอบแบบจำลอง



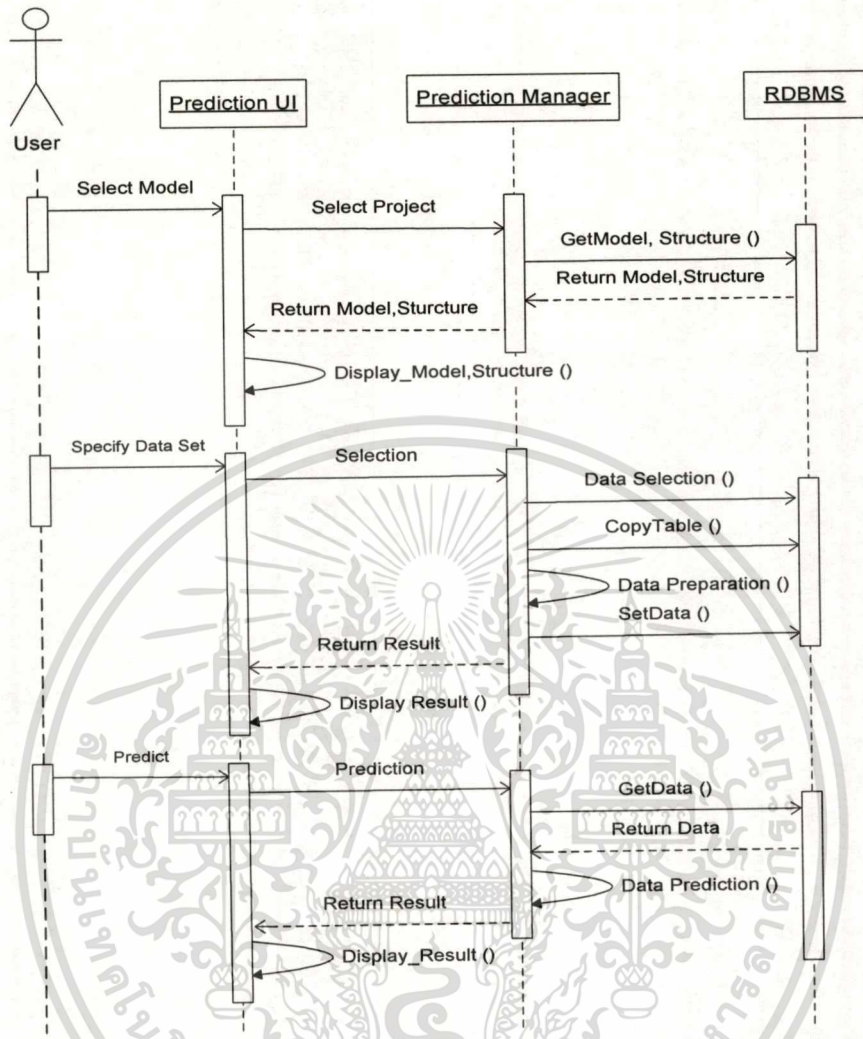
รูปที่ 4.7 ซีควเอนซ์ไดอะแกรมของการบันทึกโปรเจกต์

Actor: User

Description: ผู้ใช้ทำการบันทึกแบบจำลอง

Basic Flow:

1. ถ้าผลของการเรียนรู้และการทดสอบแบบจำลองเป็นที่น่าพอใจ ผู้ใช้ทำการบันทึกแบบจำลอง
2. ระบบทำการบันทึกโปรเจกต์ โดยในการบันทึกจะประกอบด้วยแบบจำลองและโครงสร้างของข้อมูลที่ใช้ในการสร้างแบบจำลอง



รูปที่ 4.8 ซีควเอนซ์ไดอะแกรมของการพยากรณ์ข้อมูล

Actor: User

Description: ผู้ใช้ทำการพยากรณ์ข้อมูล

Basic Flow:

1. ผู้ใช้ระบุแบบจำลองที่ใช้ในการพยากรณ์ข้อมูล
2. ระบบทำการโหลดแบบจำลองและ โครงสร้างของข้อมูลที่ใช้สร้างแบบจำลอง
3. ผู้ใช้ระบุชุดข้อมูลที่ต้องการพยากรณ์
4. ระบบสร้างตารางใหม่โดยการคัดลอกข้อมูลตามที่ใช้เลือกจากตารางต้นฉบับ
5. ระบบทำการเตรียมข้อมูลที่จะนำมาพยากรณ์

6. ระบบทำการพยากรณ์ข้อมูล และแสดงผลการพยากรณ์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาวิจัยเท่านั้น ไม่สามารถเผยแพร่โดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 4.3 รายละเอียดเกี่ยวกับตารางที่ใช้ในระบบ

รายละเอียดเกี่ยวกับตารางที่ใช้ในระบบค่าดำไม่มันิ่งโดยใช้โครงข่ายประสาทเทียม ซึ่งใช้ในการเก็บโปรเจกต์ ที่ประกอบด้วยแบบจำลอง และ โครงสร้างของข้อมูลที่ใช้ในการสร้างแบบจำลอง มีทั้งหมด 6 ตาราง ดังต่อไปนี้

#### ตารางที่ 4.3 ตารางที่ใช้ในการเก็บโปรเจกต์

ลำดับ	ชื่อตาราง	คำอธิบาย
1.	ProjectDesc	เก็บรายละเอียดของโปรเจกต์
2.	ConfigModel	เก็บตัวแปรของแบบจำลอง
3.	ProjectModel	เก็บแบบจำลอง (ค่าน้ำหนัก) ของแต่ละโปรเจกต์
4.	ColumnsDesc	เก็บชื่อคอลัมน์ และชนิดข้อมูลที่ใช้ในการสร้างแบบจำลอง
5.	ColumnsDomain	เก็บค่าที่เป็นไปได้ทั้งหมดของคอลัมน์ที่มีชนิดข้อมูลที่ไม่เป็นตัวเลข
6.	ColumnsMaxMin	เก็บค่ามากที่สุดและค่าต่ำที่สุดของแต่ละคอลัมน์ที่ใช้ในการสร้างแบบจำลอง

#### ตารางที่ 4.4 ตาราง ProjectDesc

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	PK
2.	ProjectName	nvarchar	50	ชื่อโปรเจกต์	

#### ตารางที่ 4.5 ตาราง ConfigModel

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	FK
2.	InputUnit	int	4	จำนวน โหนดในชั้นอินพุต	
3.	OutputUnit	int	4	จำนวน โหนดในชั้นเอาต์พุต	
4.	HiddenUnit	int	4	จำนวน โหนดในชั้นฮิดเดน	
5.	HidLayer	int	4	จำนวนชั้นของชั้นฮิดเดน	
6.	SumError	float	8	ค่าความผิดพลาดของแบบจำลอง	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 ตาราง ProjectModel

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	FK
2.	FstHidWgt	float	8	ค่าน้ำหนักของชั้นฮิดเดนชั้นแรก	
3.	HiddenWgt	float	8	ค่าน้ำหนักของชั้นฮิดเดนที่อยู่ระหว่างชั้นแรกกับชั้นเอาต์พุต	
4.	OutputWgt	float	8	ค่าน้ำหนักของชั้นเอาต์พุต	
5.	FstHidBias	float	8	ค่าไบแอสของชั้นฮิดเดนชั้นแรก	
6.	HiddenBias	float	8	ค่าไบแอสของชั้นฮิดเดนที่อยู่ระหว่างชั้นแรกกับชั้นเอาต์พุต	
7.	OutputBias	float	8	ค่าไบแอสของชั้นเอาต์พุต	

ตารางที่ 4.7 ตาราง ColumnsDesc

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	FK
2.	ColumnName	nvarchar	50	ชื่อคอลัมน์ที่ใช้ในการสร้างแบบจำลอง	
3.	ColumnType	nvarchar	50	ชนิดข้อมูลของแต่ละคอลัมน์	

ตารางที่ 4.8 ตาราง ColumnsDomain

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	FK
2.	ColumnName	nvarchar	50	ชื่อคอลัมน์ที่มีชนิดข้อมูลไม่เป็นตัวเลข	
3.	RealValue	nvarchar	50	ค่าที่มากที่สุดของแต่ละคอลัมน์	
4.	NewValue	float	8	ค่าน้อยที่สุดของแต่ละคอลัมน์	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.9 ตาราง ColumnsMaxMin

ลำดับ	ชื่อคอลัมน์	ชนิดข้อมูล	ขนาด	คำอธิบาย	ข้อสังเกต
1.	ProjectID	int	4	รหัสโปรเจกต์	FK
2.	ColumnName	nvarchar	50	ชื่อคอลัมน์	
3.	MaxValue	float	8	ค่าที่มากที่สุดของแต่ละคอลัมน์	
4.	MinValue	float	8	ค่าที่น้อยที่สุดของแต่ละคอลัมน์	



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

### การพัฒนาระบบ

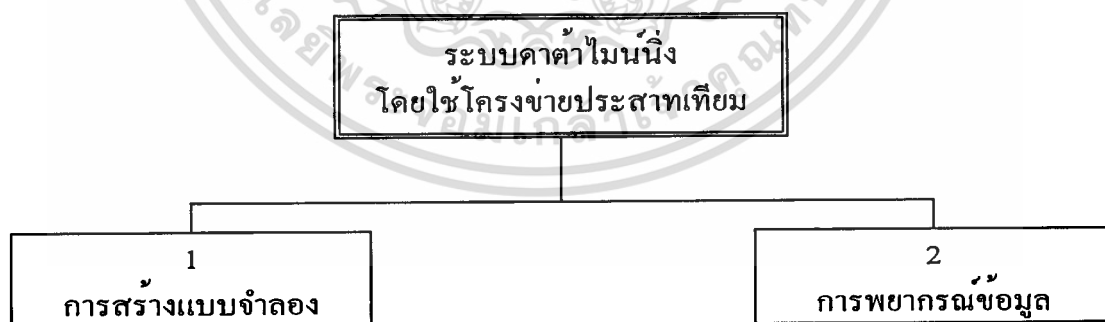
ในโครงการพัฒนาระบบนี้ เป็นการนำทฤษฎีาคำไม้นิ่งและหลักการของโครงข่ายประสาทเทียมมาใช้ในลักษณะของงานด้านการพยากรณ์ข้อมูล โดยใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชันในการสร้างแบบจำลองให้กับระบบ และพัฒนาระบบโดยใช้โปรแกรม Microsoft Visual Studio.net 2003 ร่วมกับใช้ Microsoft SQL Server 2000 เป็นฐานข้อมูล

ซึ่งในบทนี้จะอธิบายถึงผังโครงสร้างของระบบ อัลกอริทึมที่สำคัญที่ใช้ในการพัฒนาระบบ ขั้นตอนการใช้งานระบบ รวมถึงตัวอย่างหน้าจอที่ได้จากการพัฒนาระบบ

#### 5.1 ผังโครงสร้างของระบบ

ระบบคำคำไม้นิ่ง โดยใช้โครงข่ายประสาทเทียม มีผังโครงสร้างของการพัฒนาระบบดังต่อไปนี้

##### 5.1.1 ผังโครงสร้างหลักของระบบคำคำไม้นิ่งโดยใช้โครงข่ายประสาทเทียม



รูปที่ 5.1 แสดงผังโครงสร้างหลักของระบบคำคำไม้นิ่งโดยใช้โครงข่ายประสาทเทียม

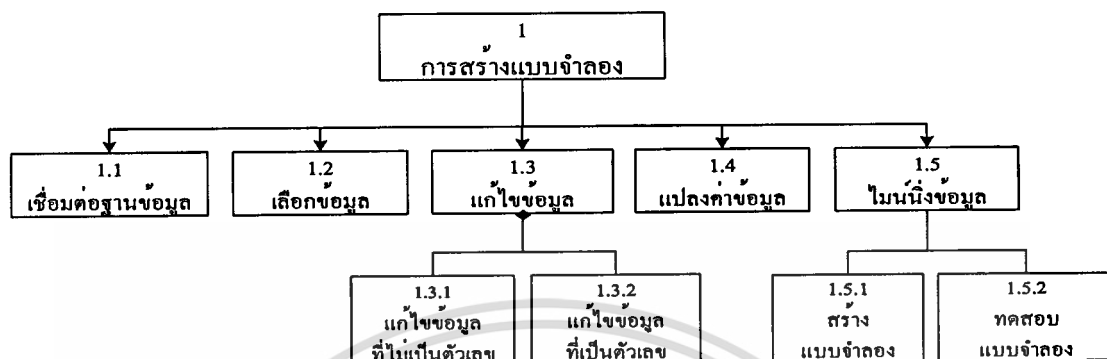
โครงสร้างหลักของการพัฒนาระบบ แบ่งออกเป็น 2 โมดูลย่อย คือ

1. การสร้างแบบจำลอง เป็นโมดูลที่ทำการสร้างแบบจำลองใหม่ให้กับระบบ

2. การพยากรณ์ข้อมูล เป็นโมดูลที่ทำการพยากรณ์ข้อมูล โดยการใช้แบบจำลองที่มีอยู่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## 1. โครงสร้างของการสร้างแบบจำลอง



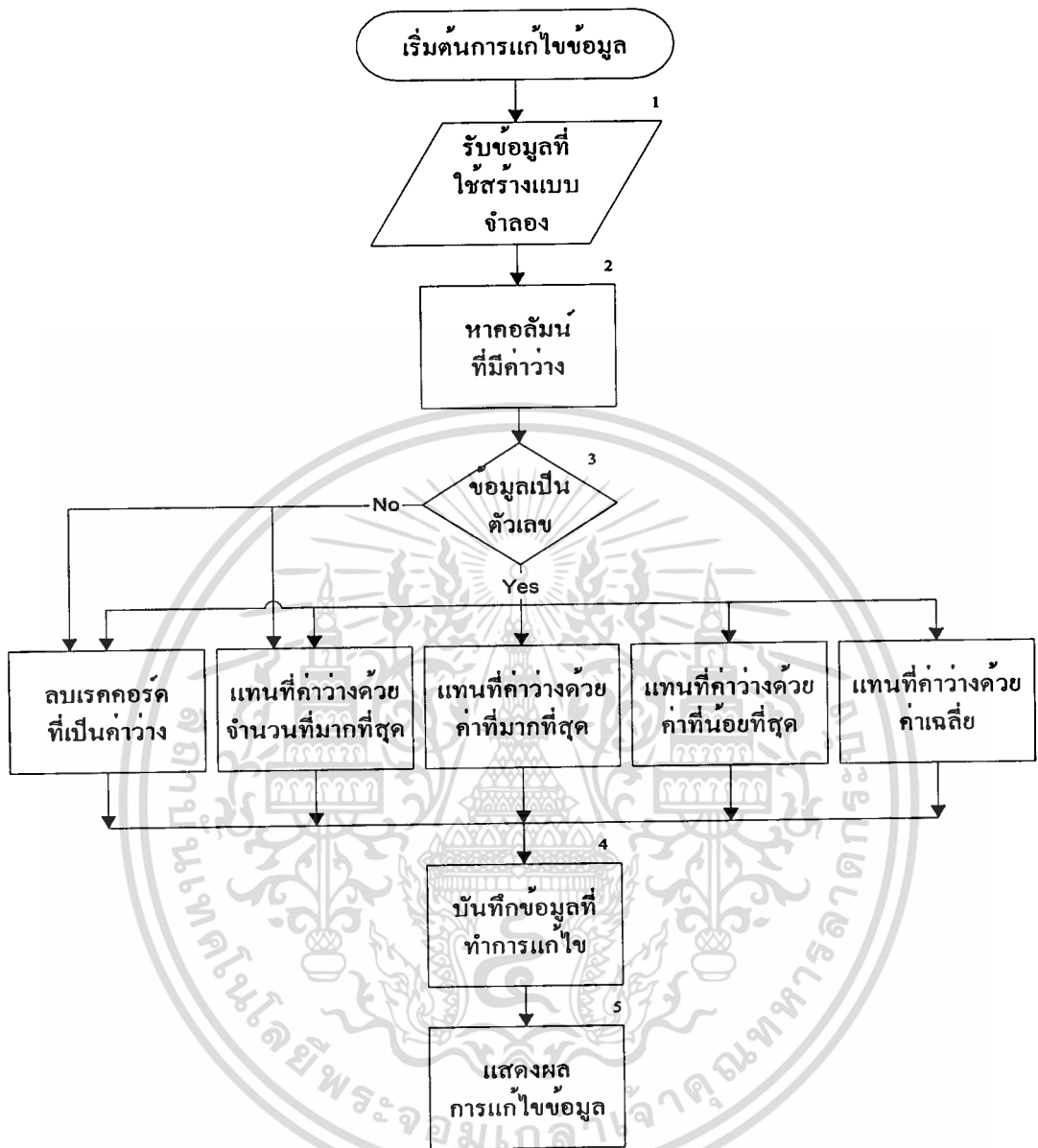
### รูปที่ 5.2 แสดงผังโครงสร้างของการสร้างแบบจำลอง

การสร้างแบบจำลอง มีการทำงานแบ่งออกเป็นโมดูลย่อยๆ ดังนี้

1.1 การเชื่อมต่อฐานข้อมูล เป็นโมดูลที่ใช้ในการเชื่อมต่อกับฐานข้อมูล ที่เก็บข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลอง

1.2 การเลือกข้อมูล เป็นโมดูลที่ใช้ในการเลือกข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลอง

1.3 การแก้ไขข้อมูล เป็นโมดูลที่ใช้ในการจัดการกับข้อมูลที่เป็นค่าว่าง ซึ่งจะมีการเลือกทำระหว่างข้อมูลที่มีค่าเป็นตัวเลข กับข้อมูลที่มีค่าไม่เป็นตัวเลข โดยในโมดูลของการแก้ไขข้อมูลนั้นใช้อัลกอริทึมการแก้ไขข้อมูล ดังแสดงที่หัวข้อ 5.2.1 และสามารถแสดงรายละเอียดด้วยผังงานโฟว์ชาร์ต ดังรูปที่ 5.3



รูปที่ 5.3 แสดงผังงานโปรแกรมการแก้ไขข้อมูล

รายละเอียดการทำงานของโปรแกรมการแก้ไขข้อมูล สามารถอธิบายการทำงานได้ดังต่อไปนี้

1. รับข้อมูลที่ได้จากโมดูลการเลือกข้อมูล ซึ่งเป็นข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลอง
2. ระบบทำการตรวจหาข้อมูลในแต่ละคอลัมน์ที่มีค่าว่าง
3. ตรวจสอบข้อมูลในแต่ละคอลัมน์ที่มีค่าว่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ได้ 2 แบบ ดังนี้

- ถ้าข้อมูลมีค่าไม่เป็นตัวเลขจะสามารถจัดการกับข้อมูลที่เป็นค่าว่าง

1. การลบเรคคอร์ดที่มีค่าว่าง
2. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด

- ถ้าข้อมูลมีค่าเป็นตัวเลขจะสามารถจัดการกับข้อมูลที่เป็นค่าว่างได้

5 แบบ ดังนี้

6. การลบเรคคอร์ดที่มีค่าว่าง
7. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด
8. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่ามากที่สุด
9. การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่าน้อยที่สุด
10. การแทนค่าที่ว่าง ด้วยข้อมูลที่เป็นค่าเฉลี่ย

4.ระบบทำการบันทึกข้อมูลที่ได้ทำการแก้ไข

5.แสดงผลการแก้ไขข้อมูล

1.4 การแปลงค่าข้อมูล เป็นโมดูลที่ใช้ในการแปลงค่าข้อมูลให้อยู่ในช่วงที่เหมาะสม คือ 0 - 1 โดยใช้สูตรตามที่แสดงไว้ที่อัลกอริทึมการแปลงค่าข้อมูล หัวข้อที่ 5.2.2

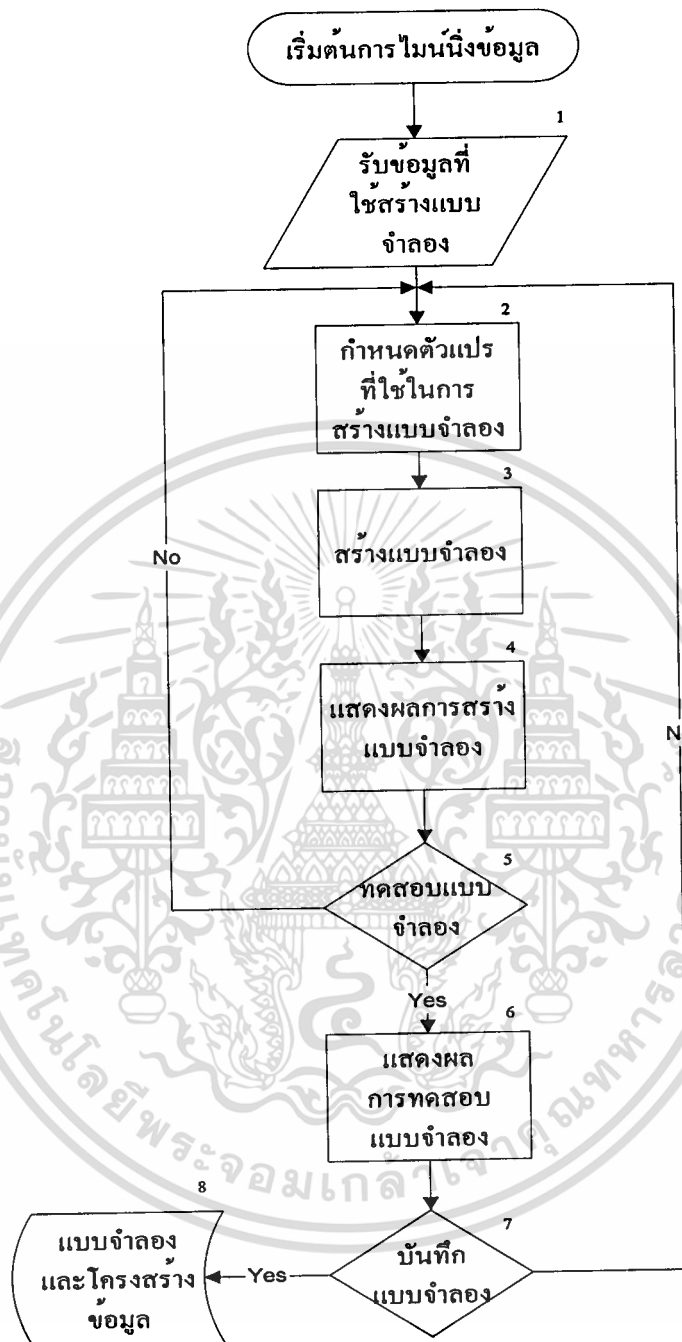
1.5 การไม้นิ่งข้อมูล เป็นโมดูลที่ประกอบด้วยโมดูลย่อยๆ 2 โมดูล คือ

1.5.1 การสร้างแบบจำลอง เป็นโมดูลที่ใช้ในการสร้างแบบจำลอง

1.5.2 การทดสอบแบบจำลอง เป็นโมดูลที่ใช้ในการทดสอบ

ประสิทธิภาพของแบบจำลอง

ซึ่งในโมดูลของการไม้นิ่งข้อมูลนั้น สามารถแสดงรายละเอียดด้วยผังงานโฟว์ชาร์ต ดังรูปที่ 5.4



รูปที่ 5.4 แสดงผังงานโปรแกรมไมนิ่งข้อมูล

รายละเอียดการทำงานของโปรแกรมไมนิ่งข้อมูล สามารถอธิบายการทำงานได้ดังต่อไปนี้

1. รับข้อมูลที่จะนำมาใช้ในการสร้างแบบจำลอง ซึ่งจะต้องผ่านการทำ

การแก้ไขข้อมูล และการแปลงค่าข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. กำหนดตัวแปรที่ใช้ในการสร้างแบบจำลอง คือ จำนวนโหนดในชั้นฮิดเดน ค่าอัตราการเรียนรู้ ค่าจำนวนรอบสูงสุดที่ใช้ในการเรียนรู้ ค่าวัดความผิดพลาดที่ยอมรับได้ และจำนวนเรคคอร์ดที่ใช้ในการเรียนรู้

3. ระบบทำการสร้างแบบจำลองตามอัลกอริทึมการสร้างแบบจำลอง ดังแสดงที่หัวข้อ 5.2.3

4. แสดงผลที่ได้จากการสร้างแบบจำลอง คือ จำนวนรอบสูงสุดที่ใช้ในการเรียนรู้ ค่าวัดความผิดพลาดของแบบจำลอง และแสดงการเปรียบเทียบค่าจริงกับค่าเอาท์พุทที่ได้จากการคำนวณ ในรูปแบบของกราฟและรายงาน

5. ตรวจสอบว่าผู้ใช้ต้องการทำการทดสอบแบบจำลองหรือไม่

- ถ้าต้องการ ระบบจะทำการทดสอบแบบจำลอง ด้วยการนำข้อมูลที่เหลือจากส่วนของการสร้างแบบจำลองมาประมวลผลกับแบบจำลองที่ได้ ตามอัลกอริทึมการทดสอบแบบจำลอง ดังแสดงที่หัวข้อ 5.2.4

- ถ้าไม่ต้องการ ผู้ใช้สามารถกลับไปทำการกำหนดตัวแปร และสร้างแบบจำลองใหม่ได้

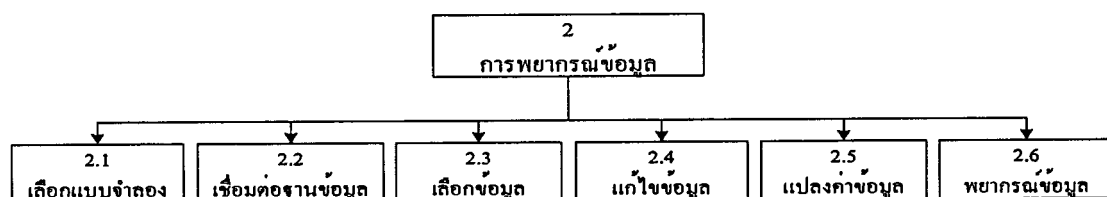
6. แสดงผลที่ได้จากการทดสอบแบบจำลอง คือ ค่าวัดความผิดพลาดของการทดสอบแบบจำลอง และแสดงการเปรียบเทียบค่าจริงกับค่าเอาท์พุทที่ได้จากการคำนวณ ในรูปแบบของกราฟและรายงาน

7. ตรวจสอบว่าผู้ใช้ต้องการบันทึกแบบจำลองหรือไม่

- ถ้าผู้ใช้ต้องการ ระบบจะทำการบันทึกเป็นโปรเจกตลงฐานข้อมูล ซึ่งโปรเจกตประกอบด้วย แบบจำลองและโครงสร้างของข้อมูลที่ใช้ในการสร้างแบบจำลอง

- ถ้าผู้ใช้ไม่ต้องการ ผู้ใช้สามารถกลับไปทำการกำหนดตัวแปร และสร้างแบบจำลองใหม่ได้

## 2. โครงสร้างของการพยากรณ์ข้อมูล



เอกสารนี้เป็นเอกสารที่สงวนไว้รูปที่ 5.5 แสดงผังโครงสร้างของการพยากรณ์ข้อมูลไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การพยากรณ์ข้อมูล มีการทำงานแบ่งออกเป็น โมดูลย่อยๆ ดังนี้

2.1 การเลือกแบบจำลอง เป็น โมดูลที่ใช้ในการเลือกแบบจำลอง เพื่อรับข้อมูลที่เกี่ยวข้องกับแบบจำลอง ซึ่งจะนำมาใช้ในการพยากรณ์ข้อมูล

2.2 การเชื่อมต่อฐานข้อมูล เป็น โมดูลที่ใช้ในการเชื่อมต่อกับฐานข้อมูล ซึ่งผู้ใช้ต้องทำการระบุชื่อเซิร์ฟเวอร์ และชื่อฐานข้อมูล

2.3 การเลือกข้อมูล เป็น โมดูลที่ใช้ในการเลือกข้อมูลที่จะนำมาทำการพยากรณ์

2.4 การแก้ไขข้อมูล เป็น โมดูลที่ใช้ในการจัดการกับข้อมูลที่เป็นค่าว่าง ซึ่งจัดการโดยการลบเรคคอร์ดที่มีค่าว่าง

2.5 การแปลงค่าข้อมูล เป็น โมดูลที่ใช้ในการแปลงค่าข้อมูลให้อยู่ในช่วงที่เหมาะสม โดยใช้สูตรตามที่แสดงไว้ที่อัลกอริทึมการแปลงค่าข้อมูล หัวข้อที่ 5.2.2

2.6 การพยากรณ์ข้อมูล เป็น โมดูลที่ใช้ในการพยากรณ์ข้อมูล โดยให้นำข้อมูลที่รับมาทำการประมวลกับแบบจำลองที่เลือก ตามอัลกอริทึมการพยากรณ์ข้อมูล ดังแสดงที่หัวข้อ 5.2.5

## 5.2 อัลกอริทึมที่ใช้ในการพัฒนาระบบ

สำหรับอัลกอริทึมของการทำงานหลักๆ ของระบบสามารถแสดงได้ ดังต่อไปนี้

### 5.2.1 อัลกอริทึมสำหรับการแก้ไขข้อมูล (Data Cleaning)

เป็นอัลกอริทึมที่ทำการจัดการกับข้อมูลที่เป็นค่าว่าง

Step 1 : Read Data Set from Database

Step 2 : Calculate detail Data Set : No. of TotalRecord, No. of Missing Value,

MaxNo., MaxValue, MinValue and AvgValue

Step 3 : If Missing Value  $\neq$  0 that do Step 4

Step 4 : Select method for Cleaning

Step 4.1 : If Method = 1 do

Deleted all Missing Record

Step 4.2 : If Method = 2 do

Replace Missing Value with Value of Maximum No.

Step 4.3 : If Method = 3 do

Replace Missing Value with Maximum Value

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Step 4.4 : If Method = 4 do

Replace Missing Value with Minimum Value

Step 4.5 : If Method = 5 do

Replace Missing Value with Average Value

### 5.2.2 อัลกอริทึมสำหรับการแปลงค่าข้อมูล (Data Transformation)

เป็นอัลกอริทึมที่ทำการแปลงค่าข้อมูลให้เหมาะสมกับอัลกอริทึมที่ใช้ในการสร้างแบบจำลอง และการแปลงค่าข้อมูลกลับเป็นค่าจริง โดยการใช้สูตรดังนี้

$$(1) \text{ NewValue} = (((\text{Value} - \text{min}) / (\text{max} - \text{min})) * (\text{NewMax} - \text{NewMin})) + \text{NewMin}$$

$$(2) \text{ RealValue} = (((\text{NewValue} - \text{Newmin}) / (\text{Newmax} - \text{Newmin})) * (\text{Max} - \text{Min})) + \text{Min}$$

Transformation :

Step 1 : Read Data Set

Step 2 : Data Transformation

Calculate new value by equation (1)

Replace real value with new value

Re-Transformation :

Step 1 : Read Data Set

Step 2 : Data Re-Transformation

Calculate real value by equation (2)

Replace value with real value

### 5.2.3 อัลกอริทึมสำหรับการสร้างแบบจำลอง (Training)

เป็นอัลกอริทึมที่ทำการสร้างแบบจำลองให้กับระบบ โดยใช้โครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

Step 1 : Identify no. of input node, no. of output node, no. of hidden node, no. of hidden layer, Learning Rate, Error Rate, Maximum Loop and Training Records

Step 2 : Random all weights in network

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Step 3 : While terminate condition is not satisfied ( SumError < ErrorRate or Loop > MaxLoop) Do Step 3.1 to 3.8

Step 3.1 : For each training sample S in Samples do

// Propagate the input forward

Step 3.2 : For each hidden layer units do

$$\text{net}_j = \sum W_{ij} x_i + \theta_j$$

$$O_j = (1 + e^{-\text{net}_j})^{-1}$$

Step 3.3 : For each output layer units k do

$$\text{net}_k = \sum W_{jk} O_j + \theta_k$$

$$O_k = (1 + e^{-\text{net}_k})^{-1}$$

//Backpropagation

Step 3.4 : Calculate Error for each output layer unit k

$$\text{Err}_k = O_k(1 - O_k)(Y_k - O_k)$$

Step 3.5 : Calculate Error for each hidden layer units j

$$\text{Err}_j = O_j(1 - O_j) \sum \text{Err}_k W_{jk}$$

//Update weight

Step 3.6 : For each output layer weight ( $W_{jk}$ ) do

$$\Delta W_{jk} = (l) \text{Err}_k O_j$$

$$W_{jk}(t+1) = W_{jk}(t) + \Delta W_{jk}$$

For each output layer bias ( $\theta_k$ ) do

$$\Delta \theta_k = (l) \text{Err}_k$$

$$\theta_k(t+1) = \theta_k(t) + \Delta \theta_k$$

Step 3.7 : For each hidden layer weight ( $W_{ij}$ ) do

$$\Delta W_{ij} = (l) \text{Err}_j x_i$$

$$W_{ij}(t+1) = W_{ij}(t) + \Delta W_{ij}$$

For each output layer bias ( $\theta_j$ ) do

$$\Delta \theta_j = (l) \text{Err}_j$$

$$\theta_j(t+1) = \theta_j(t) + \Delta \theta_j$$

Step 3.8 : Increase loop and check for terminate condition

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.2.4 อัลกอริทึมสำหรับการทดสอบแบบจำลอง (Testing)

เป็นอัลกอริทึมที่ทำการทดสอบแบบจำลองที่ได้จากการสร้างโดยใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน

Step 1 : For each testing sample S in samples do

Step 2 : For each hidden layer units do

$$\text{net}_j = \sum W_{ij} x_i + \theta_j$$

$$O_j = (1 + e^{-\text{net}_j})^{-1}$$

Step 3 : For each output layer units k do

$$\text{net}_k = \sum W_{jk} O_j + \theta_k$$

$$O_k = (1 + e^{-\text{net}_k})^{-1}$$

Step 4 : Calculate Error for each output layer unit k

Step 5 : Calculate Sum Square Error for all samples

### 5.2.5 อัลกอริทึมสำหรับการพยากรณ์ข้อมูล (Prediction)

เป็นอัลกอริทึมที่ทำการคำนวณค่าเอาต์พุตจากแบบจำลองที่เลือกมาใช้

Step 1 : Read model parameter from project (InputUnit, HiddenUnit, MaxValue, MinValue, All weights in network)

Step 2 : For I = 1 to InputUnit

Read Input data set from user

Step 3 : Calculate Result

For each hidden layer units do

$$\text{net}_j = \sum W_{ij} x_i + \theta_j$$

$$O_j = (1 + e^{-\text{net}_j})^{-1}$$

For each output layer units k do

$$\text{net}_k = \sum W_{jk} O_j + \theta_k$$

$$O_k = (1 + e^{-\text{net}_k})^{-1}$$

Step 4 : Display Result

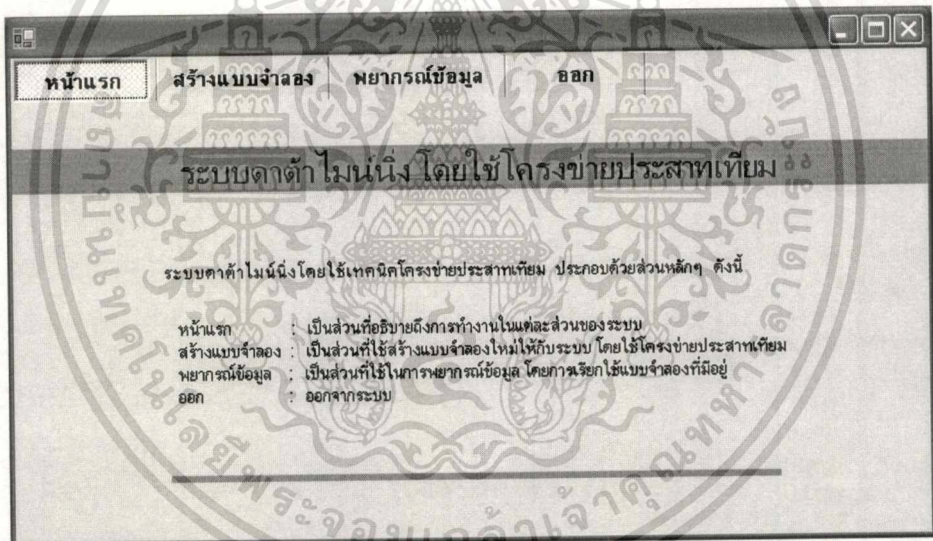
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.3 การใช้งานระบบ และตัวอย่างหน้าจอที่ได้จากการพัฒนาระบบ

สำหรับขั้นตอนการใช้งานระบบสามารถสรุปได้ ดังนี้

หน้าจอที่ปรากฏเมื่อเริ่มใช้งาน คือ หน้าแรก ซึ่งจะอธิบายถึงการทำงานในแต่ละส่วนของระบบ ดังนี้

- หน้าแรก : เป็นส่วนที่อธิบายถึงการทำงานในแต่ละส่วนของระบบ
- สร้างแบบจำลอง : เป็นส่วนที่ใช้สร้างแบบจำลองใหม่ให้กับระบบ โดยใช้โครงข่ายประสาทเทียมในการสร้างแบบจำลอง
- พยากรณ์ข้อมูล : เป็นส่วนที่ใช้ในการพยากรณ์ข้อมูล โดยการเรียกใช้แบบจำลองที่มีอยู่
- ออก : ออกจากระบบ



รูปที่ 5.6 หน้าจอแรกที่ปรากฏเมื่อเริ่มใช้งาน

### 5.3.1 การใช้งานระบบในส่วนการสร้างแบบจำลอง

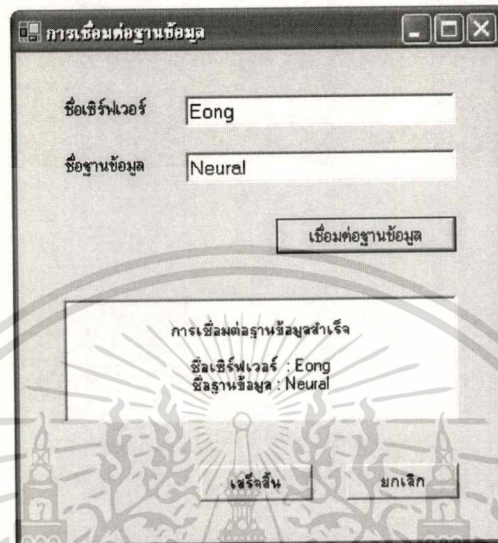
เมื่อผู้ใช้คลิกที่สร้างแบบจำลองจะปรากฏหน้าจอหลักในส่วนของการสร้างแบบจำลองใหม่ให้กับระบบ ซึ่งจะบอกถึงกระบวนการในการทำดาต้าไมนิ่ง ดังนี้

- ขั้นตอนที่ 1 ทำการเชื่อมต่อฐานข้อมูล
- ขั้นตอนที่ 2 ทำการเลือกชุดข้อมูล
- ขั้นตอนที่ 3 ทำความแก้ไขข้อมูล
- ขั้นตอนที่ 4 ทำการแปลงค่าข้อมูล
- ขั้นตอนที่ 5 ทำการไมนิ่งข้อมูล



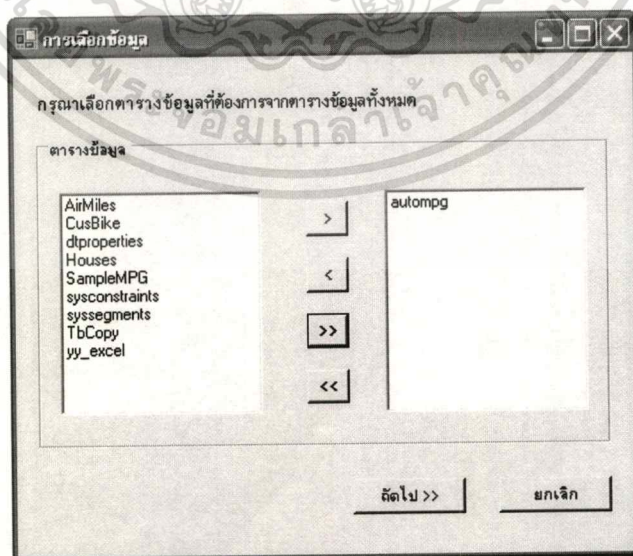
รูปที่ 5.7 หน้าจอหลักในส่วนของการสร้างแบบจำลองใหม่

ขั้นตอนที่ 1 ทำการเชื่อมต่อฐานข้อมูล โดยกำหนดชื่อเซิร์ฟเวอร์และชื่อฐานข้อมูลที่ใช้เก็บข้อมูลที่จะนำมาสร้างแบบจำลองให้กับระบบ ซึ่งในระบบนี้ใช้ฐานข้อมูล SQL Server 2000



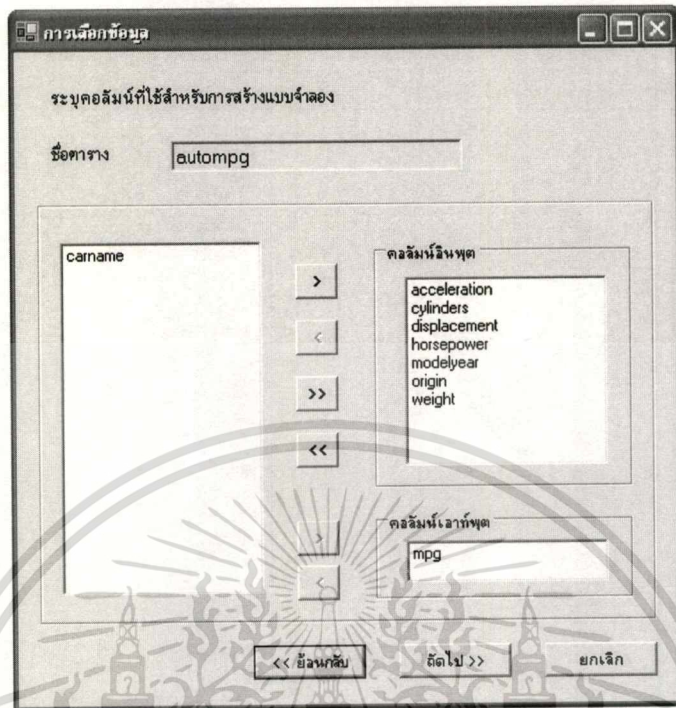
รูปที่ 5.8 หน้าจอการเชื่อมต่อฐานข้อมูล SQL Server 2000

ขั้นตอนที่ 2 เลือกชุดข้อมูล โดยทำการเลือกตาราง และระบุคอลัมน์ที่ใช้ในการสร้างแบบจำลอง

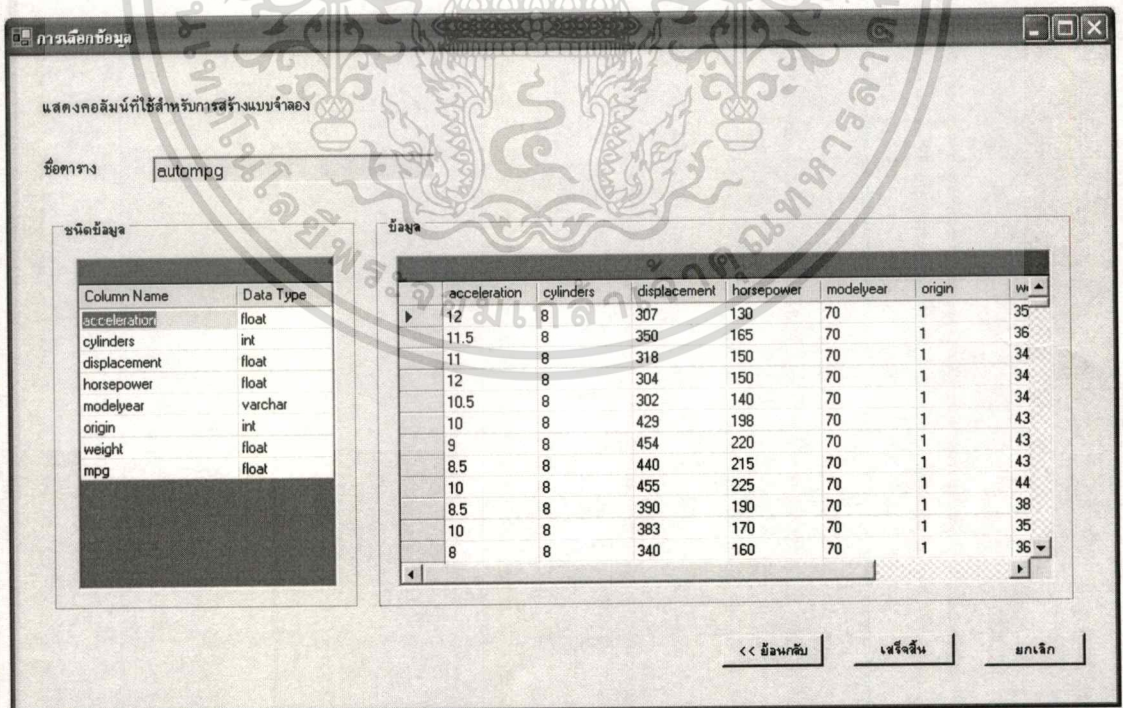


รูปที่ 5.9 หน้าจอการเลือกตาราง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.10 หน้าจอการระบุคอลัมน์



รูปที่ 5.11 หน้าจอแสดงชนิดข้อมูลและข้อมูลที่นำมาสร้างแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

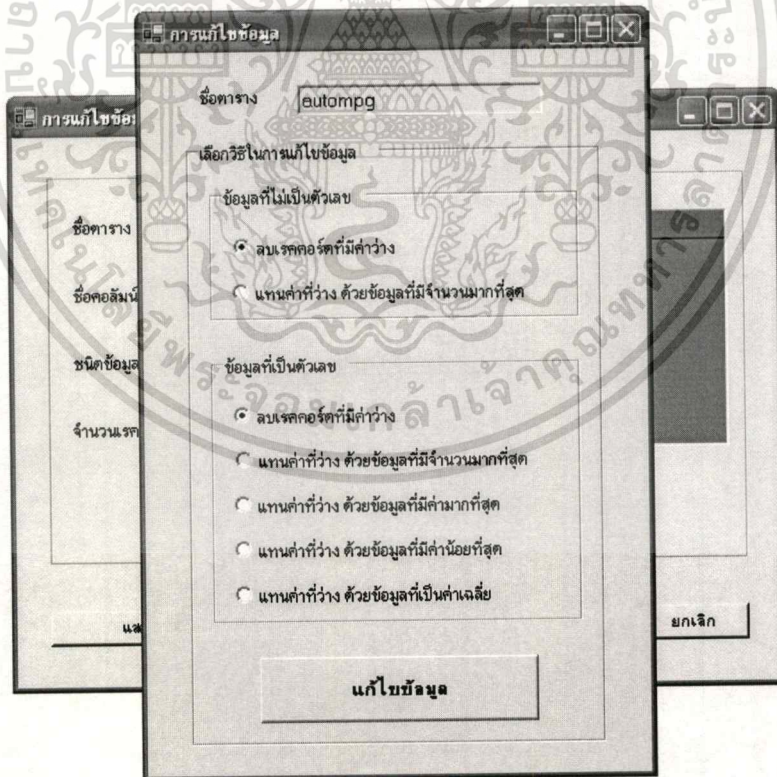
ขั้นตอนที่ 3 เลือกวิธีการแก้ไขข้อมูลกรณีที่มีข้อมูลเป็นค่าว่าง ซึ่งผู้ใช้สามารถเลือกแก้ไขข้อมูลแต่ละคอลัมน์ หรือแก้ไขทุกคอลัมน์พร้อมกันก็ได้ โดยผู้ใช้งานสามารถเลือกได้ตามความเหมาะสม ดังนี้

ข้อมูลที่ไม่ได้เป็นตัวเลขมีวิธีการจัดการ 2 รูปแบบ คือ

- การลบเรคคอร์ดที่มีค่าว่าง
- การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด

ข้อมูลที่เป็นตัวเลขมีวิธีการจัดการ 5 รูปแบบ คือ

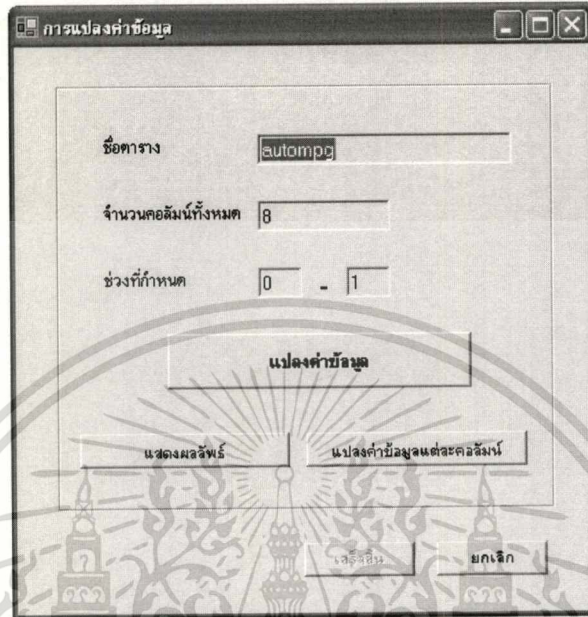
- การลบเรคคอร์ดที่มีค่าว่าง
- การแทนค่าที่ว่าง ด้วยข้อมูลที่มีจำนวนมากที่สุด
- การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่ามากที่สุด
- การแทนค่าที่ว่าง ด้วยข้อมูลที่มีค่าน้อยที่สุด
- การแทนค่าที่ว่าง ด้วยข้อมูลที่เป็นค่าเฉลี่ย



รูปที่ 5.12 หน้าจอการเลือกวิธีการแก้ไขข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 4 ทำการแปลงค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1



รูปที่ 5.13 หน้าจอการแปลงค่าข้อมูล

	acceleration	cylinders	displacement	horsepower	modelyear	origin	weight	mpg
▶	0.2380952380	1	0.617571059	0.456521739	0	0	0.536149702	0.2393
	0.208333333	1	0.728682170	0.646739130	0	0	0.589736319	0.1595
	0.178571428	1	0.645994832	0.565217391	0	0	0.516869861	0.2393
	0.238095238	1	0.609819121	0.565217391	0	0	0.516019279	0.1861
	0.148809523	1	0.604651162	0.510869565	0	0	0.520555713	0.2127
	0.119047619	1	0.932816537	0.826086956	0	0	0.773461865	0.1595
	0.059523809	1	0.997416020	0.945652173	0	0	0.777147717	0.1329
	0.029761904	1	0.961240310	0.918478260	0	0	0.765239580	0.1329
	0.119047619	1	1	0.972826086	0	0	0.797278140	0.1329
	0.029761904	1	0.832041343	0.782608695	0	0	0.634250070	0.1595
	0.119047619	1	0.813953488	0.673913043	0	0	0.552877799	0.1595
	0	1	0.702842377	0.619565217	0	0	0.565920045	0.1329

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ โดย ร.ร. 5.14 หน้าจอแสดงผลการแปลงค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1 ระเบียบข้อดำเนินการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนที่ 5 ทำการสร้างแบบจำลอง โดยมีขั้นตอนดังนี้

- กำหนดค่าตัวแปรเพื่อใช้ในการสร้างแบบจำลอง ดังนี้ จำนวนโหนดในชั้นอินพุต จำนวนโหนดในชั้นเอาต์พุต จำนวนโหนดในชั้นฮิดเดน และจำนวนของชั้นฮิดเดน
- ทำการเรียนรู้ข้อมูลเพื่อสร้างแบบจำลอง โดยการกำหนดค่าอัตราการเรียนรู้ ค่าวัดความผิดพลาดที่ยอมรับได้ จำนวนรอบที่มากที่สุดในการเรียนรู้ และจำนวนเรคคอร์ดที่ใช้ในการเรียนรู้
- ทำการทดสอบแบบจำลองเพื่อประเมินประสิทธิภาพของแบบจำลอง โดยการกำหนดจำนวนเรคคอร์ดที่ใช้ในการทดสอบ

The screenshot shows a software window titled "การไม่ทิ้งข้อมูล" (Don't lose data). It is divided into several sections for configuring a neural network model.

**ชื่อตาราง (Table Name):** autompg  
**จำนวนเรคคอร์ดทั้งหมด (Total Records):** 391

	Column Name	Data Type
Input1	acceleration	float
Input2	cylinders	int
Input3	displacement	float
Input4	horsepower	float
Input5	modelyear	varchar
Input6	origin	int
Input7	weight	float
Output	mpg	float

**การสร้างแบบจำลอง (Model Construction):**  
**อัตราการเรียนรู้ (Learning Rate):** 0.1  
**จำนวนรอบที่มากที่สุด (Maximum Epochs):** 5000  
**ค่าวัดความผิดพลาด (Error Tolerance):** 0.001  
**จำนวนเรคคอร์ด (Number of Records):** 274  
**สร้างแบบจำลอง (Build Model)** [แสดงผลลัพธ์]

**ผลของการสร้างแบบจำลอง (Model Construction Results):**  
**จำนวนรอบ (Epochs):** 5000  
**ค่าวัดความผิดพลาดเฉลี่ย (Average Error):** 0.0026283

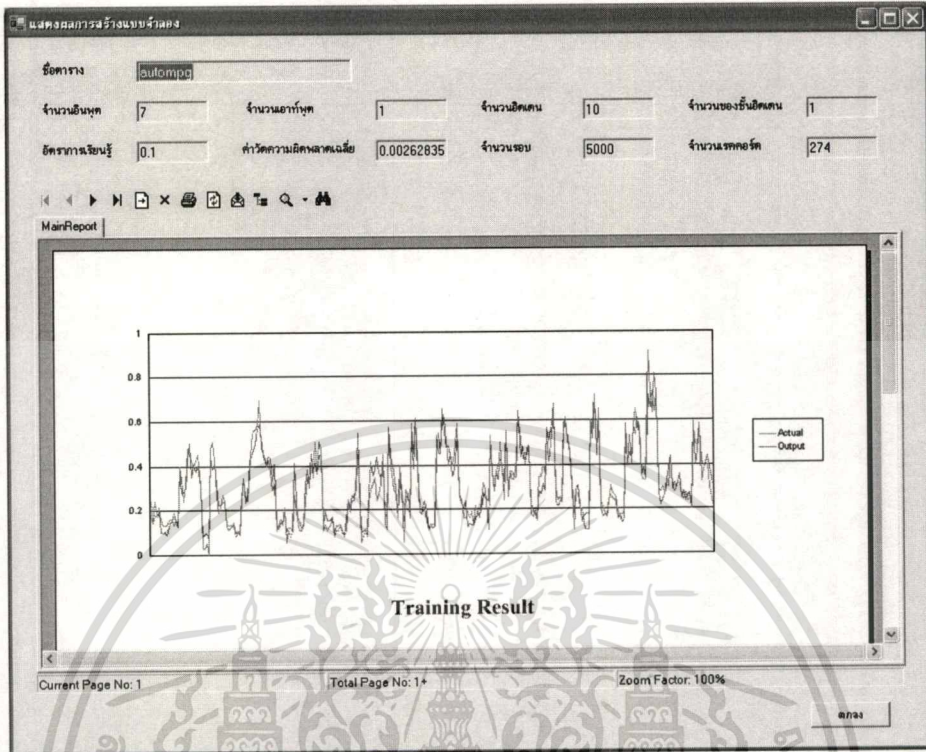
**กำหนดโครงสร้างของแบบจำลอง (Define Model Structure):**  
**จำนวนโหนดในชั้นอินพุต (Input Nodes):** 7  
**จำนวนโหนดในชั้นเอาต์พุต (Output Nodes):** 1  
**จำนวนโหนดในชั้นฮิดเดน (Hidden Nodes):** 10  
**จำนวนของชั้นฮิดเดน (Number of Hidden Layers):** 1

**การทดสอบแบบจำลอง (Model Testing):**  
**จำนวนเรคคอร์ด (Number of Records):** 117  
**ค่าวัดความผิดพลาดเฉลี่ย (Average Error):** 0.0119353  
**ทดสอบแบบจำลอง (Test Model)** [แสดงผลลัพธ์]

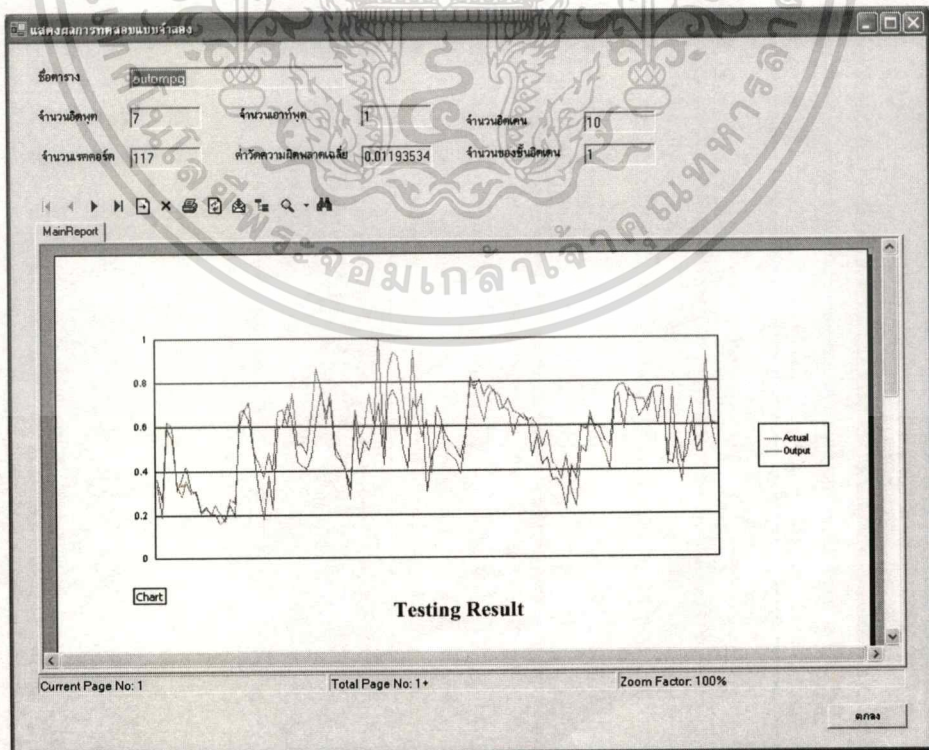
[เสร็จสิ้น] [ยกเลิก]

รูปที่ 5.15 หน้าจอการสร้างแบบจำลอง และการทดสอบแบบจำลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.16 หน้าจอแสดงผลการสร้างแบบจำลองในรูปแบบกราฟ



เอกสารนี้เป็นเอกสารที่ **รูปที่ 5.17 หน้าจอแสดงผลการทดสอบแบบจำลองในรูปแบบกราฟ** ผลิตขึ้นเพื่อใช้ในการศึกษาวิจัยเท่านั้น ไม่สามารถนำออกจำหน่ายหรือทำซ้ำโดยไม่ได้รับอนุญาตจากเจ้าของลิขสิทธิ์ หากมีข้อผิดพลาดประการใดขออภัยเป็นอย่างสูง และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แสดงผลการทดสอบแบบจำลอง

ชื่อตาราง: autompg

จำนวนอิคทุก: 7      จำนวนเอาท์ทุก: 1      จำนวนอิคเคน: 10

จำนวนรทกอร์ค: 117      ค่าวิคความผิดพลาดเฉลี่ย: 0.01193534      จำนวนของชั้นอิคเคน: 1

MainReport

### Testing Result

No.	Actual	Output	Difference	Error Rate	Sum Square Error
1	0.3351064	0.3623184	-0.0272120	0.0007405	0.0007405
2	0.1914894	0.2635039	-0.0720145	0.0051861	0.0059266
3	0.5984043	0.6243097	-0.0259054	0.0006711	0.0065977
4	0.5452128	0.5997663	-0.0545535	0.0029761	0.0095738
5	0.3324468	0.3092193	0.0232275	0.0005395	0.0101133
6	0.2872340	0.3711867	-0.0839527	0.0070481	0.0171613
7	0.3537234	0.4193881	-0.0656647	0.0043119	0.0214732
8	0.2978723	0.3137327	-0.0158604	0.0002516	0.0217247
9	0.3085106	0.2979773	0.0105333	0.0001110	0.0218357
10	0.2127660	0.2099481	0.0028179	0.0000079	0.0218436
11	0.2287234	0.2354623	-0.0067389	0.0000454	0.0218890
12	0.1994681	0.2027468	-0.0032787	0.0000107	0.0218998
13	0.2446809	0.2045180	0.0401629	0.0016131	0.0235128
14	0.2101064	0.1613291	0.0487773	0.0023792	0.0258921
15	0.1728723	0.1865315	-0.0136592	0.0001866	0.0260786
16	0.2712766	0.2454295	0.0258471	0.0006681	0.0267467
17	0.2526596	0.1945103	0.0581493	0.0033813	0.0301281
18	0.6090426	0.6647472	-0.0557046	0.0031030	0.0332311
19	0.6675532	0.6795227	-0.0119695	0.0001433	0.0333743
20	0.7101064	0.6340955	0.0760109	0.0057777	0.0391520

Current Page No: 1      Total Page No: 1      Zoom Factor: 100%

ตกลง

รูปที่ 5.18 หน้าจอแสดงผลการทดสอบแบบจำลองในรูปแบบรายงาน

ขั้นตอนที่ 6 วิเคราะห์ผลของแบบจำลอง ถ้ามีประสิทธิภาพเป็นที่น่าพอใจ ผู้ใช้สามารถบันทึกแบบจำลองและโครงสร้างของข้อมูลที่ใช้สร้างแบบจำลองเป็นโปรเจกต์เก็บไว้เพื่อนำไปใช้ในสถานการณ์ข้อมูลต่อไป

การสร้างแบบจำลอง

กรุณาใส่ชื่อแบบจำลอง

ชื่อแบบจำลอง: autompg\_5000

บันทึก      ยกเลิก

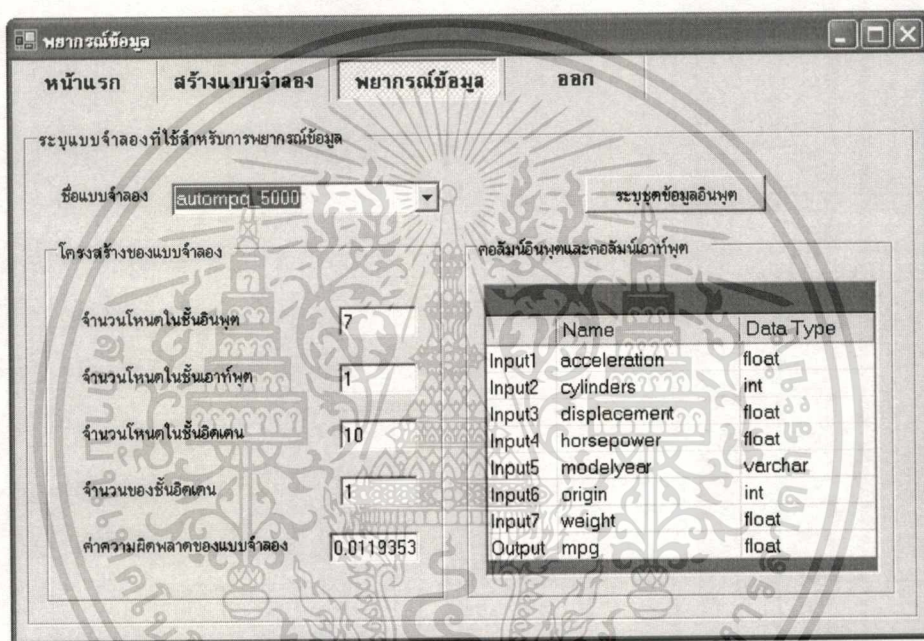
รูปที่ 5.19 หน้าจอการบันทึกแบบจำลองใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.3.2 การใช้งานระบบในส่วนการพยากรณ์ข้อมูล

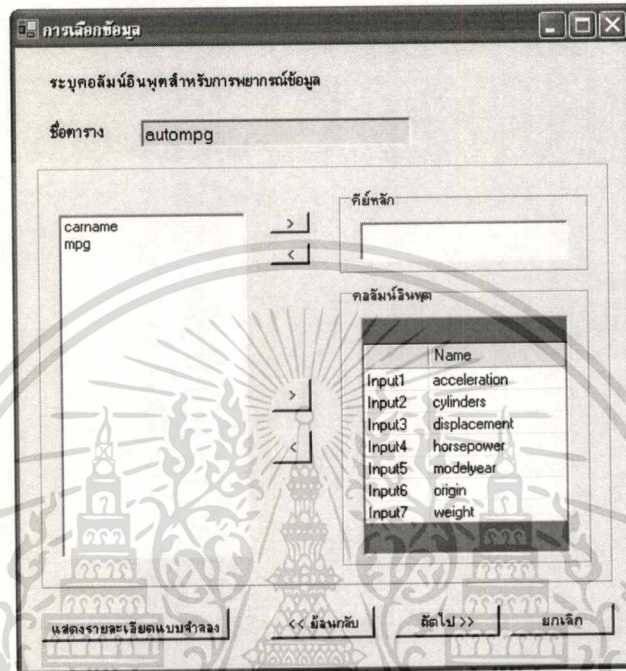
เมื่อผู้ใช้คลิกที่พยากรณ์ข้อมูลจะปรากฏหน้าจอหลักในส่วนของการพยากรณ์ข้อมูล โดยมีขั้นตอนในการพยากรณ์ข้อมูล ดังนี้

ขั้นตอนที่ 1 เลือกแบบจำลองที่มีอยู่ โดยที่จะต้องเลือกแบบจำลองให้เหมาะสมกับข้อมูลที่นำมาพยากรณ์ เมื่อเลือกแบบจำลองแล้วโปรแกรมจะแสดงโครงสร้างของแบบจำลอง พร้อมกับข้อมูลอินพุตและข้อมูลเอาต์พุต



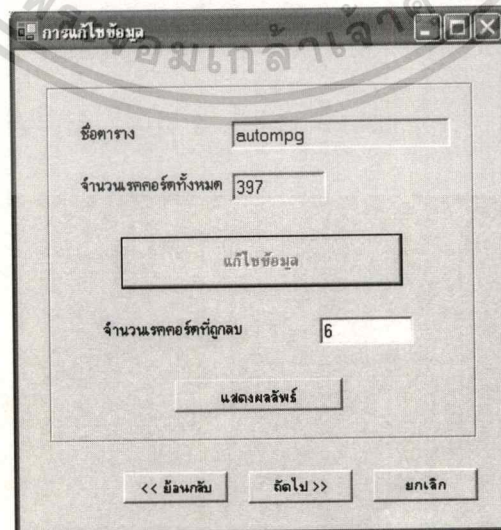
รูปที่ 5.20 หน้าจอแสดง โครงสร้างของแบบจำลอง พร้อมกับข้อมูลอินพุตและข้อมูลเอาต์พุต

ขั้นตอนที่ 2 คลิกที่ปุ่มระบุชุดข้อมูลอินพุตเพื่อเลือกชุดข้อมูล โดยทำการเลือกตารางและระบุคอลัมน์ที่นำมาใช้ในการพยากรณ์ข้อมูล ซึ่งจะต้องสอดคล้องกับแบบจำลองที่เลือก



รูปที่ 5.21 หน้าจอการระบุคอลัมน์ที่นำมาใช้พยากรณ์ข้อมูล

ขั้นตอนที่ 3 ทำการแก้ไขข้อมูลด้วยวิธีการลบเรคคอร์ดที่มีค่าว่าง



เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ซึ่งการนำเอกสารนี้ไปใช้โดยไม่ได้รับอนุญาตถือว่าผิดกฎหมาย ผู้ใช้ควรปฏิบัติตามเงื่อนไขการใช้งานที่ปรากฏในเอกสารฉบับนี้

รูปที่ 5.22 หน้าจอการแก้ไขข้อมูลที่นำมาใช้พยากรณ์ข้อมูล

ขั้นตอนที่ 4 ทำการแปลงค่าของข้อมูลให้อยู่ในช่วง 0 ถึง 1

การแปลงค่าข้อมูล

ชื่อตาราง: autompd

จำนวนคอลัมน์: 7

ช่วงที่กำหนด: 0 - 1

แปลงค่าข้อมูล

แสดงผลลัพธ์

<< ย้อนกลับ    ถัดไป >>    ยกเลิก

รูปที่ 5.23 หน้าจอการแปลงค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1

การแปลงค่าข้อมูล

ชื่อตาราง: autompd

	acceleration	cylindeis	displacement	horsepower	modelyear	origin	weight	rec
▶	0.238095238	1	0.617571059	0.456521739	0	0	0.536149702	1
	0.208333333	1	0.728682170	0.646739130	0	0	0.589736319	2
	0.178571428	1	0.645994832	0.565217391	0	0	0.516869861	3
	0.238095238	1	0.609819121	0.565217391	0	0	0.516019279	4
	0.148809523	1	0.604651162	0.510869565	0	0	0.520555713	5
	0.119047619	1	0.932816537	0.826086956	0	0	0.773461865	6
	0.059523809	1	0.997416020	0.945652173	0	0	0.777147717	7
	0.029761904	1	0.961240310	0.918478260	0	0	0.765239580	8
	0.119047619	1	1	0.972826086	0	0	0.797278140	9
	0.029761904	1	0.832041343	0.782608695	0	0	0.634250070	10
	0.119047619	1	0.813953488	0.673913043	0	0	0.552877799	11
	0	1	0.702842377	0.619565217	0	0	0.565920045	12

ตกลง

รูปที่ 5.24 หน้าจอแสดงผลการแปลงค่าข้อมูลให้อยู่ในช่วง 0 ถึง 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## ขั้นตอนที่ 5 ทำการคำนวณหาผลลัพธ์

Prediction

ชื่อตาราง  ชื่อคอลัมน์ที่ทำนายพยากรณ์

จำนวนเรคคอร์ด  จำนวนเรคคอร์ดที่ทำนายพยากรณ์

**พยากรณ์ข้อมูล**

rec	mpg	acceleration	cylinders	displacement	horsepower	modely
1	15.7425770604	12	8	307	130	70
2	14.3121099952	11.5	8	350	165	70
3	15.6234154812	11	8	318	150	70
4	15.4252472392	12	8	304	150	70
5	16.2722027548	10.5	8	302	140	70
6	12.7827595487	10	8	429	198	70
7	12.5939629733	9	8	454	220	70
8	12.8902617878	8.5	8	440	215	70
9	12.2600841390	10	8	455	225	70
10	14.1703737729	8.5	8	390	190	70
11	14.5664751351	10	8	383	170	70
12	15.8865953401	8	8	340	160	70
13	14.6568876141	9.5	8	400	150	70
14	13.5958238028	10	8	455	225	70
15	23.1964366694	15	4	113	95	70
16	19.2214089838	15.5	6	198	95	70
17	19.3103884607	15.5	6	199	97	70
18	20.0440082104	16	6	200	85	70
19	25.2330383736	14.5	4	97	88	70
20	27.5952444771	20.5	4	97	46	70

เสร็จสิ้น

รูปที่ 5.25 หน้าจอแสดงผลลัพธ์ที่ได้จากการพยากรณ์ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 6

### การทดสอบระบบ และสรุปผลการทดสอบ

จากการพัฒนาระบบবাদ้าไมน์นิ่ง โดยใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบแบคพรอพพาเกชัน สามารถทำการทดสอบระบบและสรุปผลการทดสอบได้ดังมีรายละเอียดต่อไปนี้

#### 6.1 การทดสอบระบบ

- แหล่งข้อมูล เป็นข้อมูลเกี่ยวกับการใช้เชื้อเพลิง (ไมล์ / แกลลอน) ของรถในแต่ละแบบ ซึ่งประกอบด้วย 391 รูปแบบ โดยค่าอินพุต คือ การเร่ง (acceleration) ลูกสูบ (cylinders) การทำให้เคลื่อนที่ (displacement) แรงม้า (horsepower) ปี (modelyear) รูปแบบ(origin) น้ำหนัก (weight) และค่าเอาต์พุต คือ อัตราเชื้อเพลิงที่ใช้ (mpg)

- จำนวนข้อมูลที่ใช้ในการสร้างแบบจำลอง 274 แถว

- จำนวนข้อมูลที่ใช้ในการสร้างทดสอบแบบจำลอง 117 แถว

การทดสอบที่ 1 ทำการปรับจำนวนรอบที่ใช้ในการสร้างแบบจำลอง โดยที่ตัวแปรอื่นๆ ของแบบจำลองเหมือนกัน

- ค่าอัตราการเรียนรู้ 0.1

- จำนวนโหนดในชั้นอินพุต 7 โหนด

- จำนวนโหนดในชั้นฮิดเดน 10 โหนด

ตารางที่ 6.1 แสดงผลจากการทดสอบที่ทำการปรับจำนวนรอบที่ใช้สร้างแบบจำลอง

ตัวแปรของแบบจำลอง	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
จำนวนรอบ	5000	10,000	15,000
ค่าความผิดพลาดเฉลี่ยของการเรียนรู้	0.0026283	0.0022076	0.0021856
ค่าความผิดพลาดเฉลี่ยของการทดสอบ	0.0119353	0.0117538	0.0113542

การทดสอบที่ 2 ทำการปรับโครงสร้างภายในโครงข่าย ได้แก่ จำนวนโหนดในชั้นฮิดเดน โดยที่ตัวแปรอื่นๆ ของแบบจำลองเหมือนกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

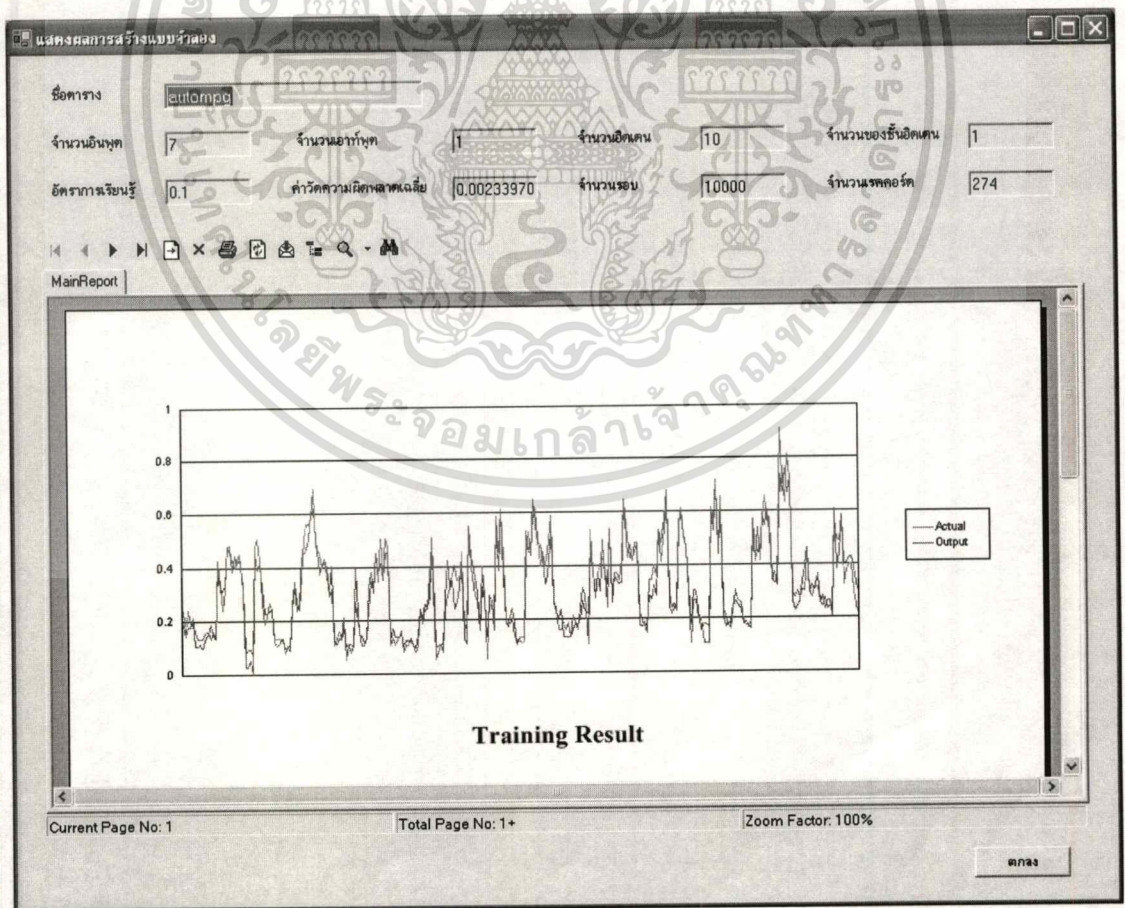
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- จำนวนรอบที่ใช้ในการสร้างแบบจำลอง 10,000 รอบ
- ค่าอัตราการเรียนรู้ 0.1
- จำนวนโหนดในชั้นอินพุต 7 โหนด

ตารางที่ 6.2 แสดงผลจากการทดสอบแบบจำลองที่ทำการปรับจำนวนโหนดในชั้นฮิดเดน

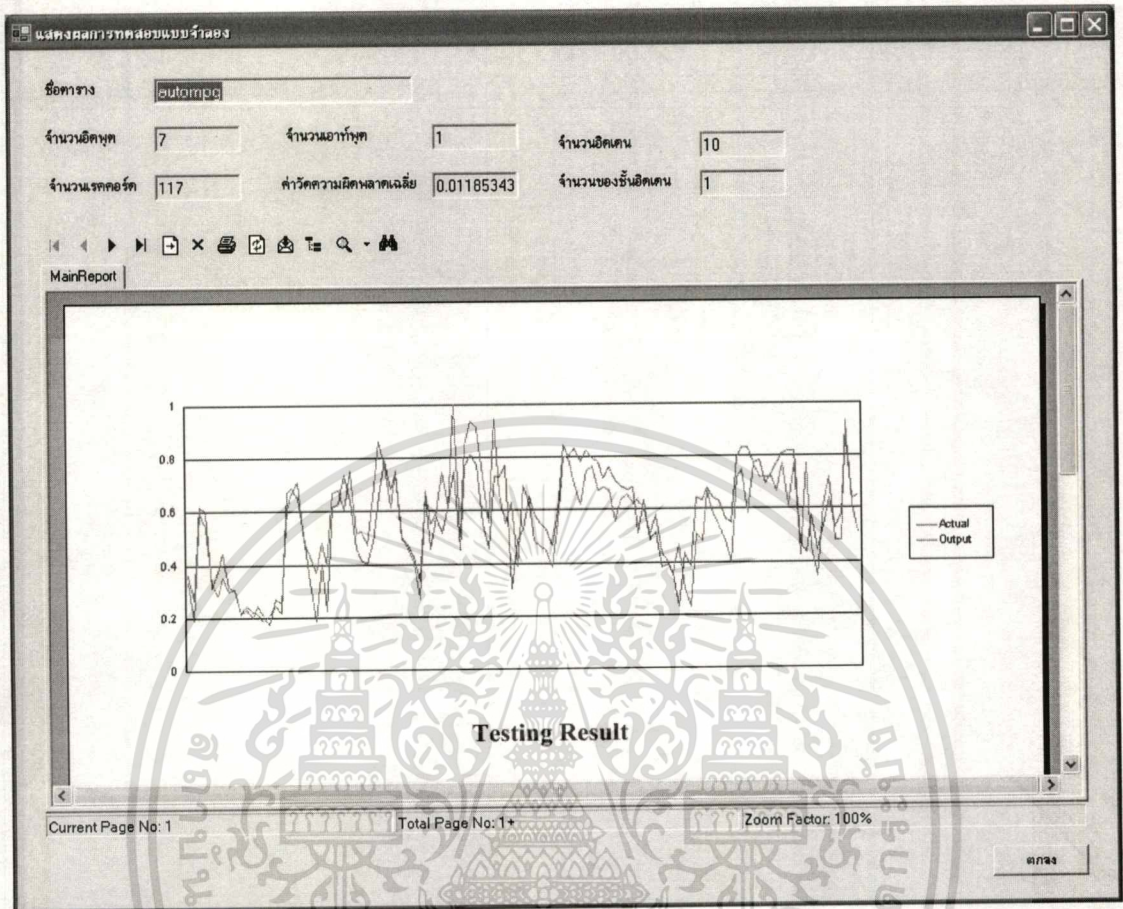
ตัวแปรของแบบจำลอง	ครั้งที่ 1	ครั้งที่ 2	ครั้งที่ 3
จำนวนโหนดในชั้นฮิดเดน	5	10	15
ค่าความผิดพลาดเฉลี่ยของการเรียนรู้	0.0024927	0.0023397	0.0022782
ค่าความผิดพลาดเฉลี่ยของการทดสอบ	0.0129329	0.0118534	0.0145878

จากการทดสอบสามารถแสดงตัวอย่างของผลที่ได้จากการสร้างแบบจำลอง และการทดสอบแบบจำลอง ดังรูปที่ 6.1 และ 6.2



รูปที่ 6.1 หน้าจอแสดงผลการสร้างแบบจำลองที่ทำการเรียนรู้จำนวน 10,000 รอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 6.2 หน้าจอแสดงผลการทดสอบแบบจำลองที่ทำการเรียนรู้จำนวน 10,000 รอบ

## 6.2 การสรุปผลการทดสอบ

จากการสร้างแบบจำลองโดยใช้อัลกอริทึมของโครงข่ายประสาทเทียมแบบแบคพรอพพา เกชัน สามารถสรุปได้ว่าค่าความผิดพลาดของการสร้างแบบจำลองและการทดสอบแบบจำลองนั้น จะขึ้นอยู่กับตัวแปรที่กำหนดให้กับระบบในการสร้างแบบจำลอง เช่น จำนวนรอบที่ใช้ในการสร้างแบบจำลอง และจำนวน โหนดในชั้นอิตศน เป็นต้น ซึ่งจะเห็นว่าหากจำนวนรอบที่ใช้ในการสร้างแบบจำลองยิ่งมาก ค่าความผิดพลาดจะยิ่งมีค่าต่ำลง แต่ในขณะเดียวกันการเพิ่มจำนวนรอบมากขึ้นก็ จะทำให้เวลาที่ใช้ในการสร้างแบบจำลองนานขึ้นตามไปด้วย ส่วนในการกำหนดจำนวน โหนดในชั้นอิตศนนั้น ขึ้นอยู่กับจำนวนของข้อมูลที่นำมาทำการสร้างแบบจำลอง โดยถ้าหากกำหนดจำนวน โหนดมากเกินไปเมื่อเทียบกับข้อมูลที่นำมาสร้างแบบจำลองก็อาจทำให้เกิดปัญหาว่า ค่าความผิดพลาดที่ได้จากการสร้างแบบจำลองนั้นมีค่าต่ำ แต่ค่าความผิดพลาดที่ได้จากการทดสอบแบบจำลองนั้นกลับมีค่าสูง ซึ่งทำให้แบบจำลองที่ได้ไม่เหมาะสมที่จะนำไปใช้งาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจัยอีกปัจจัยที่มีความสำคัญก็คือ ข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง เพราะแม้ว่าจะสามารถกำหนดตัวแปรที่ใช้สร้างแบบจำลองได้อย่างเหมาะสม แต่ถ้าข้อมูลที่นำมาใช้สร้างแบบจำลองไม่สมบูรณ์ ไม่ครอบคลุมก็จะทำให้ไม่สามารถสร้างแบบจำลองที่มีประสิทธิภาพได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บรรณานุกรม

cs.kku.ac.th. **Backpropagation In Neural Networks**. [Online]. Available:

<http://project.cs.kku.ac.th/2545/seminar/normal/8/download%5Cdoc.doc>

Gordon Linoff, Michael Berry. 2000. **Customer Relationship Management Through Data Mining**. Cambridge : SAS Institute Inc.

Jiawei Han, Micheline Kamber. 2001. **Data Mining Concept and Techniques**. San Franciscr : Morgan Kaufmann.

Michael Negnevitsky. 2002. **Artificial Intelligence**. England : Addison wesley.

Microsoft. 2005. **SQL Server 2005 Books Online**. [Online]. Available:

<http://www.microsoft.com/downloads/details.aspx?FamilyId=F0D182C1-C3AA-4CAC-B45C-BD15D9B072B7&displaylang=en>

## ประวัติผู้เขียน

ชื่อผู้เขียน	นางสาวปาณิสรา เย็นนะสา
วัน เดือน ปีเกิด	13 พฤศจิกายน พ.ศ.2524
สถานที่เกิด	จังหวัดกรุงเทพมหานคร
ประวัติการศึกษา	
ปริญญาตรี	ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้