

บทที่ 3

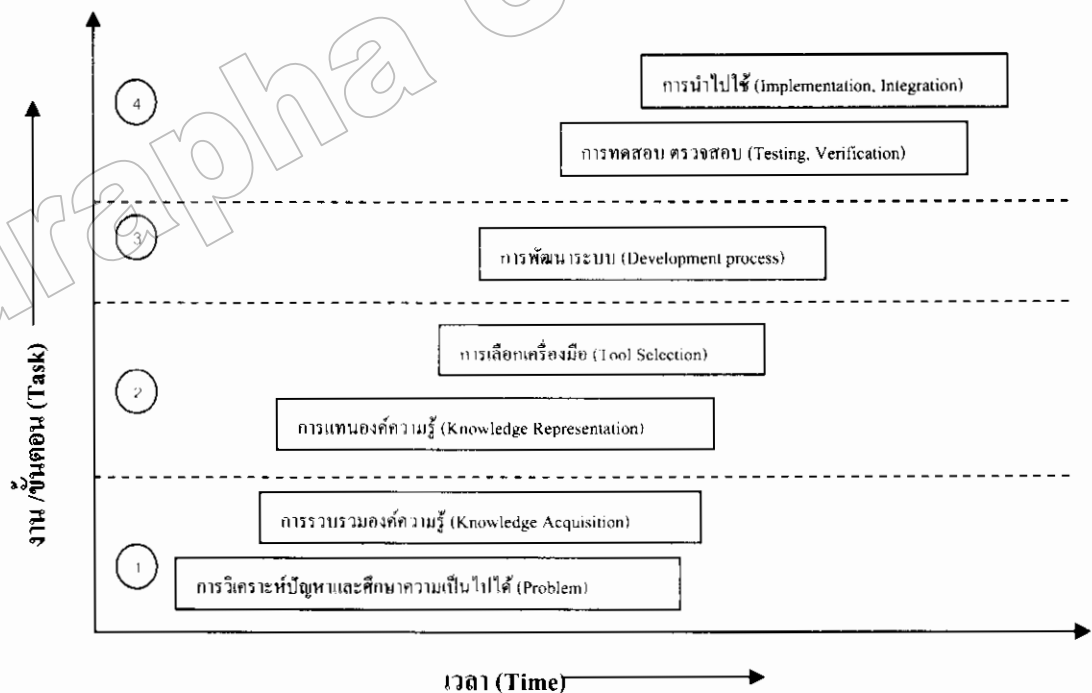
วิธีดำเนินการวิจัย

บทนำ

ในการวิจัยสร้างระบบผู้เชี่ยวชาญด้านการพยากรณ์ผู้วิจัยทำการกำหนดกระบวนการตามหลักการในการพัฒนาระบบผู้เชี่ยวชาญที่สำคัญ อันได้แก่

- 3.1 การวิเคราะห์และศึกษาแนวทางในการพัฒนาผู้เชี่ยวชาญ
- 3.2 การวิเคราะห์กฎผู้เชี่ยวชาญ
- 3.3 การสร้างต้นแบบตามโครงสร้างกฎที่ออกแบบไว้
- 3.4 การทดสอบ การตรวจสอบ และการนำไปใช้

แสดงโดยแผนภาพกระบวนการสำคัญในการพัฒนาระบบผู้เชี่ยวชาญดังต่อไปนี้



ภาพที่ 3.1 กระบวนการที่สำคัญในการพัฒนาระบบผู้เชี่ยวชาญ

3.1 การวิเคราะห์และศึกษาแนวทางในการพัฒนาภูมิปัญญา

ในขั้นตอนนี้เป็นการศึกษาและรวบรวมความรู้ข้อสรุปที่สำคัญเกี่ยวกับการพยากรณ์จากนักพยากรณ์ ข้อดีข้อเสีย ความเหมาะสมของทฤษฎีการพยากรณ์กับข้อมูลในรูปแบบต่าง ๆ และศึกษาความเป็นไปได้ในการพัฒนาระบบดังกล่าว

จากการศึกษาเอกสารงานวิจัยและหลักการในการพยากรณ์ประเด็นสำคัญที่จะช่วยให้การพยากรณ์มีความแม่นยำได้นั้นขึ้นอยู่กับ การเลือกเครื่องมือหรือวิธีการในการพยากรณ์ให้เหมาะสมกับลักษณะของข้อมูล ทั้งนี้เนื่องจากความต้องการมีการเปลี่ยนแปลงอยู่ตลอดเวลาตามปัจจัยภายนอก ซึ่งไม่อาจคาดการณ์ล่วงหน้าได้ ดังนั้นการปรับเครื่องมือการพยากรณ์ที่เหมาะสมตามปัจจัยที่เปลี่ยนแปลงโดยอัตโนมัติจึงเป็นสิ่งสำคัญ ยกตัวอย่างเช่น เมื่อแนวโน้มความต้องการสินค้าหรือการบริการมีค่าเพิ่มขึ้น/ ลดลงอย่างรวดเร็ว เราอาจปรับผลการพยากรณ์ตามอัตราค่าเพิ่มขึ้น/ ลดลงของข้อมูลที่เกิดขึ้นได้ หรือกล่าวอีกนัยหนึ่ง เราประยุกต์ใช้กฎในการปรับวิธีการพยากรณ์ตามลักษณะแนวโน้มของข้อมูลที่เปลี่ยนแปลงไป โดยมีงานวิจัยของ Adya et al. (2001) มีการพัฒนาภูมิปัญญาด้วยการตรวจสอบลักษณะเด่นของข้อมูลแบบอัตโนมัติ 6 ลักษณะ โดยในงานวิจัยนี้เราให้ความสนใจกับการตรวจสอบลักษณะรูปแบบข้อมูลที่สำคัญ 7 ลักษณะได้แก่

1. ข้อมูลที่มีฤดูกาล (Seasonality) เป็นข้อมูลที่มีลักษณะขึ้นลงตามช่วงจังหวะเวลาที่ใกล้เคียงกันในช่วงเวลาเดียวกัน การพยากรณ์จำเป็นต้องใช้โมเดลการพยากรณ์ที่สนับสนุนในเรื่องนี้ด้วย (สิริพร พูลสวัสดิ์ และคณะ, 2552)

2. การเกิดแนวโน้มเพิ่มขึ้นของข้อมูล (Positive trend) สินค้าหรือบริการบางอย่างมีลักษณะยอดขายเติบโตขึ้นอย่างต่อเนื่อง ด้วยองค์ประกอบที่สำคัญทั้งภายในและภายนอกต่าง ๆ ดังนั้นการกำหนดกฎขึ้นมาตรวจสอบข้อมูลในลักษณะนี้ เพื่อทำการปรับการพยากรณ์ให้มีความเหมาะสมกับรูปแบบข้อมูลที่เกิดขึ้น เนื่องจากบางวิธีอาจจะไม่เหมาะสมกับข้อมูลในลักษณะนี้ โดยเฉพาะเทคนิค MC ซึ่งมักจะให้ค่าการพยากรณ์ข้อมูลในลักษณะนี้ได้ไม่ทันต่อความเปลี่ยนแปลงที่เกิดขึ้น

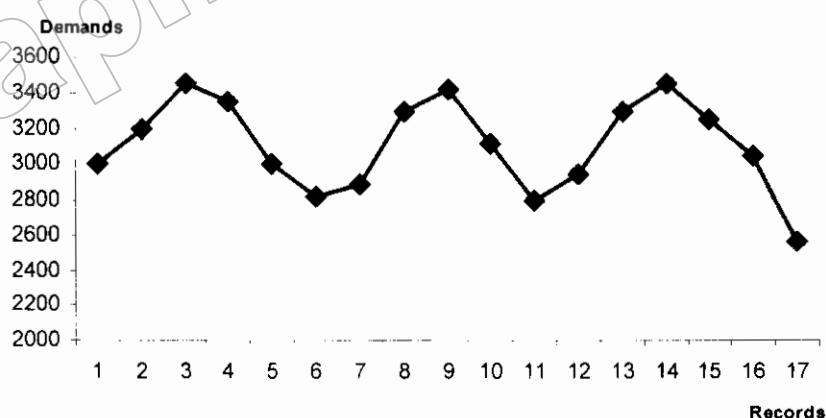
3. การเกิดแนวโน้มลดลงของข้อมูล (Negative trend) ซึ่งข้อมูลมีลักษณะตรงข้ามกับข้อ (2) แต่มีลักษณะการกำหนดกฎหรือการปรับวิธีพยากรณ์เช่นเดียวกัน

4. ข้อมูลที่มีลักษณะ โคตมากผิดปกติ (Outlier) เป็นการตั้งกฎขึ้นเพื่อตรวจสอบข้อมูลที่มีการเพิ่มขึ้น/ ลดลงอย่างผิดสังเกตในบางจุดของข้อมูล ข้อมูลในลักษณะนี้เป็นข้อมูลที่ไม่เกิดขึ้นบ่อย อาจเกิดจากผลกระทบภายนอกหรือภายในบางประการทำให้ข้อมูลเกิดความผิดปกติ ดังนั้น จึงควรปรับค่าข้อมูลเหล่านี้ให้มีลักษณะใกล้เคียงกับรูปแบบข้อมูลพื้นฐานเดิม

5. การเปลี่ยนระดับข้อมูล (Level shift) เป็นการตั้งกฎขึ้นเพื่อตรวจสอบข้อมูลที่มีระดับ (Level) ของข้อมูลผิดไปจากข้อมูลในอดีตอย่างชัดเจน โดยแบ่งเป็น 2 ลักษณะที่สำคัญคือ ข้อมูลที่มีระดับเพิ่มขึ้นและลดลงจากเดิม ดังนั้น การนำข้อมูลไปใช้ในการพยากรณ์จึงควรตัดข้อมูลเดิมช่วงที่ยังไม่มีการเปลี่ยนแปลงระดับออก และใช้เฉพาะข้อมูลช่วงที่มีการเปลี่ยนแปลงระดับแล้วมาพยากรณ์ เพื่อไม่ให้ข้อมูลรูปแบบเดิมมามีผลต่อการพยากรณ์ข้อมูลที่มีลักษณะเปลี่ยนแปลงไป

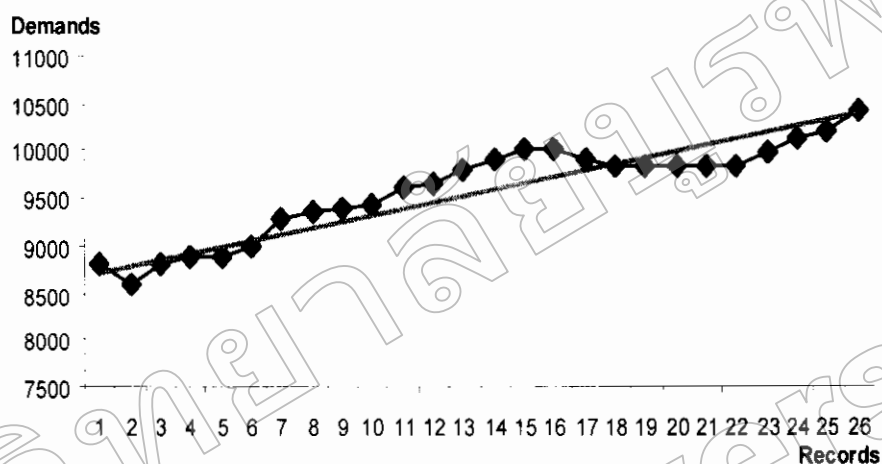
6. ข้อมูลที่มีลักษณะ ไม่น่าสนใจ (Unusual) เป็นการตั้งกฎขึ้นเพื่อตรวจสอบข้อมูลที่มีลักษณะขึ้น ๆ ลง ๆ มีลักษณะแปรปรวนมากและมีลักษณะการเกิดข้อมูลเป็นเชิงสุ่มดังนั้นจึงควรใช้วิธีการในการพยากรณ์โดยอาศัยวิธีหลักความน่าจะเป็นแบบไม่ต่อเนื่อง

7. ข้อมูลที่มีลักษณะค่อนข้างแน่นอน (Usual) เป็นการตั้งกฎขึ้นเพื่อตรวจสอบข้อมูลที่มีลักษณะการเกิดในระดับที่ใกล้เคียงกัน โดยเกิดขึ้นลงเล็กน้อยในแนวราบ ซึ่งเป็นรูปแบบข้อมูลที่ไม่จำเป็นต้องใช้วิธีการพยากรณ์ที่ซับซ้อนมากนัก



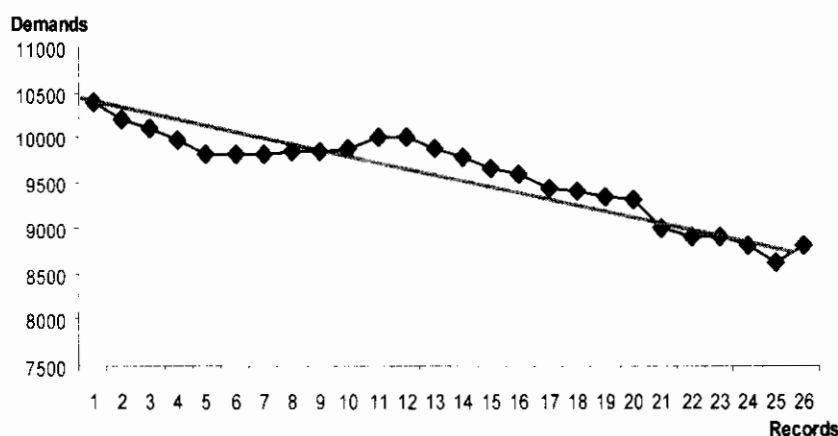
ภาพที่ 3.2 ข้อมูลที่มีฤดูกาล (Seasonality)

จากภาพที่ 3.2 แสดงตัวอย่างลักษณะข้อมูลที่มีลักษณะของฤดูกาล ซึ่งมีลักษณะการเกิดหรือความต้องการสินค้า/ บริการในลักษณะที่ใกล้เคียงกันในช่วงระยะเวลาหรือรอบเวลาเดียวกันในแต่ละปี จากภาพแสดงข้อมูลมีลักษณะเพิ่มขึ้นลดลงในทุก ๆ ช่วง 6 ระเบียบ (สิริพร พูลสวัสดิ์ และคณะ, 2552)



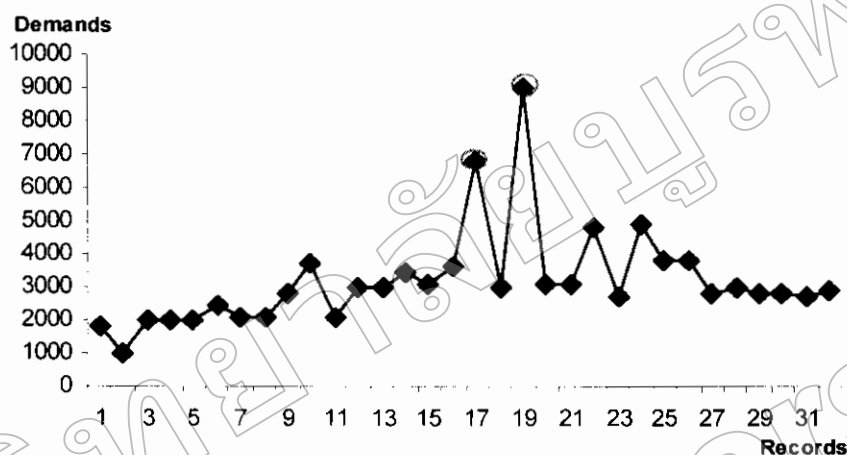
ภาพที่ 3.3 การเกิดแนวโน้มเพิ่มขึ้นของข้อมูล (Positive trend)

จากภาพที่ 3.3 แสดงลักษณะข้อมูลที่มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง จากภาพจะเห็นว่าเส้นตรงใช้แทนแนวโน้มข้อมูลมีค่าความชันไปในทางบวก ซึ่งอัตราการเติบโตของแนวโน้มขึ้นอยู่กับระดับความชันของข้อมูลด้วย ข้อมูลในลักษณะนี้แสดงถึงยอดขายที่มีแนวโน้มเพิ่มขึ้นตามช่วงเวลา



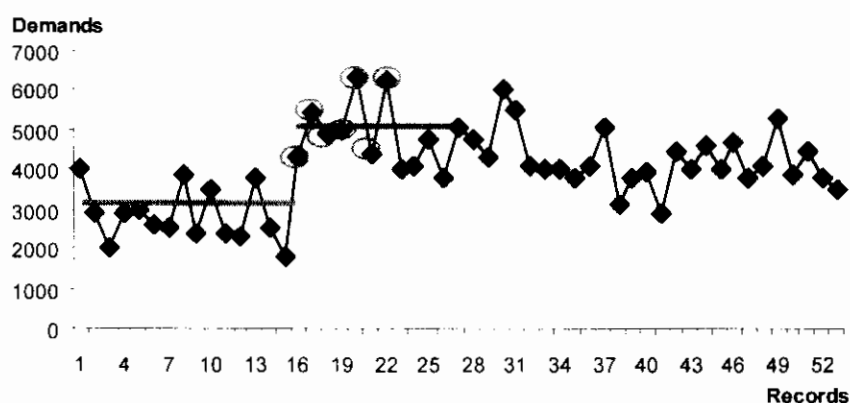
ภาพที่ 3.4 การเกิดแนวโน้มลดลงของข้อมูล (Negative trend)

และในภาพที่ 3.4 ข้อมูลมีลักษณะตรงข้ามกับข้อมูลในภาพที่ 3.3 โดยข้อมูลมีลักษณะแนวโน้มลดลงอย่างต่อเนื่อง จากภาพจะเห็นได้ว่าเส้นตรงที่ใช้แทนแนวโน้มข้อมูลมีค่าความชันไปในทางลบ ซึ่งอัตราการลดลงของแนวโน้มขึ้นอยู่กับระดับความชันของข้อมูลด้วย ข้อมูลในลักษณะนี้แสดงถึงยอดขายที่มีแนวโน้มลดลงเรื่อย ๆ ตามช่วงเวลา



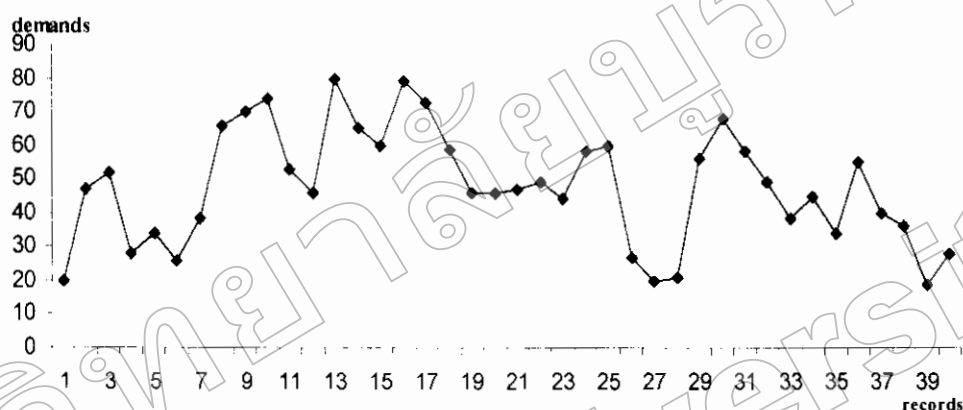
ภาพที่ 3.5 ลักษณะ โดคมากผิดปกติ (Outlier) ของข้อมูล

จากภาพที่ 3.5 แสดงตัวอย่างข้อมูลที่มีโดคมากผิดปกติของระยะตอนที่ 17 และ 19 ซึ่งเป็นข้อมูลที่มีลักษณะผิดไปจากฐานข้อมูลเดิมค่อนข้างมากและเห็นได้ชัดเจนอาจเกิดขึ้นนาน ๆ ครั้ง โดยเกิดจากปัจจัยภายนอกบางอย่างทำให้ยอดขายเพิ่มมากขึ้นจนผิดปกติหรือลดลงมากผิดปกติ โดยทั่วไปข้อมูลแบบนี้มักเป็นปัญหาต่อการนำมาเป็นฐานข้อมูลที่ใช้ในการพยากรณ์ เพราะจะทำให้ผลการพยากรณ์คลาดเคลื่อนได้ง่าย



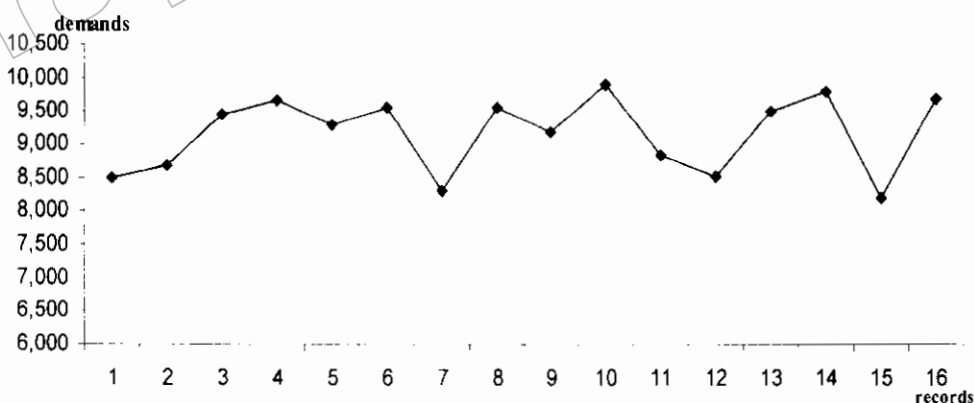
ภาพที่ 3.6 ลักษณะการเปลี่ยนแปลงระดับ (Level Shift) ของข้อมูล

จากภาพที่ 3.6 แสดงข้อมูลที่มีการเปลี่ยนแปลงระดับของข้อมูลไปจากเดิมโดยสิ้นเชิง จากภาพจะเห็นได้ว่าข้อมูลเคยอยู่ในระดับ 2,000 ถึง 4,000 ในช่วง 16 ระเบียบแรก แต่หลังจากนั้น ข้อมูลกลับอยู่ในช่วง 4,000 ถึง 6,000 ในช่วงหลังข้อมูลในลักษณะนี้อาจเกิดจากการที่สินค้ามีการพัฒนาจนผู้ซื้อหันมานิยม หรือการได้ลูกค้ากลุ่มใหม่เป็นกลุ่มใหญ่ทำให้ยอดขายเดิมเปลี่ยนแปลงไปจากระดับเดิม



ภาพที่ 3.7 ข้อมูลที่มีลักษณะไม่แน่นอน (Unusual)

จากภาพที่ 3.7 แสดงข้อมูลที่มีลักษณะขึ้น ๆ ลง ๆ ไม่แน่นอนหรือมีลักษณะการเกิดเป็นเชิงสุ่ม เช่นในบางช่วงของข้อมูลระเบียบที่ 3 ถึง 6 ข้อมูลมีค่าลดลงและเมื่อถึงระเบียบที่ 7 ถึง 10 ข้อมูลกลับเพิ่มขึ้นเรื่อย ๆ และข้อมูลเริ่มตกลงอีกในช่วงที่ 11 และ 12



ภาพที่ 3.8 ข้อมูลที่มีลักษณะแน่นอน (Usual)

จากภาพที่ 3.8 แสดงข้อมูลที่มีลักษณะขึ้นลง ไม่แตกต่างกันมากนัก จากภาพจะเห็นว่าข้อมูลขึ้นลงแตกต่างกันอยู่ในช่วง 8,000 กว่าถึง 9,000

3.2 การวิเคราะห์กฎผู้เชี่ยวชาญ

ในส่วนของการวิเคราะห์ระบบประกอบไปด้วยส่วนงานและขั้นตอนสำคัญต่าง ๆ ได้แก่ การเลือกเครื่องมือที่ใช้ในการพัฒนาระบบ ชุดข้อมูลที่ใช้ในการทดลองการกำหนดข้อมูลในส่วนของการนำเข้า (Input) และส่วนของการส่งออก (Output) การถอดความรู้ออกเป็นโมเดลการพยากรณ์ที่ใช้ในงานวิจัยด้วยรูปแบบชุดของกฎการพยากรณ์ และเพิ่มเติมในส่วนของการปรับปรุงโมเดลการพยากรณ์บางตัวให้มีประสิทธิภาพมากขึ้น แบ่งออกเป็นหัวข้อสำคัญดังต่อไปนี้

3.2.1 เครื่องมือที่ช่วยในการสร้างโปรแกรม ผู้วิจัยเลือกใช้โปรแกรมภาษา Visual Basic for Application โดยทำงานผ่านโปรแกรมตารางงาน เนื่องจากเป็นโปรแกรมที่มีใช้กันโดยทั่วไป เพียงเขียนกลไกการอนุมานเกี่ยวกับกฎการพยากรณ์ ผังไว้ที่ตัวโปรแกรมแล้วเรียกใช้งานผ่านโปรแกรมหาคำว่า

3.2.2 ชุดข้อมูลที่ใช้ในงานวิจัย ชุดของข้อมูลที่ใช้ในการวิจัยครั้งนี้แบ่งออกเป็นสองส่วนสำคัญได้แก่

ส่วนแรก เป็นข้อมูลทางธุรกิจเพื่อใช้ในการทดลองเบื้องต้น ได้แก่ ข้อมูลยอดขายรายรถยนต์ยี่ห้อโตโยต้า (เข้าถึงได้จาก <http://www.toyota.co.th>) ซึ่งข้อมูลที่มีลักษณะเติบโตขึ้นเรื่อย ๆ ข้อมูลยอดขายผักกะหล่ำปลีจากห้างค้าปลีกสมัยใหม่แห่งหนึ่ง (จิรัฐ สงวัฒนา, 2548) ซึ่งข้อมูลมีลักษณะขึ้นลง ไม่แน่นอน ความต้องการใช้บริการรถประจำทาง สาย 4 (ไม่ปรับอากาศ) ในกรุงเทพฯ ช่วงวันจันทร์ และวันอาทิตย์ (หน่วย : คน) ตามลำดับ ซึ่งข้อมูลวันจันทร์มีลักษณะค่อนข้างคงที่ และข้อมูลวันอาทิตย์มีความแปรปรวน และ ข้อมูลจาก NN3 (Artificial Neural Networks & Computational Intelligence Forecasting Competition , 2007) ชุด A11 และ A20 ซึ่งเป็นข้อมูลทางธุรกิจแบบรายเดือน โดยข้อมูลชุด A11 มีลักษณะเป็นแนวโน้มเพิ่มขึ้นอย่างช้า ๆ ไม่ชัดเจนมากนัก และ A20 มีลักษณะเป็นแนวโน้มขึ้นลงไม่แน่นอน (เข้าถึง ได้จาก <http://www.neuralforecasting-competition.com>)

ส่วนที่สอง ข้อมูลในการวิจัยครั้งนี้เป็นข้อมูลจาก M-3 Competition ซึ่งมีข้อมูลทั้งหมด 3,003 ระเบียบ ซึ่งข้อมูลมีทั้งรายปี รายสามเดือน รายเดือน และอื่น ๆ รวม 4 ชุดสำคัญ โดยแต่ละชุดประกอบไปด้วย ข้อมูลประเภทต่าง ๆ ทางธุรกิจ อีก 6 รูปแบบ ได้แก่ ข้อมูลจุลภาค (Micro) ข้อมูลด้านอุตสาหกรรม (Industry) ข้อมูลมหภาค (Macro) ข้อมูลการเงิน (Finance) ข้อมูลประชากร (Demographic) และข้อมูลอื่น ๆ (Other) โดยในการทดลองครั้งนี้เลือกทดลองเฉพาะข้อมูลราย

เดือนซึ่งมีลักษณะของข้อมูลที่หลากหลายและแตกต่างกันทางสถิติมากกว่าข้อมูลชุดอื่น ๆ ก่อนข้างมาก (Flores & Pearce, 2000) รวม 1,428 ชุดข้อมูลดังรายละเอียดตามตารางที่ 3.1

ตารางที่ 3.1 รายละเอียดข้อมูลรายเดือนของ M3-Competition

Type of Time Series Data						
M3-Competition (Monthly)						
Micro	Industry	Macro	Finance	Demographic	Other	Total
474	334	312	145	111	52	1,428

ในส่วนของพารามิเตอร์นำเข้าที่กำหนดให้ระบบผู้เชี่ยวชาญประกอบไปด้วย

- 1) Sampling: หรือการกำหนดจำนวนครั้งในการจำลอง (สุ่ม) ข้อมูล การจำลองสถานการณ์ของเทคนิคมอนติคาร์โล
 - 2) Decay: เป็นการกำหนดจำนวนการหาความสัมพันธ์ของข้อมูลในอดีตที่ผ่านมา เพื่อคัดเลือกข้อมูลที่จะนำมาใช้ในการจำลองสถานการณ์ โดยเรากำหนดข้อมูลที่จะมีความสัมพันธ์เมื่อไม่เกิดข้อมูลที่สังเกตติดต่อกัน 30 ครั้ง
 - 3) Outlier: เป็นการกำหนดจำนวนข้อมูลที่ใช้ในการตรวจสอบความผิดปกติของข้อมูลที่เกิดขึ้นโดยใช้ข้อมูลก่อนหน้า และหลังข้อมูลระยะเบี่ยง ที่สังเกตอย่างละ 6 ระยะเบี่ยง รวมเป็น 12 ระยะเบี่ยง
 - 4) Trend: เป็นการกำหนดจำนวนข้อมูลที่ใช้ในการสังเกตการเติบโต/ ลดลงของข้อมูลอย่างต่อเนื่อง เพื่อตรวจสอบแนวโน้มของข้อมูล โดยกำหนดจำนวนการสังเกตที่ระยะเบี่ยงก่อนหน้าว่ามีการเพิ่มขึ้น/ ลดลงอย่างต่อเนื่อง ติดต่อกัน 5 ครั้งขึ้นไปหรือไม่
 - 5) Level shift: เป็นการกำหนดจำนวนการเกิด Outlier ติดต่อกัน เพื่อตรวจสอบลักษณะการเปลี่ยนแปลงระดับข้อมูล โดยกำหนดจำนวนการสังเกตอยู่ที่การเกิด Outlier ติดต่อกัน 6 ระยะเบี่ยง
 - 6) Seasonality: เป็นการกำหนดค่าที่สำคัญทางสถิติสำหรับทดสอบค่าไคสแควร์ (Chi-Square) ในการทดสอบค่าดัชนีฤดูกาลของข้อมูลแต่ละเดือน
- ค่าพารามิเตอร์ที่ใช้ในกฎการพยากรณ์สามารถสรุปได้ดังตารางที่ 3.2

ตารางที่ 3.2 ค่าพารามิเตอร์ที่ใช้ในกฎการพยากรณ์

No.	Parameters	Symbols	Description	Adjustment
1	Sampling	N	ปรับตามข้อมูลที่เปลี่ยนแปลงไปตามเวลา	Adaptive
2	Decay	D	ข้อมูลเสื่อมเมื่อไม่เกิดขึ้นติดต่อกัน 30 ครั้ง	Static
3	Outlier	O	ใช้ข้อมูล 6 ระเบียบ หน้าหลัง ตัวที่สังเกต	Static
4	Trend	T	เติบโต/ลดลง ติดต่อกัน 5 ครั้ง ถือว่าเกิดแนวโน้ม	Static
5	Level Shift	L	เกิด Outlier ติดต่อกัน 6 ระเบียบ	Static
6	Seasonal	S	ใช้ค่าไคสแควร์ (Chi-Square Test) ในการทดสอบค่าดัชนีฤดูกาลของข้อมูลแต่ละเดือนที่ระดับนัยสำคัญ $\alpha = 0.01$	Static

จากการกำหนดค่าพารามิเตอร์ในตารางที่ 3.2 มีพารามิเตอร์ที่สำคัญ 6 พารามิเตอร์สำหรับใช้ในส่วนของการจำลองข้อมูลประกอบไปด้วยพารามิเตอร์ในการหาจำนวนครั้งที่เหมาะสมในการจำลองสถานการณ์ข้อมูล (Sampling : N) ซึ่งค่าข้อมูลมีการเปลี่ยนแปลงไปตามเวลา (Adaptive) และส่วนของพารามิเตอร์ที่ใช้ในการกำหนดระยะข้อมูลที่เกิดความเสื่อม (Decay : D) คือข้อมูลที่ไม่เกิดขึ้นติดต่อกัน 30 ครั้ง (Static) ในทุก ๆ การจำลอง และในส่วนที่เหลือเป็นพารามิเตอร์ที่กำหนดขึ้นเพื่อใช้ในการตรวจสอบลักษณะของข้อมูล (Feature) ทดลองทุกชุดข้อมูลซึ่งใช้ค่าพารามิเตอร์เดียวกัน (Static) โดยการตรวจสอบลักษณะความผิดปกติของข้อมูล (Outlier : O) ใช้ข้อมูล 6 ระเบียบ หน้าหลังข้อมูลที่สังเกตความผิดปกติ การตรวจสอบแนวโน้มข้อมูล (Trend : T) กำหนดให้ข้อมูลที่มีการเพิ่มขึ้นหรือลดลง 5 ครั้งติดต่อกันเป็นข้อมูลที่มีแนวโน้ม การตรวจสอบข้อมูลที่มีลักษณะการเปลี่ยนแปลงระดับ (Level Shift : L) กำหนดให้ดูจากการเกิดความผิดปกติของข้อมูลติดต่อกัน 6 ครั้ง และข้อมูลที่มีฤดูกาล (Seasonal : S) ใช้การทดสอบด้วยค่าไคสแควร์โดยใช้ค่าระดับนัยสำคัญ α เท่ากับ 0.01

3.2.3 กำหนดชุดของกฎการพยากรณ์หรือการถอดความรู้ ชุดของกฎการพยากรณ์นี้เป็น การนำแนวความคิดที่เกี่ยวข้องกันในการพยากรณ์มาทำการจัดเรียงลำดับขั้นตอนที่จำเป็น การไหลเวียนของตรรกะปัญหาตามเงื่อนไขต่าง ๆ ที่ได้ศึกษาและรวบรวมไว้ โดยในการออกแบบการวิจัยครั้งนี้แบ่งกฎผู้เชี่ยวชาญออกเป็นกฎในส่วนของการตรวจสอบรูปแบบข้อมูล และกฎในส่วนของการพยากรณ์ข้อมูล ดังนี้

กฎการตรวจสอบรูปแบบของข้อมูล (Demand pattern and detector rules)

กฎที่ใช้ในการพยากรณ์ประกอบด้วยกฎการตรวจสอบลักษณะของข้อมูลที่เกิดขึ้นในอดีต (Detector rules) และกฎการปรับผลการพยากรณ์ (Forecasting rules) ในงานวิจัยนี้ เราสร้างกฎการตรวจสอบข้อมูลที่เกิดขึ้นในอดีต (Detector rules) ได้แก่ Seasonal, Trend, Outlier, Level shift และ Unusual/ Usual ดังนี้

Detector rules:

R1-D. Seasonal

IF ข้อมูลมีลักษณะเป็นรูปแบบตามช่วงเวลา

Then ข้อมูลเกิด Seasonal

R2-D. Outlier

IF พบข้อมูลที่มีการเพิ่มขึ้น/ ลดลง อย่างผิดสังเกตในบางจุดของข้อมูล

Then ข้อมูลเกิด Outlier

R3-D. Level Shift

IF พบข้อมูลเกิด Outlier ติดต่อกันเกิน 4 ครั้ง

Then ข้อมูลเกิด LevelShift

R4-D. Changing basic trend

IF ข้อมูลมีแนวโน้มเพิ่มขึ้น/ ลดลงอย่างต่อเนื่อง

Then ข้อมูลเกิดการเปลี่ยนแปลงแนวโน้มพื้นฐาน

R5-D. Unusual

IF ข้อมูลมีค่าความแปรปรวนสูง

Then ข้อมูลเกิด Unusual

R6-D. Usual

IF ข้อมูลมีลักษณะขึ้นลงเล็กน้อยในแนวนราบ

Then ข้อมูลแบบ Usual

กฎการพยากรณ์ข้อมูล (Forecasting rules)

Forecasting rules:

R1-F. IF ข้อมูลเกิด Seasonal

Then ค่าพยากรณ์ = Seasonal index * Deseason value

R2-F. IF พบ Outlier

Then แทนค่าข้อมูลในจุดที่เกิด Outlier ด้วย ค่าเฉลี่ย $\pm$ ค่าเบี่ยงเบนมาตรฐาน
ตรวจสอบลักษณะข้อมูลด้วยกฎ **R3-D**.

ค่าพยากรณ์ ข้อมูลด้วย MCRB

R3-F. IF พบ Level Shift

Then ตัดข้อมูลส่วนก่อนเกิด Level Shift ออก

ค่าพยากรณ์ ข้อมูลด้วย MCRB

R4-F1. IF ข้อมูลเกิดแนวโน้มเพิ่มขึ้น

Then ค่าพยากรณ์ = ค่าข้อมูลล่าสุดที่เกิดขึ้น (X_t) + ค่าจากอัตราการเพิ่มขึ้น
(Average increase rate)

R4-F2. IF ข้อมูลเกิดแนวโน้มลดลง

Then ค่าพยากรณ์ = ค่าข้อมูลล่าสุดที่เกิดขึ้น (X_t) - ค่าจากอัตราการลดลง
(Average decrease rate)

R5-F. IF ข้อมูลเกิด Unusual

Then ค่าพยากรณ์ ข้อมูลด้วย MCRB

R6-F. IF ข้อมูลเกิด Usual

Then ค่าพยากรณ์ ข้อมูลด้วย MA

R1-D. เป็นกฎในการตรวจสอบ Seasonal ด้วยวิธีทางสถิติโดยใช้ค่าไคสแควร์ (Chi-Square-Test) ในการทดสอบค่าดัชนีฤดูกาลของข้อมูลแต่ละเดือน ในการทดลองนี้กำหนดขอบเขตการตัดสินใจโดยเลือกระดับนัยสำคัญ $\alpha = 0.01$ และใช้วิธีการพยากรณ์กรณีพบว่าเป็นข้อมูลที่มีฤดูกาลโดยใช้วิธีค่าดัชนีฤดูกาลในการพยากรณ์ด้วยกฎการพยากรณ์ **R1-F**.

R2-D. เป็นกฎในการตรวจสอบ Outlier ของข้อมูลที่มีลักษณะโคคมากผิดปกติ ซึ่งสามารถตรวจสอบได้โดยใช้ค่าส่วนเบี่ยงเบนมาตรฐาน, σ เรากำหนดค่าพารามิเตอร์ของจำนวนข้อมูลที่เกิดขึ้นในอดีต n_{r2} เพื่อใช้หาค่า σ และจัดการข้อมูลที่มีลักษณะนี้ด้วยกฎ **R2-F**.

R3-D. เป็นกฎในการตรวจสอบ Level Shift ของข้อมูลที่มีลักษณะ โคดมากผิดปกติ ติดต่อกันตามพารามิเตอร์ที่กำหนด ซึ่งสามารถตรวจสอบได้โดยใช้ผลของกฎใน R2-D. และจัดการข้อมูลที่มีลักษณะนี้ด้วยกฎ R3-F.

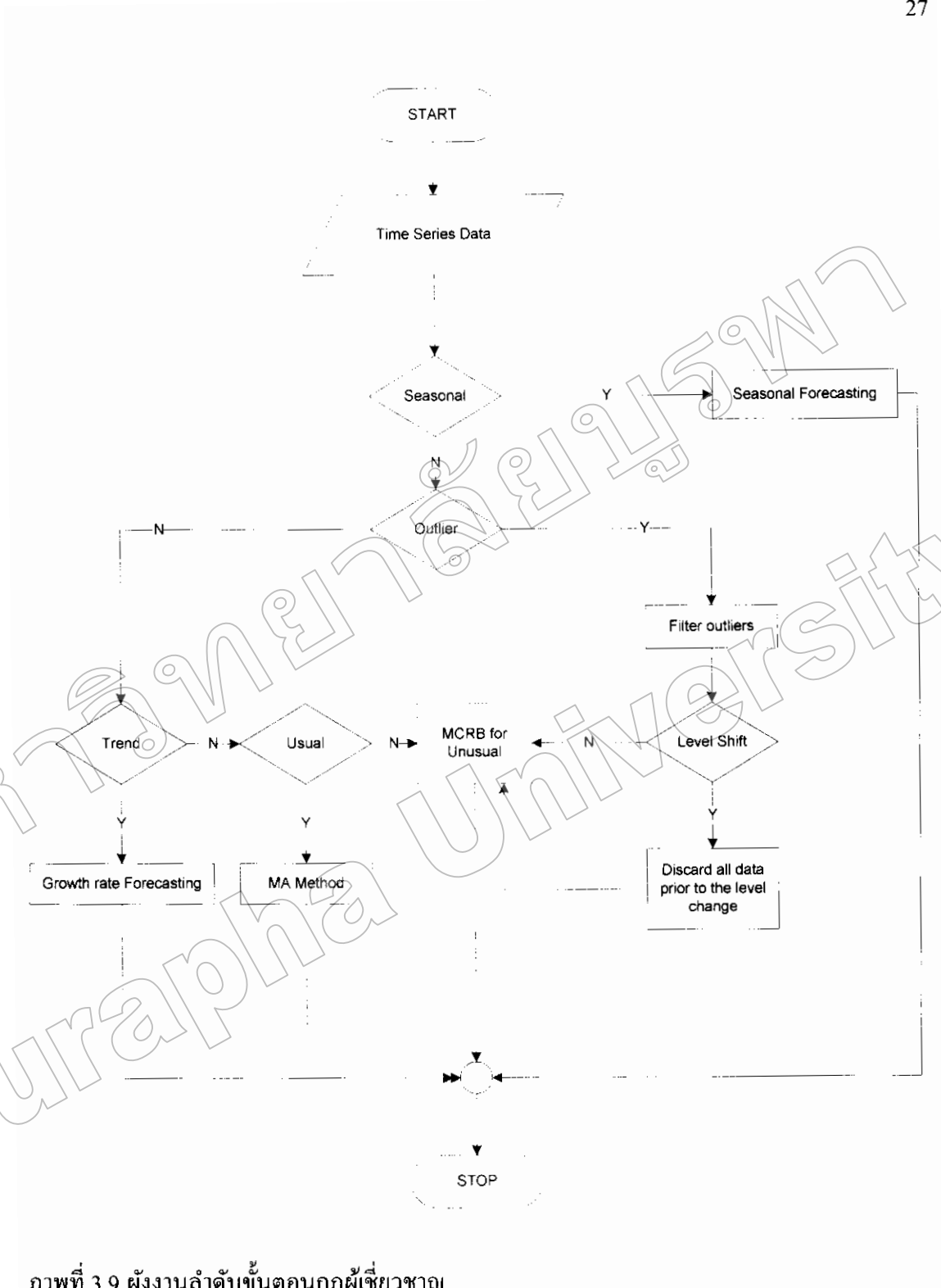
R4-D. เป็นกฎในการตรวจสอบการเกิดแนวโน้มของข้อมูลเป็นลักษณะการเกิดการเปลี่ยนแปลงแนวโน้มพื้นฐาน โดยข้อมูลอาจเปลี่ยนแปลงในทิศทางที่เพิ่มขึ้นหรือลดลง เรากำหนดค่าพารามิเตอร์ของจำนวนข้อมูลที่เกิดขึ้นในอดีต n_{r4} ว่ามีการเพิ่มขึ้นหรือลดลงอย่างต่อเนื่อง ติดต่อกันเป็นจำนวนเท่าไร ยกตัวอย่างเช่น เรากำหนด $n_{r4} = 2$ ถ้าข้อมูลเพิ่มขึ้น 2 ระเบียบต่อเนื่องกันถือว่าเกิดแนวโน้มเพิ่มขึ้น และใช้กฎการพยากรณ์ R4-F1. เพื่อพยากรณ์ข้อมูลตามอัตราการเพิ่มขึ้นของข้อมูล แต่ถ้าหากข้อมูลมีแนวโน้มลดลงจะใช้กฎการพยากรณ์ R4-F2. เพื่อพยากรณ์ข้อมูลตามอัตราการลดลง

R5-D. เป็นกฎในการตรวจสอบ Unusual ของข้อมูลโดยใช้ค่าความแปรปรวนของข้อมูล เป็นพารามิเตอร์ในการตัดสินใจ และใช้วิธีการพยากรณ์กรณีพบว่าเป็นข้อมูลที่มี Unusual ด้วยวิธี Monticaro Simulation ด้วยกฎการพยากรณ์ R5-F.

R6-D. เป็นกฎในการตรวจสอบ Usual ของข้อมูลโดยใช้ค่าความแปรปรวนของข้อมูล เป็นพารามิเตอร์ในการตัดสินใจ และใช้วิธีการพยากรณ์กรณีพบว่าเป็นข้อมูลที่มี Usual ด้วยวิธี Moving Average ด้วยกฎการพยากรณ์ R6-F.

โดยขั้นตอนการพยากรณ์ในภาพรวมแสดงดังภาพที่ 3.9 เริ่มต้นจากการตรวจสอบลักษณะฤดูกาลของข้อมูลซึ่งกำหนดการตรวจสอบเป็นขั้นตอนแรกตามงานวิจัยของ Monica et al. (2000) และจากการศึกษาเพิ่มเติมพบว่าควรตรวจสอบฤดูกาลก่อนการตรวจสอบข้อมูลที่มีความผิดปกติเนื่องจากหากข้อมูลมีการกำจัดข้อมูลที่ผิดปกติแล้วอาจทำให้ไม่พบฤดูกาลข้อมูล โดยถ้าพบข้อมูลเป็นแบบฤดูกาล จะใช้วิธีพยากรณ์ด้วยวิธีค่าดัชนีฤดูกาล (Seasonal Forecasting) หากไม่พบก็ จะใช้กฎในการตรวจสอบลักษณะของข้อมูลในลักษณะอื่น ๆ นั่นคือ การตรวจสอบข้อมูลที่มีลักษณะ โคดมากผิดปกติ (Outlier) เป็นลำดับถัดมาเพื่อจัดการข้อมูล (Filter Outlier) ให้มีความเหมาะสมในการตรวจสอบลักษณะอื่นและพยากรณ์ต่อเนื่องจากหากมีข้อมูลที่มีลักษณะผิดปกติ มากเกินไปจะดึงหรือมีส่วนให้ผลการพยากรณ์คลาดเคลื่อน ส่วนของการเกิดการเปลี่ยนแปลงระดับข้อมูล (Level Shift) เป็นส่วนที่ใช้ในการตรวจสอบกรณีที่พบความผิดปกติของข้อมูลเนื่องจากการตรวจสอบการเปลี่ยนแปลงระดับข้อมูลในงานวิจัยนี้อาศัยลักษณะความผิดปกติของข้อมูล และการตรวจสอบข้อมูลที่มีลักษณะเกิดขึ้นไม่แน่นอน (Unusual) ซึ่งจะทำการพยากรณ์ด้วยวิธี MCRB (Monte Carlo Simulation ร่วมกับวิธี กฎผู้เชี่ยวชาญ) หรือ หากข้อมูลไม่มีลักษณะ โคดมากผิดปกติ จะทำการตรวจสอบการมีแนวโน้มขึ้น/ลงอย่างต่อเนื่อง (Trend) ตามงานวิจัยของ Benito

and Stephen (2000) ซึ่งงานวิจัยนี้ได้นำเสนอไว้ในการตรวจสอบลักษณะของข้อมูลต่อการตรวจสอบลักษณะผิดปกติของข้อมูล และใช้วิธีการพยากรณ์ด้วยอัตราการเติบโตเพิ่มขึ้น/ ลดลง (สิริพร พูลสวัสดิ์ และคณะ, 2551b) และหากข้อมูลมีลักษณะขึ้นลงเล็กน้อย (Usual) ก็จะใช้วิธี MA (Moving Average) โดยลำดับขั้นตอนทั้งหมดที่ได้ทำการออกแบบนี้ผู้วิจัยได้ทำการทดลองสลับขั้นตอนในบางขั้นตอนเช่นการตรวจสอบแนวโน้มข้อมูลก่อนการตรวจสอบฤดูกาล หรือการตรวจสอบฤดูกาลก่อนการตรวจสอบความผิดปกติของข้อมูลกับข้อมูลในลักษณะต่าง ๆ เพื่อทดสอบประสิทธิภาพในการตรวจสอบข้อมูลในลักษณะต่าง ๆ จนได้ลำดับขั้นตอนที่มีความเหมาะสมและเป็นไปตามข้อเสนอแนะของงานวิจัยก่อนหน้านี้ตามที่แสดงในภาพ



ภาพที่ 3.9 ผังงานลำดับขั้นตอนกฎผู้เชี่ยวชาญ

3.2.4 MCRB Forecasting Models MCRB นี้ผู้วิจัยได้ทำการพัฒนารูปแบบของโมเดล เพื่อเป็นการเพิ่มประสิทธิภาพวิธี MC ให้มีความยืดหยุ่นและเหมาะสมกับสถานการณ์ที่เคยเกิดขึ้น โดยมีแนวคิดที่สำคัญ 2 ประเด็นคือ

ประเด็นแรก หากข้อมูลที่เกิดขึ้นบ่อย ย่อมมีความน่าจะเป็นและแนวโน้มที่จะเกิดขึ้นอีกในอนาคต

ประเด็นที่สอง ถ้าข้อมูลเกิดขึ้นแล้วไม่เกิดขึ้นอีกในระยะเวลาหนึ่ง ข้อมูลนั้นก็จะเสื่อมไป (Decay)

โดยขั้นตอนในการคัดเลือกข้อมูลโดย Decay procedure แสดงในภาพที่ 3.10

ALGORITHM DecayProcedure ($D[1...n]$)

การคัดเลือกข้อมูลเพื่อใช้ในการจำลองสถานการณ์เทคนิคมอนติคาร์โล

Input: แถวลำดับข้อมูลความต้องการในอดีต $D[1...n]$

Output: ข้อมูลที่ถูกคัดเลือกเพื่อใช้ในการจำลองสถานการณ์เทคนิคมอนติคาร์โล $DF[1..*]$

```

1       $D[1...n]$ 
2       $k \leftarrow 0$ 
3      For  $j \leftarrow n-30$  to  $n$ 
4          For  $i = 1$  to  $n$ 
5              If  $D[i] = D[j]$ 
6                  Then {  $Sum[k] \leftarrow Sum[k] + 1$ 
7                       $DF[k] \leftarrow D[j]$ 
8                       $k \leftarrow k + 1$  }
9              End for
10     End for
  
```

ภาพที่ 3.10 ขั้นตอนการคัดเลือกข้อมูลที่มีความเสื่อมเพื่อใช้ในการจำลองสถานการณ์เทคนิคมอนติคาร์โล

เรากำหนดตัวแปรและพารามิเตอร์ต่าง ๆ ดังนี้

D คือ เซตข้อมูลปริมาณความต้องการสินค้าที่ใช้ในการคัดเลือก
โดย $D = \{1 \dots n\}$

DF คือ เซตข้อมูลปริมาณความต้องการสินค้า ที่ถูกคัดเลือก โดย $DF = \{1 \dots k\}$

Sum คือ เซตจำนวนครั้งของความต้องการสินค้า ของเซตที่ถูกคัดเลือก (DF)
ไปใช้ในการจำลองสถานการณ์โดย $Sum = \{1 \dots k\}$

จากภาพที่ 3.10 แสดงขั้นตอนการคัดเลือกข้อมูลเพื่อใช้ในการจำลองสถานการณ์เทคนิคมอนติคาร์โล (MC) โดยขั้นตอนการคัดเลือกเริ่มจากการนำข้อมูล 30 ระเบียบสุดท้ายมาเปรียบเทียบกับข้อมูลทั้งหมดที่มีอยู่ที่ละค่าว่าเคยเกิดขึ้นมาก่อนหรือไม่ หากพบว่าค่าข้อมูลดังกล่าวเคยเกิดขึ้นมาก่อนก็จะเก็บค่าความถี่ของข้อมูลไว้เพื่อนำไปคิดเป็นค่าความน่าจะเป็นที่ข้อมูลดังกล่าวจะเกิดขึ้นอีกในการจำลองสถานการณ์ แต่ถ้าหากข้อมูลใดที่ไม่พบความถี่ในการเกิดขึ้นก็จะถูกตัดออกไม่ถูกนำมาจำลองสถานการณ์เพื่อการพยากรณ์

ส่วนการกำหนดจำนวนครั้งในการสุ่มและการจำลองปัญหามีรูปแบบดังต่อไปนี้

การกำหนดจำนวนครั้งในการสุ่ม สามารถหาได้จากสมการ (3.1)

$$iteration = \frac{z^2 s^2}{d^2} \quad (3.1)$$

โดยที่ $iteration$ คือ จำนวนครั้งในการจำลอง (สุ่ม) ข้อมูล

s คือ ค่าความเบี่ยงเบนมาตรฐาน

z คือ ค่าความเชื่อมั่น (เช่น ค่าความเชื่อมั่นที่ระดับ 98% ค่า

$$z = 2.05)$$

d คือ ค่าความคลาดเคลื่อนที่มากที่สุดที่ยอมรับได้ (เช่น 2% ของค่าเฉลี่ย)

ทำการจำลองปัญหา (Simulation) ในงานวิจัยนี้ เราพัฒนาวิธี MC ให้มีความยืดหยุ่นและเปลี่ยนแปลงตามข้อมูลความต้องการอย่างอัตโนมัติ (Dynamic MC) โดยมีขั้นตอนดังนี้

1) ปรับการกระจายค่าความน่าจะเป็นของข้อมูล(Frequency update) ในทุก ๆ ครั้งที่มีการเปลี่ยนแปลงชุดข้อมูล

2) หาจำนวนครั้งในการสุ่ม (*iteration*) ในทุก ๆ ครั้งที่มีการเปลี่ยนแปลงข้อมูล
(สมการ 3.1)

เราสามารถเขียนขั้นตอนการทำงานของวิธี Dynamic MC ดังแสดงในภาพที่ 3.11

มหาวิทยาลัยบูรพา
Burapha University

ALGORITHM DynamicMC ($DF[1...n], Sum[1...n]$)

จำลองสถานการณ์พยากรณ์ข้อมูลแบบอัตโนมัติด้วยเทคนิคมอนติคาโล

Input: แถวลำดับข้อมูลความต้องการ $DF[1...n]$ และข้อมูลความถี่ $Sum[1...n]$

ใช้ในการจำลอง สถานการณ์สำหรับพยากรณ์ข้อมูล

Output: ข้อมูลผลการพยากรณ์ *Forecastvalue*

```

1       $DF[1...n]$ 
2       $Sum[1...n]$ 
3       $SORT(DF[1...n], Sum[1...n])$ 
4      For  $i = 1$  to  $n$ 
5          For  $j \leftarrow 1$  to  $n$ 
6               $Sum \leftarrow Sum + Sum[j]$ 
7          End For
8          For  $k \leftarrow 1$  to  $n$ 
9               $Prop[k] \leftarrow Sum[k] / Sum$ 
10             If  $(k > 1)$  then
11                  $CPI[k] \leftarrow CPI[k-1] + Prop[k]$ 
12                  $CPu[k-1] \leftarrow CPI[k]$ 
13             Else
14                  $CPI[k] \leftarrow 0$ 
15             End For
16         End For
17         For  $i = 1$  to iteration
18              $rand \leftarrow Random()$ 
19             For  $j = 1$  to  $n$ 
20                 If  $(rand \geq CPI[j] \text{ and } rand < CPu[j])$ 
21                     Then  $FC[i] \leftarrow DF[j]$ 
22                 End For
23              $total = total + FC[i]$ 
24         End For
25          $Forecastvalue = total / iteration$ 

```

ภาพที่ 3.11 ขั้นตอนของวิธี Dynamic MC

เรากำหนดตัวแปรและพารามิเตอร์ต่าง ๆ ดังนี้

<i>DF</i>	คือ เซตข้อมูลปริมาณความต้องการสินค้าที่ถูกคัดเลือก โดย $DF = \{1..n\}$
<i>Sum</i>	คือ เซตจำนวนความถี่ของความต้องการสินค้าของเซตที่ถูกคัดเลือก (<i>DF</i>) เพื่อไปใช้ในการหาความน่าจะเป็น โดย $Sum = \{1..n\}$
<i>Prop</i>	คือ เซตข้อมูลความน่าจะเป็นของเซตข้อมูลความถี่ (<i>Sum</i>) เพื่อไปใช้ในการจำลองสถานการณ์โดย $Prop = \{1..n\}$
<i>CPI, CPu</i>	คือ เซตขอบจำกัดล่าง, และขอบจำกัดบนของความน่าจะเป็นสะสมเซต <i>Prop</i>
<i>FC</i>	คือ เซตค่าพยากรณ์จำนวน <i>iteration</i>
<i>total</i>	คือ ผลรวมค่าพยากรณ์จำนวน <i>iteration</i>
<i>forecastvalue</i>	คือ ค่าพยากรณ์ที่ได้
<i>rand</i>	คือ ค่าเลขสุ่มที่ได้ในแต่ละ <i>iteration</i>

จากภาพที่ 3.11 แสดงขั้นตอนวิธีการจำลองสถานการณ์แบบอัตโนมัติ (Dynamic) โดยอัลกอริทึมเริ่มจากการนำข้อมูลทั้งหมดที่ได้จากการคัดเลือกตามขั้นตอนในภาพที่ 3.10 มากำหนดอันตรภาคชั้นและสร้างค่าความน่าจะเป็นสะสมในแต่ละชั้น ขั้นตอนต่อมาคือการจำลองสถานการณ์ตามจำนวนรอบ (Iteration) ที่หาได้ โดยทุก ๆ รอบจะทำการสุ่ม (Rand) ค่าความน่าจะเป็น และทำการตรวจสอบว่าค่าความน่าจะเป็นดังกล่าวอยู่ในช่วงอันตรภาคชั้นใด ซึ่งหากตกอยู่ในอันตรภาคชั้นใด จะถือว่าค่าพยากรณ์ (FC) ในครั้งนั้นเท่ากับข้อมูลที่แทนในชั้นนั้น และทำการเก็บค่าพยากรณ์สะสมไว้เพื่อหาผลพยากรณ์ (Forecastvalue) จากการจำลองสถานการณ์ โดยนำค่าพยากรณ์ที่ได้จากการจำลองสถานการณ์หารกับจำนวนรอบที่ใช้จำลอง

3.3 การสร้างต้นแบบ

การสร้างต้นแบบตามโครงสร้างกฎที่ออกแบบไว้เป็นขั้นตอนในการนำส่วนประกอบต่าง ๆ ดังกล่าวข้างต้นมาสร้างตัวแบบระบบพยากรณ์ด้วยการเขียนโปรแกรมภาษา Visual Basic for Application โดยเขียนกลไกการอนุมานเกี่ยวกับกฎการพยากรณ์ ส่งไว้ที่ตัวโปรแกรมแล้วเรียกใช้งานผ่านโปรแกรมตารางงาน ซึ่งสามารถที่จะใส่ชุดข้อมูลความต้องการที่เคยเกิดขึ้นให้ป้อนฐานในการพยากรณ์ และมีส่วนของการกำหนดค่าพารามิเตอร์ต่าง ๆ ผ่านโปรแกรมตารางงานให้ส่งค่าไปทำการวิเคราะห์ข้อมูล และทำการพยากรณ์ผลออกมาในรูปของปริมาณความต้องการที่น่าจะเกิดขึ้นในอนาคต

3.4 การทดสอบ การตรวจสอบ

หลังจากที่ต้นแบบได้ถูกสร้างขึ้นแล้วต้องทำการทดสอบการทำงานว่าสามารถตรวจสอบลักษณะของข้อมูลและทำงานตามกฎต่าง ๆ ที่สร้างไว้หรือไม่ เพื่อที่จะสามารถนำไปใช้ในสภาวะการณ์จริงโดยการทดสอบผลการพยากรณ์ด้วยตัววัดในรูปแบบต่าง ๆ โดยมีรูปแบบการเปรียบเทียบ ดังนี้

- เปรียบเทียบความแม่นยำของแต่ละตัววัด ซึ่งประกอบไปด้วยค่า MAPE, MdAPE และ sMAPE
- การเปรียบเทียบผลการพยากรณ์ของวิธีการที่พัฒนาขึ้นกับวิธีการพยากรณ์อื่น ๆ ด้วยวิธีการทางสถิติ

เมื่อระบบที่พัฒนาขึ้นผ่านการทดสอบแก้ไขจนสามารถทำงานได้ตามที่ออกแบบไว้ และได้ผลการพยากรณ์ที่ถูกต้องแม่นยำในระดับที่น่าพอใจแล้วจะถือได้ว่าการพัฒนาระบบผู้เชี่ยวชาญด้านการพยากรณ์ประสบความสำเร็จในระดับหนึ่ง โดยหากมีองค์ความรู้ใหม่ ๆ ที่น่าสนใจ ก็สามารถที่จะนำมาทำการปรับปรุงเพิ่มเติมประสิทธิภาพการทำงานของระบบได้อีกตามความเหมาะสม