



ใบรับรองวิทยานิพนธ์
บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตร์มหาบัณฑิต (การจัดการเทคโนโลยีอุตสาหกรรมเกษตร)

ปริญญา

การจัดการเทคโนโลยีอุตสาหกรรมเกษตร

เทคโนโลยีอุตสาหกรรมเกษตร

สาขา

ภาควิชา

เรื่อง การทำนายคุณลักษณะของวัตถุดิบเพื่อการเตรียมสารบรรจุในสายการผลิตสับปะรดบรรจุ
กระป๋อง

Raw Material Characteristic Prediction for Packing Media Preparation in Canned
Pineapple Production Line

นามผู้วิจัย นางสาวนฤมล ชมบุญมี

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์ปรารณา ปรารณาดิ, Ph.D.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์จรัชย์ พุทธกุลสมศิริ, Ph.D.)

อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

(ผู้ช่วยศาสตราจารย์วิพิมพ์ จวีสุข, Ph.D.)

หัวหน้าภาควิชา

(ผู้ช่วยศาสตราจารย์วิพิมพ์ จวีสุข, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญญา วีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่..... เดือน..... พ.ศ.

สิงสัทธี มทาวิตยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

การทำนายคุณลักษณะของวัตถุดิบเพื่อการเตรียมสารบรรจุในสายการผลิตสับประคบรรจุกระป๋อง

Raw Material Characteristic Prediction for Packing Media Preparation in Canned Pineapple
Production Line

โดย

นางสาวนฤมล ชมบุญมี

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (การจัดการเทคโนโลยีอุตสาหกรรมเกษตร)

พ.ศ. 2553

นฤมล ชมบุญมี 2553: การทำนายคุณลักษณะของวัตถุดิบเพื่อการเตรียมสารบรรจุในสายการผลิต
สับปะรดบรรจุกระป๋อง ปริญญาวิทยาศาสตรมหาบัณฑิต (การจัดการเทคโนโลยีอุตสาหกรรมเกษตร)
สาขาการจัดการเทคโนโลยีอุตสาหกรรมเกษตร ภาควิชาเทคโนโลยีอุตสาหกรรมเกษตร อาจารย์ที่ปรึกษา
วิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์ปรารณา ปรารณาคี, Ph.D. 134 หน้า

การเตรียมสารบรรจุ (Packing Medium) จัดเป็นกระบวนการสำคัญกระบวนการหนึ่งในสายการผลิต
สับปะรดบรรจุกระป๋อง เนื่องจากความหวานสุดท้าย (Cut-out Strengths) ของผลิตภัณฑ์จะขึ้นอยู่กับความเข้มข้นของ
สารบรรจุที่บรรจุลงในกระป๋อง ปริมาณของแข็งที่ละลายได้ (Total Soluble Solids) ของวัตถุดิบ และสัดส่วนของ
ปริมาณวัตถุดิบต่อปริมาณสารบรรจุในกระป๋อง ปริมาณของแข็งที่ละลายได้ ซึ่งวัดเป็นองศาบริกซ์ เป็นค่าประมาณ
ของปริมาณน้ำตาลในน้ำผลไม้หรือค่าความหวานของน้ำผลไม้ ซึ่งจะขึ้นอยู่กับความสุกและความน้ำของผลไม้ สาร
บรรจุที่มีค่าความหวานที่เหมาะสมกับค่าความหวานของวัตถุดิบจะต้องถูกเตรียมไว้ล่วงหน้า เพื่อลดความล่าช้าที่สูญ
เสียไปในกระบวนการผลิต การทำนายค่าความหวานของวัตถุดิบที่ถูกต้องจะช่วยสนับสนุนการเตรียมสารบรรจุได้
อย่างมาก งานวิจัยนี้นำเสนอการสร้างแบบจำลองสำหรับการทำนายค่าความหวานของวัตถุดิบสับปะรดและระบบ
สนับสนุนการตัดสินใจอย่างง่ายสำหรับการเตรียมสารบรรจุในสายการผลิตสับปะรดบรรจุกระป๋องของโรงงานผลิต
ผลไม้บรรจุกระป๋องของบริษัท มาลีสามพราน จำกัด (มหาชน) แบบจำลองที่พิจารณา ประกอบด้วย แบบจำลองการ
วิเคราะห์การถดถอยแบบพหุ แบบจำลองเครือข่ายประสาทเทียม และแบบจำลองต้นไม้การตัดสินใจ 2 แบบ ได้แก่
Classification And Regression Trees (CART) และ Chi-square Automatic Iteration Detector (CHAID) ผลการศึกษา
พบว่า แบบจำลองเครือข่ายประสาทเทียมที่กำหนดตัวแปรอิสระ คือ เดือนที่รับวัตถุดิบ (แสดงถึงฤดูกาลปลูก) เกรดสี
วัตถุดิบ (แสดงถึงระดับความสุก) อำเภอที่ปลูกและผู้รวบรวมวัตถุดิบ (แสดงถึงแหล่งที่มาของวัตถุดิบ) ให้ค่าความ-
คลาดเคลื่อนในการทำนายน้อยที่สุด อย่างไรก็ตาม แบบจำลองอื่นที่สร้างขึ้นก็ให้ค่าความคลาดเคลื่อนที่ไม่แตกต่าง
กันมากนัก ดังนั้น ผู้วิจัยจึงกำหนดให้ใช้ค่าทำนายความหวานของวัตถุดิบจากทุกแบบจำลองแสดงในระบบสนับสนุน
การตัดสินใจที่พัฒนาขึ้น และเพื่อเพิ่มความสามารถของโปรแกรมในการทำนายค่าให้ครอบคลุมชุดข้อมูลอื่น นอก
เหนือจากชุดข้อมูลที่ใช้ในการสร้างแบบจำลอง ผู้วิจัยได้แบ่งระดับความน่าเชื่อถือของการทำนายออกเป็น 5 ระดับ
ได้แก่ ระดับสูงมากและระดับสูง เป็นการทำนายค่าจากชุดข้อมูลที่มีตัวแปรอิสระตรงกับชุดข้อมูลในอดีตที่นำมาใช้ในการ
การสร้างแบบจำลอง จึงทำให้การทำนายค่ามีความแม่นยำสูง ระดับปานกลาง เป็นการทำนายค่าจากชุดข้อมูลที่มีการ
จับคู่ระหว่างอำเภอที่ปลูกและผู้รวบรวมวัตถุดิบไม่ตรงกับชุดข้อมูลที่นำมาใช้ในการสร้างแบบจำลอง แต่ยังคงอยู่ในขอบ
เขตที่ใช้ในการสร้างแบบจำลอง ระดับต่ำและระดับต่ำมาก เป็นการทำนายค่าจากชุดข้อมูลที่มีการรับวัตถุดิบเข้ามา
จากแหล่งใหม่ ทำให้ความแม่นยำในการทำนายค่าลดน้อยลง นอกจากนี้ ระบบสนับสนุนการตัดสินใจที่พัฒนาขึ้นยัง
แสดงค่าทำนายที่ได้จากแบบจำลองที่ดีที่สุด และค่าสถิติที่สำคัญ รวมทั้งการคำนวณค่าความหวานของสารบรรจุที่ต้อง
เตรียม เพื่อเป็นข้อมูลประกอบการตัดสินใจของผู้ใช้ต่อไป

Narumon Chomboonmee 2010: Raw Material Characteristic Prediction for Packing Media Preparation in Canned Pineapple Production Line. Master of Science (Agro-Industrial Technology Management), Major Field: Agro-Industrial Technology Management, Department of Agro-Industrial Technology. Thesis Advisor: Assistant Professor Parthana Parthanadee, Ph.D. 134 pages.

Packing medium (PM) preparation is an important process in canned pineapple production lines, as the cut-out strengths of the canned pineapple products depend on the strength of the PM filled in the cans, the total soluble solids (TSS) of the raw material (RM) juice, and the ratio of RM to PM. The TSS, measured as degree Brix, is an estimate of sugar content in the fruit juice, which greatly depends on the ripeness and succulence of the fruits. The PM, which matches with the degree Brix of the RM, needs to be prepared in advance in order to reduce the delay in the production process. The accurate RM degree Brix prediction would support this advance PM preparation greatly. This research presented the developments of RM degree Brix prediction models and a simple decision support system (DSS) for PM degree Brix preparation for canned pineapple production lines of the Malee Sampran PCL plant. The models considered consisted of multiple regression models, artificial neural network (ANN) models and decision trees models, including the classification and regression tree (CART) and Chi-square Automatic Iteration Detector (CHAID). The results showed that the ANN model with Month (indicating the season of cultivation), Color grade of the pineapples (indicating the pineapple ripeness), Amphoe and Supplier of raw materials (indicating the pineapple source) input as independent variables was found to be the best prediction model. However, there were no significant differences among the prediction errors obtained from all considered models. The DSS was therefore developed to present all the predicted RM degree Brix values from all models. To enhance the program ability in predicting the values from any other data inputs, the prediction reliability was classified into five levels. The very high and high reliability levels were resulted from the inputs of the independent variables that matched with the data sets used to construct the prediction models, hence yielding high prediction accuracy. The medium reliability level was resulted from the data inputs in which the Amphoe and Supplier variables were alternated from those presented in the actual data. The low and very low reliability levels were presented when raw materials were supplied from new sources, hence giving high prediction error values. This developed DSS program also showed the best prediction value and other important statistics, along with the required PM degree Brix calculation, to support the users' decisions.

Student's signature

Thesis Advisor's signature

กิตติกรรมประกาศ

ขอกราบขอบพระคุณ ผศ.ดร. ปรรณนา ปรรณนาดี อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ผู้ให้คำปรึกษาและความรู้ต่าง ๆ ที่ใช้ในงานวิจัยนี้ ตลอดจนกรุณาให้คำแนะนำ ตรวจสอบแก้ไขข้อบกพร่องต่าง ๆ จนสำเร็จเป็นวิทยานิพนธ์ได้ ขอกราบขอบพระคุณ ผศ.ดร. จิรัชัย พุทธกุลสมศิริ อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมที่กรุณาให้คำแนะนำและความรู้ต่าง ๆ และแก้ไขข้อบกพร่องเกี่ยวกับแบบจำลองทางสถิติ ขอกราบขอบพระคุณ ผศ.ดร. รวิพิมพ์ ฦวิสุข อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วมที่กรุณาให้คำแนะนำและความรู้ต่าง ๆ ที่สำคัญเกี่ยวกับแบบจำลองเครือข่ายประสาทเทียม จนงานวิจัยสำเร็จไปได้ด้วยดี และขอกราบขอบพระคุณคณาจารย์ทุกท่านที่ให้ความรู้ทางด้านต่าง ๆ ซึ่งนำมาประยุกต์ใช้ในงานวิจัยนี้

ขอกราบขอบพระคุณ คุณพ่อ คุณแม่ ที่ให้การสนับสนุนทุกด้าน ให้โอกาสทางการศึกษา ให้กำลังใจและอบรมสั่งสอนมาโดยตลอด สุดท้ายขอขอบพระคุณเพื่อน ๆ ทุกท่านที่คอยช่วยเหลือ ให้คำแนะนำและเป็นกำลังใจตลอดมา

ขอกราบขอบพระคุณทางบริษัท มาลีสามพราน จำกัด (มหาชน) ที่ให้ความอนุเคราะห์เรื่องข้อมูลที่นำมาใช้ในงานวิจัยนี้ ตลอดจนอำนวยความสะดวกในด้านการเชื่อมโยงข้อมูลที่เกี่ยวข้องต่าง ๆ และโครงการทุนวิจัยมหาบัณฑิต สกว. สาขาวิทยาศาสตร์และเทคโนโลยี ภายใต้โครงการเชื่อมโยงภาคการผลิตกับงานวิจัย ทุน สกว. – อุดสาหกรรม ปี 2551 (MRG-WI515S025) ที่ให้ทุนสนับสนุนงานวิจัยครั้งนี้

นฤมล ชมบุญมี

มกราคม 2553

สารบัญ

	หน้า
สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	3
การตรวจเอกสาร	4
อุปกรณ์และวิธีการ	47
อุปกรณ์	47
วิธีการ	47
ผลและวิจารณ์	59
สรุปและข้อเสนอแนะ	89
สรุป	89
ข้อเสนอแนะ	90
เอกสารและสิ่งอ้างอิง	91
ภาคผนวก	96
ภาคผนวก ก รายละเอียดของตัวแปรอิสระที่เป็นข้อมูลที่ใช้สร้างแบบจำลองต่างๆ	97
ภาคผนวก ข การแบ่งกลุ่มข้อมูลในการทำนายค่าโดยวิธีต้นไม้การตัดสินใจในแบบจำลองต่างๆ	107
ภาคผนวก ค การตรวจสอบค่าความคลาดเคลื่อนของแบบจำลองต่างๆ	124
ประวัติการศึกษา และการทำงาน	134

สารบัญตาราง

ตารางที่		หน้า
1	รายละเอียดของตัวแปรอิสระและตัวแปรตามที่ใช้ในการสร้างแบบจำลอง	52
2	ตัวแปรอิสระที่พิจารณาในแบบจำลองการทำนายค่าความหวานวัตถุดิบ	53
3	การกำหนดตัวแปรหุ่นที่ใช้ในการสร้างแบบจำลองการถดถอย	59
4	ตัวแปรอิสระที่พิจารณาในแบบจำลองการทำนายค่าความหวาน โดยการวิเคราะห์การถดถอย	61
5	ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่าง ๆ โดยวิธีวิเคราะห์การถดถอยเชิงพหุ	67
6	โครงสร้างที่เหมาะสมของแบบจำลองต่างๆ โดยวิธีเครือข่ายประสาทเทียม	68
7	ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธีเครือข่ายประสาทเทียม	69
8	ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธี CART	69
9	ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธี CHAID	70
10	รายละเอียดของตารางค่าความหวานจริงในอดีต	79
11	รายละเอียดของตารางค่าทำนายความหวานวัตถุดิบ	80
12	รายละเอียดตารางระดับความน่าเชื่อถือในการทำนายค่า	82
 ตารางผนวกที่		
ก1	รายละเอียดอำเภอที่ปลูกซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการทำนายค่าความหวานวัตถุดิบสับปะรด	98
ก2	รายละเอียดของผู้รวบรวมวัตถุดิบซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการทำนายค่าความหวานวัตถุดิบสับปะรด	99
ก3	รายละเอียดของแหล่งที่มาของวัตถุดิบซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการทำนายค่าความหวานวัตถุดิบสับปะรด	100

สารบัญตาราง (ต่อ)

ตารางผนวกที่		หน้า
ก4	ตัวแปรหุ่นเดือนที่รับวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม	102
ก5	ตัวแปรหุ่นอำเภอที่ปลูกวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม	103
ก6	ตัวแปรหุ่นผู้รวบรวมวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม	104
ก7	ตัวแปรหุ่นแหล่งที่มาวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม	105
ข1	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง MC โดยวิธี CART	108
ข2	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง A โดยวิธี CART	109
ข3	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง Q โดยวิธี CART	110
ข4	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง AQ โดยวิธี CART	112
ข5	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง S โดยวิธี CART	114
ข6	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง MC โดยวิธี CHAID	116
ข7	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง A โดยวิธี CHAID	117
ข8	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง Q โดยวิธี CHAID	118
ข9	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง AQ โดยวิธี CHAID	120
ข10	การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง S โดยวิธี CHAID	122

สารบัญภาพ

ภาพที่		หน้า
1	ขั้นตอนการผลิตอุตสาหกรรมสับประรดกระป๋อง	9
2	การทำงานของระบบประสาทของมนุษย์	17
3	การทำงานของเครือข่ายประสาทเทียม	18
4	เครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าชั้นเดียว	19
5	เครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าหลายชั้น	20
6	ฟังก์ชันปั้งซี (Identity Function)	21
7	ฟังก์ชันซิกมอยด์ (Sigmoid Function)	22
8	ฟังก์ชันไฮเพอร์โบลิคแทนเจน (Hyperbolic Tangent; tanH)	22
9	โครงสร้างของเครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ	23
10	การเปลี่ยนแปลงอัตราการเรียนรู้ระหว่างการฝึกเครือข่าย	25
11	โครงสร้างต้นไม้การตัดสินใจ แบบ CART	29
12	ค่าเฉลี่ยระดับความหวานวัตถุดิบสับประรดในแต่ละเดือนที่รับวัตถุดิบ	48
13	ค่าเฉลี่ยระดับความหวานวัตถุดิบสับประรดในแต่ละเกรดสีวัตถุดิบ	49
14	ค่าเฉลี่ยระดับความหวานวัตถุดิบสับประรดในแต่ละอำเภอที่ปลูก	50
15	ค่าเฉลี่ยระดับความหวานวัตถุดิบสับประรดในแต่ละผู้รวบรวมวัตถุดิบ	50
16	ค่าเฉลี่ยระดับความหวานวัตถุดิบสับประรดในแต่ละเดือนแหล่งที่มาวัตถุดิบ	51
17	แผนภาพการกระจายของค่าคลาดเคลื่อนในการวิเคราะห์การถดถอยเบื้องต้น เมื่อไม่มีเทอมยกกำลังสองของตัวแปรเกรดสี	60
18	แผนภาพการกระจายของค่าคลาดเคลื่อนในการวิเคราะห์การถดถอยเบื้องต้น เมื่อมีเทอมยกกำลังสองของตัวแปรเกรดสี	60
19	ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับสูงมาก	73
20	ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับสูง	74
21	ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับปานกลาง	76
22	ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับต่ำ	77
23	ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับต่ำมาก	78
24	หน้าหลักการใช้งานฐานข้อมูลการทำนายค่าความหวาน	83

สารบัญภาพ (ต่อ)

ภาพที่		หน้า
25	ช่องจังหวัดที่ปลูกแสดงโดยอัตโนมัติ เมื่อผู้ใช้เลือกอำเภอที่ปลูก	84
26	ตัวอย่างการแสดงค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับสูงมาก	84
27	ตัวอย่างการแสดงค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับสูง	85
28	ตัวอย่างการแสดงค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับปานกลาง	85
29	ตัวอย่างการแสดงค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับต่ำ	86
30	ตัวอย่างการแสดงค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับต่ำมาก	86
31	ตัวอย่างการแสดงค่าทำนายความหวานจริงในอดีตที่ใช้ในการสร้างแบบจำลอง	87
32	ส่วนคำนวณค่าความหวานของสารบรรจุที่ควรเตรียมในกระบวนการผลิต	87
ภาพผนวกที่		
ค1	แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูงและสูงมาก	125
ค2	แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูงและสูงมาก	126
ค3	แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูงและสูงมาก	127
ค4	แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือปานกลางและต่ำ	128

สารบัญภาพ (ต่อ)

ภาพผนวกที่	หน้า
ค5 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือปานกลางและต่ำ	129
ค6 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือปานกลางและต่ำ	130
ค7 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำมาก	131
ค8 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำมาก	132
ค9 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำมาก	133

การทำนายคุณลักษณะของวัตถุดิบเพื่อการเตรียมสารบรรจุในสายการผลิตสับประรด บรรจุกระป๋อง

Raw Material Characteristic Prediction for Packing Media Preparation in Canned Pineapple Production Line

คำนำ

สับประรดเป็นผลผลิตทางการเกษตรที่มีมูลค่าการส่งออกเป็นอันดับต้นๆของประเทศไทย โดยรูปแบบหลักที่ส่งออก คือ การแปรรูปเป็นสับประรดกระป๋อง มีมูลค่าส่งออกเพิ่มสูงขึ้นอย่างต่อเนื่อง โดยในปี 2551 มีมูลค่าส่งออกประมาณ 18,700 ล้านบาท คิดเป็นร้อยละ 30 ของผักและผลไม้กระป๋องส่งออกทั้งหมด (สำนักส่งเสริมการค้าสินค้าเกษตร กรมการค้าภายใน, 2552)

กระบวนการผลิตสับประรดกระป๋องจะมีการเติมสารบรรจุ (Packing Media) ซึ่งอาจเป็นน้ำผลไม้หรือน้ำเชื่อมอย่างใดอย่างหนึ่งหรือรวมกัน โดยกำหนดคุณลักษณะการเตรียมสารบรรจุเป็นค่าระดับความหวานหรือปริมาณของแข็งละลายได้เป็นองศาบริกซ์ เนื่องจากคุณลักษณะของวัตถุดิบสับประรดมีค่าไม่แน่นอนแตกต่างกันตามแหล่งการปลูก ระดับความสุก สภาพการปลูก ระยะเก็บเกี่ยว ฤดูกาล และปัจจัยที่เกี่ยวข้องอื่นๆ ทางผู้ผลิตจึงจำเป็นต้องเตรียมสารบรรจุที่จะเติมลงไปในสับประรดกระป๋องให้มีระดับความหวานที่เหมาะสมกับวัตถุดิบแต่ละเกรด เพื่อให้ได้ระดับความหวานของผลิตภัณฑ์ที่ตรงตามมาตรฐานที่กำหนด ซึ่งในปัจจุบันผู้ผลิตสับประรดกระป๋องจะเตรียมสารบรรจุล่วงหน้าก่อนการผลิต โดยอาศัยการประมาณค่าจากข้อมูลคุณลักษณะของวัตถุดิบในอดีต และข้อมูลจากการผลิตวันก่อนหน้า ร่วมกับประสบการณ์และความชำนาญส่วนบุคคลของพนักงาน ซึ่งมักจะเกิดความผิดพลาดจากการคาดการณ์ ส่งผลให้ต้องมีการปรับปรุงคุณลักษณะของสารบรรจุในระหว่างการผลิต ก่อให้เกิดความล่าช้าในกระบวนการผลิต

ในงานวิจัยนี้ได้ศึกษาถึงความสัมพันธ์ระหว่างปัจจัยที่เกี่ยวข้องกับคุณลักษณะของวัตถุดิบสับประรดของโรงงานกรณีศึกษา โดยการนำเทคนิคทางคณิตศาสตร์และสถิติต่างๆ เข้ามาช่วยในการสร้างแบบจำลองการทำนายคุณลักษณะและพัฒนาระบบสนับสนุนการตัดสินใจอย่างง่าย เพื่อนำไปใช้ในการกำหนดคุณลักษณะของสารบรรจุที่ต้องเตรียมไว้ล่วงหน้าต่อไป ซึ่งแบบจำลองการทำนาย

และระบบสนับสนุนการตัดสินใจอย่างง่ายนี้จะช่วยให้กระบวนการผลิตสับปะรดบรรจุกระป๋อง เป็นไปอย่างต่อเนื่องมากขึ้น สามารถลดการรอคอยในกระบวนการผลิต ทำให้ผลผลิตของโรงงาน แปรรูปสับปะรดกระป๋องสูงขึ้น ลดต้นทุนการผลิตต่อหน่วย อีกทั้งยังเป็นการรักษาคุณภาพของ สินค้าที่อาจเสียไประหว่างการรอคอยการผลิตอีกด้วย



วัตถุประสงค์

1. เพื่อศึกษาปัจจัยที่สำคัญที่มีผลต่อคุณลักษณะในด้านระดับความหวานของวัตถุดิบ
สับปะรด
2. เพื่อเปรียบเทียบประสิทธิภาพการทำนายค่าคุณลักษณะของวัตถุดิบสับปะรดด้วยการ
สร้างแบบจำลองทางคณิตศาสตร์และสถิติต่างๆ
3. เพื่อหาแบบจำลองการทำนายที่เหมาะสม ในการนำไปประยุกต์ใช้ในการสร้างระบบ
สนับสนุนการตัดสินใจแบบง่าย (Simple Decision Support System) เพื่อใช้ทำนายค่าความหวาน
ของวัตถุดิบสับปะรด และนำค่าทำนายความหวานที่ได้ไปใช้ในการกำหนดคุณสมบัติของสารบรรจุ
(Packing Media) ในกระบวนการผลิตสับปะรดกระป๋องต่อไป

การตรวจเอกสาร

อุตสาหกรรมแปรรูปสับปะรดของประเทศไทย

อุตสาหกรรมแปรรูปสับปะรดเป็นอุตสาหกรรมที่ประเทศไทยมีศักยภาพในการผลิตสูง เนื่องจากสภาพภูมิประเทศที่ตั้งอยู่ในบริเวณที่เป็นป่าเขตร้อน จึงเป็นส่วนสำคัญที่ทำให้สามารถปลูกสับปะรดที่มีรสชาติโดดเด่นและเป็นที่ต้องการของตลาดโลก สภาพการผลิตอุตสาหกรรมสับปะรดกระป๋องไทย จะทำการผลิตได้อย่างเต็มที่ในช่วงฤดูกาลที่มีวัตถุดิบออกมามากระหว่างเดือนเมษายน – พฤษภาคม และพฤศจิกายน – กุมภาพันธ์ ซึ่งในแต่ละปีโรงงานผลิตสับปะรดกระป๋องจะทำการผลิตได้ประมาณ 7 - 8 เดือนและนิยมนำวัตถุดิบอื่นๆ เช่น เงาะ ลิ้นจี่ ลำไย ฝรั่ง มะละกอ เป็นต้น มาใช้ในการผลิตเพื่อให้มีการผลิตตลอดทั้งปี

1. วัตถุดิบที่ใช้ในการผลิต

วัตถุดิบหลักที่ใช้ในการผลิต ได้แก่ สับปะรดสด เป็นพืชใบเลี้ยงเดี่ยวประเภทล้มลุก ปลูกได้ง่าย ไม่ต้องการการบำรุงรักษามากนัก และสามารถปลูกได้เกือบตลอดทั้งปี แต่ช่วงที่เหมาะสมสำหรับการเพาะปลูก คือ ระหว่างเดือนพฤศจิกายน - พฤษภาคม การปลูกแต่ละครั้งสามารถเก็บผลได้ 3 รุ่น พันธุ์ที่นิยมปลูกในประเทศไทยมี 3 พันธุ์ คือ พันธุ์อินทรีชิด พันธุ์ขาว และพันธุ์ปัตตาเวียหรือพันธุ์ศรีราชา (Smooth Cayenne) โดยประมาณร้อยละ 90 ของเนื้อที่เพาะปลูกสับปะรดในประเทศไทย จะปลูกพันธุ์ปัตตาเวีย เนื่องจากเป็นสับปะรดที่มีคุณภาพเหมาะสมในการบริโภคสด และเป็นวัตถุดิบป้อนโรงงานอุตสาหกรรมแปรรูป พื้นที่เพาะปลูกที่สำคัญ อยู่ในเขตจังหวัดประจวบคีรีขันธ์ เพชรบุรี และชุมพร นอกจากนี้ยังมีพื้นที่เพาะปลูกในแหล่งอื่นที่สำคัญ ได้แก่ จังหวัดระยอง ชลบุรี ตราด หนองคาย นครพนม ลำปาง และกาญจนบุรี เป็นต้น (สำนักงานเศรษฐกิจอุตสาหกรรม, 2545)

วิธีการที่ทางโรงงานแปรรูปจะได้สับปะรดมาเป็นวัตถุดิบ 3 วิธี (ทศพร, 2546) ดังนี้

1.1 การรับซื้อจากชาวไร่ทั่วไปหรือจากพ่อค้าคนกลาง (ระบบการซื้อขายแบบอิสระ)

โรงงานแปรรูปจะจัดซื้อสับปะรดสดจากเกษตรกรที่นำมาขายหน้าโรงงานหรือเข้าไปรับซื้อถึงไร่ โดยจะมีการกำหนดราคารับซื้อในแต่ละวันเมื่อมีการรับซื้อ วิธีนี้จะทำให้โรงงานแปรรูป

รูปไม่ต้องรับผิดชอบ หรือมีเงื่อนไขการซื้อมากนัก เพียงแต่มีการตรวจสอบคุณภาพและมาตรฐานให้ได้ตามความต้องการเท่านั้น โดยราคาสับปะรดจะถูกกำหนดโดยกลไกตลาด ทั้งทางด้านอุปสงค์และอุปทานในแต่ละฤดูกาลผลิต

1.2 การซื้อขายโดยผ่านการทำสัญญา (Contract Farming)

โรงงานแปรรูปสับปะรดจะทำสัญญากับเกษตรกรผู้ผลิตกลุ่มใหญ่ ผ่านสหกรณ์หรือกลุ่มเกษตรกร หรือเกษตรกรรายใหญ่ เพื่อความสะดวกในการทำสัญญาและลดต้นทุนในการทำธุรกรรม (Transaction Cost) โดยทำสัญญาซื้อขายล่วงหน้าในแต่ละปี มีการระบุปริมาณและราคา วิธีนี้จะช่วยให้โรงงานแปรรูปทราบถึงจำนวนพื้นที่เพาะปลูก สามารถคาดคะเนผลผลิตที่แน่นอนได้ ทำให้ลดความเสี่ยงด้านผลผลิต (Supply Risk) อันเป็นปัญหาสำคัญของอุตสาหกรรมแปรรูป การเกษตรที่ต้องการความสม่ำเสมอของปริมาณวัตถุดิบ เพื่อป้อนสู่โรงงานแปรรูป นอกจากนี้แล้วเงื่อนไขของสัญญาทางการช่วยเหลือและส่งเสริมความรู้ทางด้านวิชาการแก่เกษตรกรในระบบลูกไร่ ในสัญญาจะช่วยให้ทางโรงงานแปรรูปสามารถควบคุมคุณภาพของวัตถุดิบได้ในระดับหนึ่ง

การทำสัญญานี้ เกษตรกรจะได้รับผลประโยชน์ในแง่ของความแน่นอนทางการตลาด (Assured Market) เป็นการลดภาระความเสี่ยงด้านการตลาด โดยเฉพาะอย่างยิ่งในช่วงที่สับปะรดล้นตลาด เกษตรกรสามารถทราบล่วงหน้าถึงราคาที่จะได้รับ ในขณะที่ทางโรงงานแปรรูปสามารถทราบต้นทุนทางวัตถุดิบอย่างคร่าวๆ ล่วงหน้า โดยไม่ต้องเสี่ยงกับการแปรเปลี่ยนทางด้านราคาวัตถุดิบ แต่ปัญหาที่เกิดขึ้นจากการซื้อขายในระบบสัญญา มักเกิดจากการไม่ปฏิบัติตามข้อผูกมัดทางสัญญา เช่น กรณีที่สับปะรดล้นตลาด ทางโรงงานแปรรูปก็มักจะจุกจิกในการตีราคาและคัดเกรดสับปะรดที่แพงกว่าเมื่อเทียบกับเกรดที่ต้องรับซื้อตามสัญญา ในทำนองเดียวกันในฤดูกาลที่สับปะรดขาดแคลน เกษตรกรบางรายก็ไม่อยากรักษาสัญญาผูกพัน เป็นเพราะราคาตลาดอาจสูงกว่าราคาที่ตนได้ตกลงไว้ในสัญญา

1.3 การปลูกสับปะรดในไร่ของโรงงานเอง

ในกรณีนี้โรงงานแปรรูปจะทำการปลูกสับปะรดในพื้นที่ของตนเอง หรือพื้นที่เช่าเพื่อบรรเทาปัญหาในช่วงสับปะรดขาดแคลน และเพื่อควบคุมคุณภาพให้ตรงตามความต้องการ แต่เป็นการเพิ่มภาระในการควบคุมการจัดการของบริษัท ซึ่งจะมีความฉดในการแปรรูปมากกว่าการเพาะ

ปลูก รวมทั้งต้องรับภาระเรื่องการเช่าพื้นที่เพาะปลูก ค่าแรงงานให้เอาใจใส่ดูแลพืชผลให้ได้คุณภาพ โดยปกติแล้วทางบริษัทมักจะพยายามหลีกเลี่ยงวิธีนี้ และพยายามหาวัตถุดิบโดยการรับซื้อจากชาวไร่หรือการซื้อขายโดยผ่านการทำสัญญาแทน

2. ประเภทผลิตภัณฑ์

โรงงานผลิตสับปะรดส่วนใหญ่จะทำการผลิตผลิตภัณฑ์ตามการสั่งซื้อของลูกค้าแต่ละราย ที่ต้องการผลิตภัณฑ์สับปะรดกระป๋องแตกต่างกัน โดยทั่วไปสามารถนำสับปะรดสดมาผลิตเป็นผลิตภัณฑ์สับปะรดแบบต่าง ๆ (สำนักงานเศรษฐกิจอุตสาหกรรม, 2545) ได้ดังนี้

2.1 สับปะรดแว่น (Slices or Rings) เป็นผลิตภัณฑ์ที่มีราคาสูงและเป็นผลิตภัณฑ์หลักของโรงงานอุตสาหกรรมสับปะรดกระป๋อง สับปะรดผ่านเป็นแว่นชนิดที่ดีที่สุดคือ ชนิดแว่นแฟนซี (Fancy Slices) รองลงมาคือ แว่นชนิดคัด (Choice Slices) และชนิดแว่นมาตรฐาน (Standard Slices) นอกจากนี้มีชนิดครึ่งแว่น (Half Slices) ชนิดหนึ่งส่วนสี่แว่น (Quarter Slices) และชนิดแว่นหัก (Broken Slices) ซึ่งใช้สับปะรดที่มีคุณภาพไม่ถึงมาตรฐานของผลิตภัณฑ์ข้างต้นมาทำ ชิ้นที่หักอาจมีขนาดและสัดส่วนไม่เสมอกัน

2.2 สับปะรดชิ้นใหญ่ (Chunks) เป็นสับปะรดชิ้นสั้นหนาที่ตัดมาจากสับปะรดที่ผ่านเป็นแว่นกว้าง 12 มิลลิเมตรขึ้นไป และยาวไม่เกิน 38 มิลลิเมตร

2.3 สับปะรดชิ้นยาว (Spears or Fingers) ได้แก่ สับปะรดที่ตัดเป็นเสี้ยวตามแนวยาวของสับปะรดทั้งผลแต่ละชิ้นยาว 65 มิลลิเมตร หรือยาวกว่านั้น

2.4 สับปะรดลิ้ม (Tidbits) ได้แก่ สับปะรดที่ตัดเป็นเสี้ยวจากสับปะรดแว่นมีลักษณะคล้ายลิ้ม มีสัดส่วนสม่ำเสมอ หนาประมาณ 8 – 13 มิลลิเมตร

2.5 สับปะรดลูกเต๋า (Dices or Cubes) ได้แก่ สับปะรดที่มีลักษณะคล้ายลูกบาศก์ขอบด้านยาวที่สุดต้องไม่เกิน 14 มิลลิเมตร

2.6 สับประรดชิ้นละ (Pieces) เป็นสับประรดที่มีขนาดชิ้นไม่สม่ำเสมอ และไม่สามารถจัดรวมเข้าอยู่ในชนิดใดชนิดหนึ่งข้างต้นได้

2.7 สับประรดชิ้นเศษ (Chips) ได้จากเศษเนื้อสับประรดที่เหลือจากการทำสับประรดลูกเต๋า

2.8 สับประรดชิ้นย่อย (Crushed or Crisp Cut) เป็นสับประรดลูกเต๋าละเอียดๆ หรือฝานเป็นชิ้นยาวๆ ชูดหรือซอยเป็นชิ้นเล็กๆ นอกจากนี้ยังมีน้ำสับประรด (Juice) เป็นผลพลอยได้จากการผลิตสับประรดกระป๋อง

นอกจากนี้อาจมีการนำสับประรดไปผลิตเป็นผลิตภัณฑ์อื่น ๆ เช่น เงานะสอดใส่สับประรดและผลไม้รวมบรรจุกระป๋อง (Canned Tropical Fruit Salad) ซึ่งจะมีส่วนผสมของผลไม้ต่างๆ ตามสูตรที่ถูกกำหนดไว้ เช่น สับประรด มะละกอ ฝรั่ง กล้วย เชอรี่ เป็นต้น

3. กระบวนการผลิต

การผลิตสับประรดกระป๋องใช้วัตถุดิบที่สำคัญ คือ สับประรดสดและวัตถุดิบประกอบอื่นๆ เช่น น้ำตาลทราย กรดมะนาว กระป๋องที่ใช้บรรจุ และกล่องกระดาษ เป็นต้น โรงงานผลิตสับประรดกระป๋องเป็นอุตสาหกรรมที่ใช้แรงงานมาก โดยเฉพาะในส่วนการตกแต่งสับประรด และการควบคุมน้ำหนักของผลิตภัณฑ์ ขั้นตอนการผลิตสับประรดกระป๋องแสดงในภาพที่ 1 โดยมีขั้นตอนการผลิตดังนี้ (สำนักงานเศรษฐกิจอุตสาหกรรม, 2545)

3.1 เกษตรกรจะนำสับประรดที่หักก้านและยอดออกแล้วมาจำหน่ายให้โรงงาน โดยโรงงานจะทำการสุ่มตรวจคุณภาพเบื้องต้น คัดขนาด ซึ่งน้ำหนักสับประรดตามมาตรฐานที่สมาคมผู้ผลิตอาหารสำเร็จรูปกำหนดขึ้น จากนั้นจะส่งผ่านสับประรดไปตามสายพานขนาดต่างๆ ผ่านเครื่องจืดน้ำเพื่อทำความสะอาดสับประรดก่อนเข้าเครื่องตัดหัวท้าย และลำเลียงเข้าเครื่องปอกเปลือก และเอาแกนออก เปลือกสับประรดและแกนสับประรดจะถูกลำเลียงแยกออก เพื่อนำไปคั้นเป็นน้ำสับประรดต่อไป

3.2 สับประรดที่ได้จะผ่านการล้างทำความสะอาดอีกครั้งก่อนจะเคลื่อนไปตามเครื่องลำเลียง เพื่อให้คนงานตกแต่งดึงตาหรือเปลือกที่ยังติดค้างออก จากนั้นจะผ่านเข้าเครื่องฝานสับประรด

เป็นแวนส์สัปดาห์ 1 ลูกจะผ่านเป็นแวนส์ได้ประมาณ 8-10 แวนส์ ซึ่งในขั้นตอนนี้คนงานจะต้องคัดเลือกแวนส์ตามเกรดสีและขนาดความสมบูรณ์ของแวนส์ ขนาดที่ไม่สมบูรณ์จะถูกนำเข้าเครื่องขอยเป็นลิ้มลูกเต๋าหรืออื่นๆ เพื่อบรรจุกระป๋องหรือผสมกับผลไม้อื่น ๆ เพื่อผลิตผลิตภัณฑ์ผลไม้รวมบรรจุกระป๋องหรือฟรุ๊ตคอกเทลต่อไป จากนั้นคนงานจะบรรจุแวนส์สัปดาห์ลงในกระป๋อง ซึ่งสามารถแบ่งการบรรจุได้ 3 วิธีคือ

3.2.1 แบบปกติ (Regular Pack) น้ำหนักเนื้อสัปดาห์จะต้องไม่น้อยกว่าร้อยละ 58 ของความจุกระป๋อง

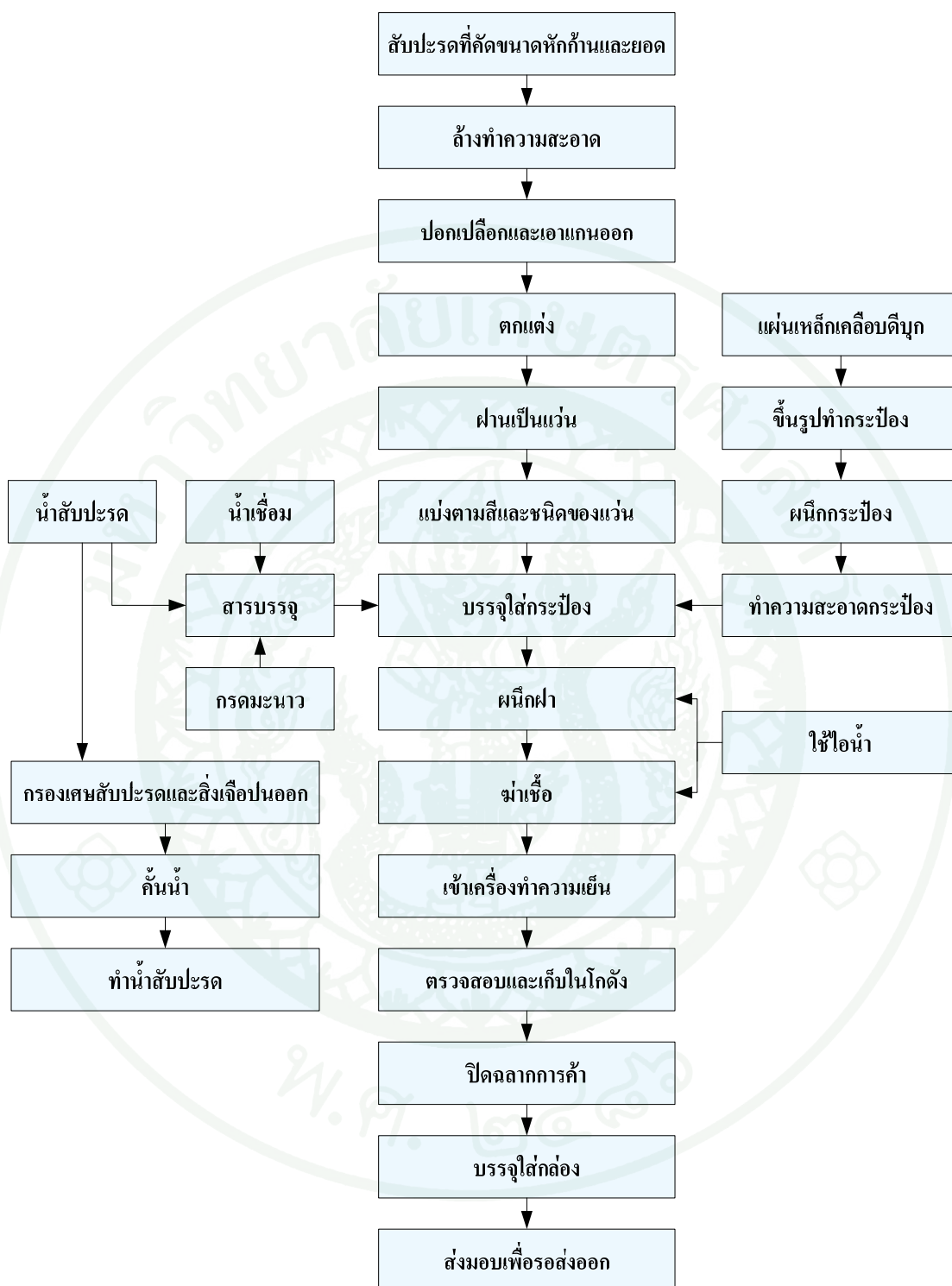
3.2.2 แบบแน่น (Heavy Pack) น้ำหนักเนื้อสัปดาห์จะต้องไม่น้อยกว่าร้อยละ 73 ของความจุกระป๋อง

3.2.3 แบบอัด (Solid Pack) น้ำหนักเนื้อสัปดาห์จะต้องไม่น้อยกว่าร้อยละ 78 ของความจุกระป๋อง

จากนั้นชั่งน้ำหนักให้ได้ตามที่ต้องการและใส่น้ำเชื่อมที่ปรุงรส หรือน้ำสัปดาห์สดตามที่ลูกค้าสั่ง โดยน้ำเชื่อมหรือน้ำสัปดาห์สดที่เติมลงไป เรียกว่า สารบรรจุ (Packing Media)

3.3 ทำการไล่อากาศ แล้วผ่านเข้าเครื่องผนึกฝากระป๋อง และส่งเข้าเครื่องฆ่าเชื้อด้วยไอน้ำ หลังจากนั้นจึงส่งเข้าเครื่องทำความเย็น เพื่อปรับสภาพให้คืนตัวและมีคุณภาพคงที่

3.4 ตรวจสอบสภาพกระป๋อง แล้วนำไปเก็บในคลังสินค้าเพื่อรอการจำหน่าย โดยจะนำสัปดาห์กระป๋องมาปิดฉลากการค้าตามที่ลูกค้าต้องการ จากนั้นจึงบรรจุสัปดาห์กระป๋องใส่กล่องเพื่อส่งออกไป



ภาพที่ 1 ขั้นตอนการผลิตอุตสาหกรรมสับปะรดกระป๋อง

ที่มา: สำนักงานเศรษฐกิจอุตสาหกรรม กระทรวงอุตสาหกรรม (2545)

คุณลักษณะที่กำหนดคุณภาพในการผลิตสับปะรดกระป๋อง

ในกระบวนการผลิตสับปะรดกระป๋องจะมีการวิเคราะห์คุณลักษณะทางเคมีเพื่อกำหนดคุณภาพในระหว่างกระบวนการผลิต โดยค่าที่นิยมใช้วัดค่า คือ ค่าระดับความหวาน การหาค่าดังกล่าวเป็นการพิจารณาในเรื่องรสชาติ การถนอมอาหารและการผลิตสินค้าให้ตรงตามมาตรฐานที่กำหนด

ค่าความหวาน ($^{\circ}\text{Brix}$)

การหาค่าปริมาณของแข็งทั้งหมดในสารละลาย (Total Soluble Solid) ซึ่งอาจประกอบไปด้วยของแข็งหลายชนิดที่ละลายได้ในสารละลายนั้น เช่น กรดต่างๆ เกลือ เป็นต้น ค่าที่ได้มีหน่วยเป็นองศาบริกซ์ ($^{\circ}\text{Brix}$) คิดเป็นร้อยละโดยน้ำหนักของของแข็งที่ละลายได้ทั้งหมดในสารละลายนั้น

ในกระบวนการผลิตสับปะรดกระป๋องจะใช้ค่านี้ในการหาค่าความหวานของวัตถุดิบ ผลิตภัณฑ์ระหว่างกระบวนการผลิตและผลิตภัณฑ์สุดท้าย โดยการวัดความเข้มข้นหรือปริมาณน้ำตาลจากดัชนีการหักเหของแสง (Refractive Index) เทียบกับน้ำตาลมาตรฐาน (นิยมใช้น้ำตาลซูโครส) โดยใช้เครื่องวัดการหักเหของแสงหรือรีแฟรคโตมิเตอร์ (Refractometer) (กล้านรงค์, 2546)

ค่าความหวานของวัตถุดิบมีความสำคัญอย่างมากในกระบวนการผลิต เนื่องจากใช้ในการกำหนดคุณลักษณะของสารบรรจุ (Packing Media) ที่จะต้องเตรียมไว้ล่วงหน้า ให้มีค่าความหวานที่เหมาะสมกับวัตถุดิบที่ใช้ในการผลิต เพื่อให้ได้ผลิตภัณฑ์ที่มีค่าความหวานตรงตามมาตรฐานที่กำหนด ซึ่งการคำนวณค่าความหวานของสารบรรจุในกระบวนการผลิต นิยมใช้ทฤษฎีการเท่ากันของสสารก่อนผสมและหลังผสม จากสมการ

$$B_1W_1 + B_2W_2 = B_3W_3$$

ดังนั้นจะได้สมการที่ (1)

$$B_2 = \frac{B_3W_3 - B_1W_1}{W_2} = \frac{B_3W_3 - B_1W_1}{(W_3 - W_1)} \quad (1)$$

โดย B_1, W_1 คือ ค่าความหวาน ($^{\circ}\text{Brix}$) และน้ำหนักของวัตถุดิบสับปะรด (กรัม) ตามลำดับ

B_2, W_2 คือ ค่าความหวาน ($^{\circ}\text{Brix}$) และน้ำหนักของสารบรรจุ (กรัม) ตามลำดับ

B_3, W_3 คือ ค่าความหวาน ($^{\circ}\text{Brix}$) และน้ำหนักของผลิตภัณฑ์สุดท้าย (กรัม) ตามลำดับ

ดังนั้น การทำนายค่าความหวานของวัตถุดิบสับปะรดที่มีประสิทธิภาพ จะช่วยทำให้การเตรียมสารบรรจุมีความถูกต้อง ลดความล่าช้าในกระบวนการผลิตในการปรับค่าความหวานสารบรรจุให้เหมาะสมกับวัตถุดิบ ทำให้กระบวนการผลิตมีความต่อเนื่องมากขึ้น

ความรู้พื้นฐานสำหรับการพยากรณ์

1. เทคนิคการพยากรณ์

การพยากรณ์ (Forecasting) หมายถึงการทำนาย (Prediction) การคาดการณ์ การประมาณค่าในอนาคตโดยอาศัยข้อมูลหรือเหตุการณ์ที่เกิดขึ้นในอดีตที่มีพื้นฐานองค์ประกอบสถานการณ์เดียวกัน การพยากรณ์เป็นส่วนสำคัญในการวางแผนงานที่สำคัญ สำหรับเทคนิคการพยากรณ์แบ่งได้เป็น 2 เทคนิคใหญ่ ๆ (สมเกียรติ, 2548) ได้แก่

1.1 เทคนิคการพยากรณ์เชิงคุณภาพ (Qualitative Forecasting Techniques)

เป็นการพยากรณ์ที่อาศัยวิจารณญาณ หรือความคิดเห็นของผู้เชี่ยวชาญหรือผู้มีประสบการณ์ในเรื่องนั้นๆ ไม่มีทฤษฎีหรือหลักเกณฑ์สนับสนุน ลักษณะการพยากรณ์จะไม่มีรูปแบบแน่นอน เทคนิคการพยากรณ์เชิงคุณภาพมีหลายวิธี ได้แก่ วิธีเดลไฟ (Delphi Method) วิธีการวิจัยตลาด (Market Research Method) เป็นต้น

1.2 เทคนิคการพยากรณ์เชิงปริมาณ (Quantitative Forecasting Techniques)

เป็นเทคนิคที่ขึ้นกับการวิเคราะห์ข้อมูลในอดีต โดยพยายามหาความสัมพันธ์หรือแนวโน้ม ซึ่งสามารถใช้สำหรับการพยากรณ์ได้ การพยากรณ์เชิงปริมาณอาจแบ่งได้เป็น 2 วิธี คือ

1.2.1 เทคนิคการวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นเทคนิคที่บ่งชี้รูปแบบหรือแนวโน้มในข้อมูล เพื่อใช้พยากรณ์อนาคตภายใต้ข้อสมมติฐานว่า แนวโน้มในอดีตจะยังคงมีต่อไปในอนาคต เทคนิคที่ใช้ในการพยากรณ์วิธีนี้ได้แก่ เทคนิคการเฉลี่ยเคลื่อนที่ (Moving Average) เทคนิคการปรับให้เรียบแบบเอกซ์โปเนนเชียล (Exponential Smoothing) เทคนิคการแยกส่วนประกอบของอนุกรมเวลา (Time Series Decomposition) เป็นต้น

1.2.2 เทคนิคการพยากรณ์ความสัมพันธ์ระหว่างตัวแปร (Causal or Associative Forecasting) เทคนิคนี้มีแนวความคิดว่าพฤติกรรมของสิ่งที่จะพยากรณ์ ถูกกำหนดโดยปัจจัยภายนอก ซึ่งมีอิทธิพลต่อสิ่งที่จะพยากรณ์ในรูปแบบความสัมพันธ์บางลักษณะ วิธีการพยากรณ์โดยเทคนิคดังกล่าว ได้แก่ เทคนิคการวิเคราะห์การถดถอย (Regression Analysis) และเทคนิคการพยากรณ์เชิงเศรษฐมิติ (Econometric Forecasting)

2. การวัดความแม่นยำของเทคนิคการพยากรณ์

การวัดความแม่นยำของเทคนิคการพยากรณ์จะเป็นสิ่งสำคัญที่จะต้องคำนวณ โดยวัดจากขนาดของความคลาดเคลื่อนสามารถวัดได้ดังนี้

กำหนดให้	y_t	แทนค่าจริงของอนุกรมเวลาที่เวลา t
	$\hat{y}_t$	แทนค่าของพยากรณ์อนุกรมเวลาที่เวลา t
	$y_t - \hat{y}_t$	แทนค่าความคลาดเคลื่อนของค่าพยากรณ์ (Forecast Error)
	n	แทนจำนวนข้อมูลที่ทดสอบนั้น

2.1 ค่าเฉลี่ยของค่าความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error, MAE) คำนวณโดยการเฉลี่ยขนาดของความคลาดเคลื่อนจากการพยากรณ์ ในรูปค่าสมบูรณ์ของค่าความคลาดเคลื่อนแต่ละค่า

$$MAE = \frac{\sum_{i=1}^n (|y_i - \hat{y}_i|)}{n} = \frac{\sum_{i=1}^n |e_i|}{n} \quad (2)$$

2.2 ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Squared Error, MSE) คำนวณโดยการเฉลี่ยค่ากำลังสองของความคลาดเคลื่อน

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (3)$$

รากที่สองของ MSE (Root Mean Squared Error, RMSE) เป็นค่าประมาณขององค์ประกอบส่วนในการพยากรณ์ สามารถใช้วัดค่าความคลาดเคลื่อนได้เช่นกัน

2.3 ค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Percentage Error, MAPE) คำนวณโดยการเฉลี่ยร้อยละผลรวมของความคลาดเคลื่อน โดยไม่คำนึงถึงเครื่องหมาย

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100}{n} \quad (4)$$

2.4 ร้อยละค่าคลาดเคลื่อน (Mean Percentage Error, MPE) คำนวณโดยการเฉลี่ยร้อยละผลรวมของความคลาดเคลื่อน

$$MPE = \frac{\sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right) \times 100}{n} \quad (5)$$

3. ตัวแปรเชิงกลุ่ม

ในการศึกษาหาความสัมพันธ์ระหว่างตัวแปรหลายๆตัวนั้น ในบางครั้งมีตัวแปรเชิงกลุ่มเข้ามาเกี่ยวข้องด้วย เช่น รายจ่ายของคนอาจจะขึ้นอยู่กับรายได้และเพศ โดยรายจ่ายเป็นตัวแปรตามซึ่ง

เป็นตัวแปรเชิงปริมาณ คือ สามารถวัดค่ามากหรือน้อยได้ และตัวแปรอิสระคือรายได้ซึ่งเป็นตัวแปรเชิงปริมาณเช่นกัน ขณะที่ตัวแปรเพศจะแบ่งเป็น เพศชาย และเพศหญิง ซึ่งไม่สามารถวัดค่าออกมาเป็นปริมาณที่มากหรือน้อยได้ เพศจึงเป็นตัวแปรเชิงคุณภาพ (Qualitative Variable) หรือตัวแปรเชิงกลุ่ม การหาความสัมพันธ์ระหว่างเพศกับรายจ่ายต้องมีการกำหนดค่าให้กับตัวแปรเพศ เช่น กำหนดให้เพศชาย มีค่าเป็น 0 และเพศหญิงมีค่าเป็น 1 ทำให้สามารถวัดค่าความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้ง่าย ซึ่งจะเรียกตัวแปรที่ทำการปรับค่านี้นี้ว่า Categorical Variable หรือตัวแปรเทียม หรือตัวแปรหุ่น (Dummy Variable) (กัลยา, 2549)

ดังนั้น ตัวแปรหุ่น (Dummy Variable) จึงเป็นตัวแปรที่สร้างขึ้น เพื่อระบุกลุ่มหรือชุดที่ค่าสังเกตนั้นอยู่ มีค่าเป็น 1 หรือ 0 โดยกำหนดค่าเท่ากับ 1 เมื่อค่าสังเกตอยู่ในกลุ่มที่สนใจและมีค่าเท่ากับ 0 เมื่อค่าสังเกตไม่อยู่ในกลุ่มที่สนใจ

ทฤษฎีที่ใช้ในการศึกษา

1. แบบจำลองการถดถอย (Regression Model)

แบบจำลองการถดถอย เป็นวิธีการทางสถิติที่ใช้ในการประมาณค่าหรือการทำนายค่าของตัวแปรตัวหนึ่ง โดยใช้ค่าของข้อมูลอีกตัวหนึ่งหรือหลายตัวเป็นตัวทำนายค่า ตัวแปรที่สนใจศึกษาหรือผลที่วัดได้ เรียกว่าตัวแปรตาม (Dependent Variable หรือ Response Variable) และตัวแปรอื่นที่แทนปัจจัยที่ทำให้ค่าของตัวแปรตามเปลี่ยนแปลงหรือตัวแปรทำนายค่า เรียกว่า ตัวแปรอิสระหรือตัวแปรต้น (Independent Variable หรือ Predictor Variable) แบบจำลองประเภทนี้มีข้อสมมติหรือเงื่อนไขทางสถิติที่ค่าความคลาดเคลื่อนสุ่มของแบบจำลอง ต้องมีการแจกแจงแบบปกติมาตรฐาน มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนคงที่ ($\varepsilon_i \sim N(0, \sigma^2)$) และหากมีตัวแปรอิสระมากกว่า 1 ตัวตัวแปรอิสระต้องไม่มีสหสัมพันธ์กัน (Multicollinearity) (กัลยา, 2549)

การวิเคราะห์การถดถอยกรณีหนึ่งตัวแปรอิสระ เรียกว่า การวิเคราะห์การถดถอยแบบง่าย (Simple Regression Analysis) หรือใช้ตัวแปรอิสระหลายตัวมาทำนายตัวแปรตามจะเรียกว่า การวิเคราะห์การถดถอยแบบพหุ (Multiple Regression Analysis) ถ้าตัวแปรอิสระต่าง ๆ มีการยกกำลัง

ต่าง ๆ จะเรียกว่า การถดถอยแบบพหุนาม (Polynomial Regression) (กัลยา, 2549; ทรงศิริ, 2548)

1.1 แบบจำลองการถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Model)

รูปแบบการถดถอยเชิงเส้นแบบง่ายเป็นรูปแบบที่กำหนดตัวแปรตามมีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระเพียงตัวแปรเดียว สามารถแสดงความสัมพันธ์ในรูปสมการเชิงเส้นได้ ดังสมการที่ (6)

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad ; i = 1, 2, \dots, n \quad (6)$$

โดยที่	Y_i	เป็นค่าของตัวแปรตามที่ i
	X_i	เป็นค่าของตัวแปรอิสระที่ i
	β_0	เป็นจุดที่เส้นการถดถอยตัดแกน Y หรือจุดตัดแกนตั้ง (Y-intercept)
	β_1	เป็นค่าความชันของเส้นตรง (Slope) แสดงอัตราการเพิ่มหรือลดของตัวแปรตาม Y เมื่อค่าของตัวแปรอิสระ X เพิ่มขึ้นหนึ่งหน่วย หรือเรียกค่าสัมประสิทธิ์การถดถอย (Regression Coefficient)
	ε_i	เป็นความคลาดเคลื่อนสุ่ม (Random Error) ของค่าสังเกตที่ i ในแบบจำลอง

1.2 แบบจำลองการถดถอยเชิงพหุ (Multiple Regression Model)

รูปแบบการถดถอยเชิงเส้นโดยการวิเคราะห์การถดถอย เมื่อมีตัวแปรที่เกี่ยวข้องหลายตัวแปร กำหนดให้ตัวแปรหนึ่งเป็นตัวแปรตาม ส่วนตัวแปรอื่นที่มีอิทธิพลต่อตัวแปรตามเป็นตัวแปรอิสระจำนวน k ตัว แสดงอยู่ในรูปแบบจำลองการถดถอยเชิงเส้นเชิงพหุ ดังสมการที่ (7)

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i \quad ; i = 1, 2, \dots, n \quad (7)$$

โดยที่	Y_i	เป็นตัวแปรตาม
	$\beta_0, \beta_1, \dots, \beta_k$	เป็นพารามิเตอร์ในแบบจำลอง $k+1$ ตัว
	$X_{i1}, X_{i2}, \dots, X_{ik}$	เป็นตัวแปรอิสระ k ตัว; $i = 1, 2, \dots, n$
	ε_i	เป็นความคลาดเคลื่อนสุ่ม (Random Error) ของค่าสังเกตที่ i ในแบบจำลอง

1.3 แบบจำลองการถดถอยแบบพหุนาม (Polynomial Regression Model)

ถ้าตัวแปรอิสระหลายตัวที่ยกกำลังสูงกว่าหนึ่ง (Higher Order) มีความสัมพันธ์กับตัวแปรตาม จะได้สมการการถดถอยแบบยกกำลัง ซึ่งแสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ดังสมการที่ (8)

$$Y_i = \sum_{j=1}^1 \sum_{k=1}^p \beta_k Z_k(X_{ij}) + \varepsilon_i \quad (8)$$

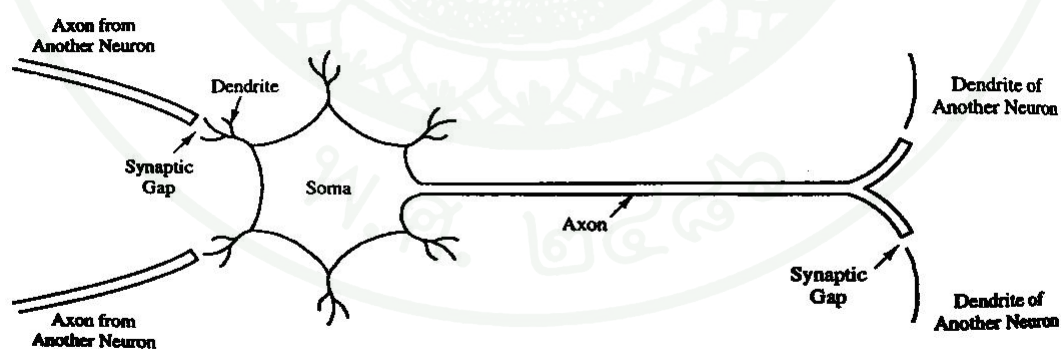
โดยที่	Y_i	ค่าของตัวแปรตามจากการทดลองที่ i
	X_{ij}	ค่าของตัวแปรอิสระที่ j จากการทดลองที่ i โดย X_{ij} เป็นอิสระต่อกัน ($j = 1, 2, 3, \dots, k$)
	β_k	สัมประสิทธิ์การถดถอยเชิงส่วน โดยที่ β_k แสดงถึงการเปลี่ยนแปลงของตัวแปรตาม Y_i เมื่อ $Z_k(X_{ij})$ เปลี่ยนไป 1 หน่วย โดยที่ตัวแปรอื่นๆ มีค่าคงที่
	$Z_k(X_{ij})$	รูปแบบของสมการที่เป็นฟังก์ชันของ X_{ij} โดยอาจอยู่ในรูปกำลังหนึ่ง กำลังสอง หรือสูงกว่า หรือปฏิสัมพันธ์ของ X_{ij}
	ε_i	ค่าความคลาดเคลื่อนแบบสุ่ม (Random Error) ของค่าสังเกตที่ i ในแบบจำลอง

อัมรินทร์ (2548) ได้สรุปถึงข้อดีของแบบจำลองการถดถอย คือ สามารถสร้างและใช้งานได้ง่าย และมีกระบวนการทางสถิติรับรอง ตลอดจนมีโปรแกรมสำเร็จรูปที่ใช้ในการสร้างแบบ

จำลองหลากหลาย แต่มีข้อเสีย คือ ในการวิเคราะห์การถดถอยแบบโพลิโนเมียลที่มีรูปแบบของตัวแปรอิสระที่ยากำลึงสูงๆ ในการใช้งาน อาจไม่เสถียรหรือมีการเหวี่ยง (Oscillation) และมีความไวต่อสมมติฐานทางสถิติ ในการหาความน่าเชื่อถือในแบบจำลองจะขึ้นกับข้อสมมติทางสถิติ เช่น กรณีที่ตัวแปรอิสระมีสหสัมพันธ์กันสูง จะต้องทำการรวมกลุ่มตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มเดียวกันก่อน เพื่อแก้ปัญหาคอตัวแปรอิสระมีสหสัมพันธ์กัน เป็นต้น

2. แบบจำลองเครือข่ายประสาทเทียม (Artificial Neural Network)

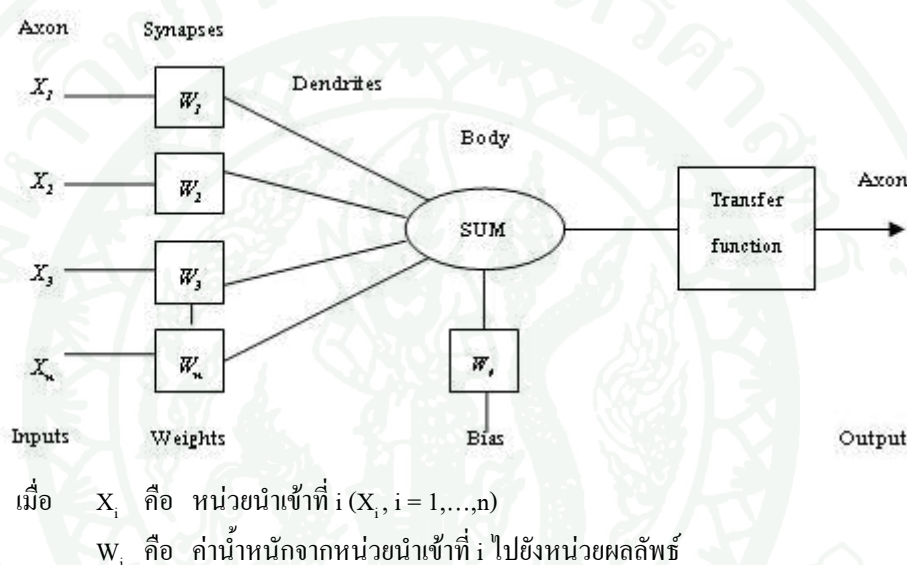
เครือข่ายประสาทเทียม เป็นระบบการประมวลผลข้อมูล แบบจำลองถูกพัฒนาขึ้นเพื่อเลียนแบบการทำงานของระบบประสาทของมนุษย์ (Biological Neural Network) หรือระบบการรับรู้ของมนุษย์ นิวรอนของระบบประสาทของมนุษย์ ประกอบด้วยองค์ประกอบที่สำคัญ 3 อย่าง ได้แก่ เดนไดรต์ (Dendrites) โซมา (Soma) หรือตัวเซลล์ (Cell Body) และแอกซอน (Axon) โดยเดนไดรต์ทำหน้าที่รับสัญญาณจากนิวรอนตัวอื่นๆ ที่อยู่ในรูปคลื่นกระแสไฟฟ้า และส่งผ่านสัญญาณนั้นไปยังช่องว่างไซแนป (Synaptic Gap) ซึ่งทำหน้าที่ปรับระดับสัญญาณเหมือนกับค่าน้ำหนักในเครือข่ายประสาทเทียม จากนั้นเดนไดรต์จะรับสัญญาณให้กับตัวเซลล์ เพื่อทำการรวมสัญญาณและส่งผ่านไปยังแอกซอนเพื่อส่งผ่านไปยังนิวรอนหน่วยอื่นๆต่อไป การทำงานของระบบประสาทของมนุษย์แสดงในภาพที่ 2



ภาพที่ 2 การทำงานของระบบประสาทของมนุษย์

ที่มา: Fausett (1994)

สำหรับการทำงานของเครือข่ายประสาทเทียมนั้น ประกอบด้วยนิวรอนหรือหน่วยประมวลผลจำนวนมากเรียงตัวเป็นชั้น ได้แก่ ชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) มีการเชื่อมโยงกันเป็นเครือข่าย แต่ละเส้นการเชื่อมโยงจะมีน้ำหนัก (Weight) กำกับอยู่ ข้อมูลหรือสัญญาณจากนิวรอนนำเข้าจะถูกส่งผ่านเส้นเชื่อมโยงพร้อมกับปรับน้ำหนักไปยังนิวรอนประมวลผลถัดไป ซึ่งจะรวมสัญญาณที่ปรับน้ำหนักแล้วและนำไปผ่านฟังก์ชันกระตุ้น (Activation Function หรือ Transfer Function) เพื่อคำนวณผลลัพธ์ต่อไป ดังแสดงในภาพที่ 3



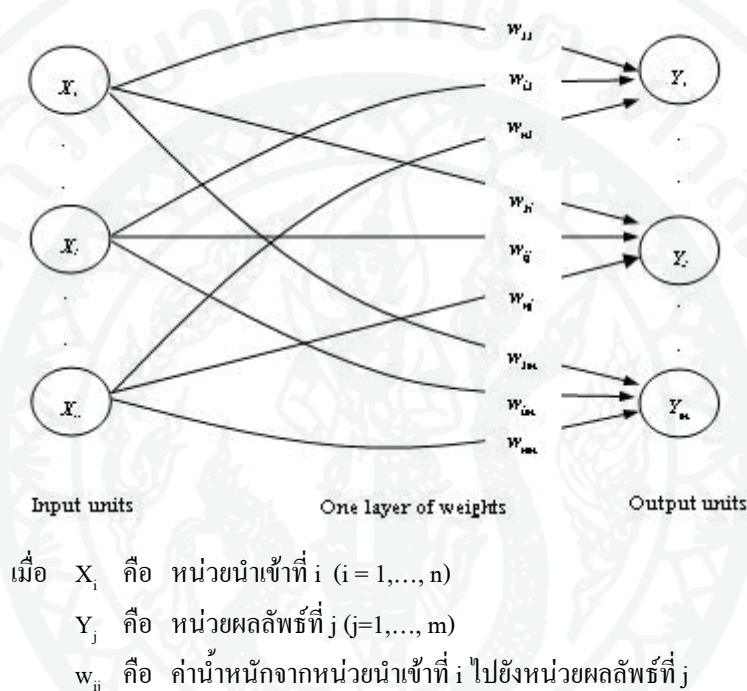
ภาพที่ 3 การทำงานของเครือข่ายประสาทเทียม

ที่มา: Page *et al* (1993)

ลักษณะสำคัญที่ใช้ในการจำแนกประเภทของเครือข่ายประสาทเทียมได้แก่ สถาปัตยกรรมหรือโครงสร้างของเครือข่ายประสาทเทียม (Architecture) การฝึกหรือการเรียนรู้ของเครือข่ายประสาทเทียม (Training/Learning) และฟังก์ชันกระตุ้น ซึ่งมีรายละเอียดดังต่อไปนี้

2.1 โครงสร้างหรือสถาปัตยกรรมของเครือข่ายประสาทเทียม เป็นการจัดเรียงชั้นของนิวรอน และรูปแบบการเชื่อมโยงระหว่างนิวรอน ภายในเครือข่ายประสาทเทียม จัดแบ่งโครงสร้างพื้นฐานตามจำนวนชั้น แบ่งออกเป็นเครือข่ายแบบชั้นเดียว (Single-layer Net) และเครือข่ายแบบหลายชั้น (Multilayer Net)

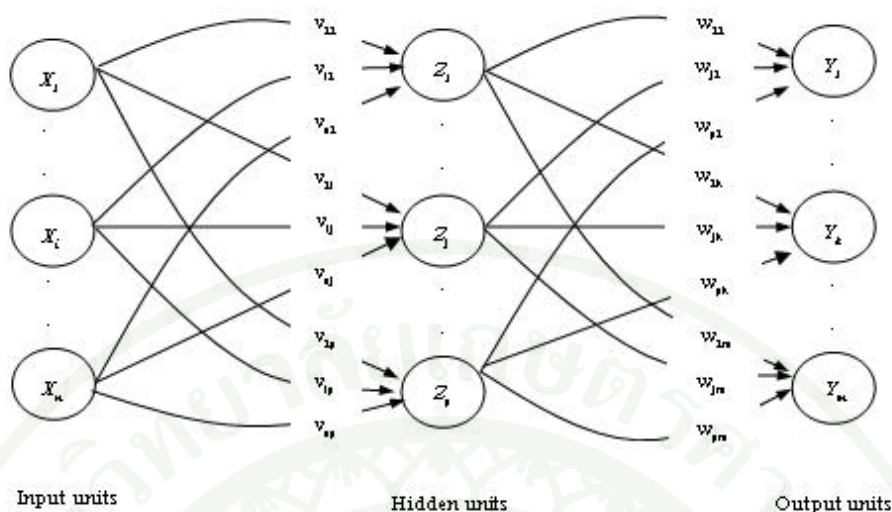
2.1.1 เครือข่ายแบบชั้นเดียว เป็นเครือข่ายประสาทเทียมอย่างง่าย ประกอบด้วยชั้นนำเข้าและชั้นผลลัพธ์ มีการเชื่อมโยงเพียง 1 ชั้นเท่านั้น การส่งผ่านข้อมูลกระทำจากชั้นนำเข้าส่งไปยังชั้นผลลัพธ์ ซึ่งจะรวบรวมข้อมูลที่ได้รับทำการปรับน้ำหนักแล้วส่งผ่านฟังก์ชันกระตุ้น เครือข่ายประสาทเทียมประเภทนี้จะใช้แก้ปัญหาเชิงเส้นที่มีความซับซ้อนน้อย ตัวอย่างเครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าชั้นเดียว (Single-layer Feedforward Networks) แสดงในภาพที่ 4



ภาพที่ 4 เครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าชั้นเดียว

ที่มา: Fausett (1994)

2.1.2 เครือข่ายแบบหลายชั้น เป็นเครือข่ายประสาทเทียมที่ประกอบด้วยชั้นซ่อน (Hidden Layer) อย่างน้อย 1 ชั้นขึ้นไปแทรกอยู่ระหว่างชั้นนำเข้าและชั้นผลลัพธ์ใช้ในการแก้ปัญหาที่มีความซับซ้อน เครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าหลายชั้น (Multiplayer Feed-forward Networks) แสดงในภาพที่ 5



เมื่อ X_i คือ หน่วยนำเข้าที่ i ($i = 1, \dots, n$)

Z_j คือ หน่วยซ่อนที่ j ($j = 1, \dots, p$)

Y_k คือ หน่วยผลลัพธ์ที่ k ($k = 1, \dots, m$)

v_{ij} คือ ค่าน้ำหนักจากหน่วยนำเข้าที่ i ไปยังหน่วยซ่อนที่ j

w_{jk} คือ ค่าน้ำหนักหน่วยซ่อนที่ j ไปยังหน่วยผลลัพธ์ที่ k

ภาพที่ 5 เครือข่ายประสาทเทียมแบบป้อนสู่ด้านหน้าหลายชั้น

ที่มา: Fausett (1994)

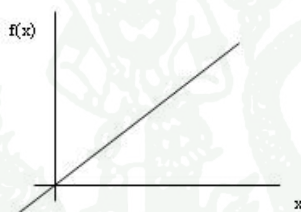
2.2 การเรียนรู้ของเครือข่ายประสาทเทียม เป็นวิธีในการกำหนดค่าน้ำหนักให้กับการเชื่อมโยง สามารถแบ่งได้เป็น 2 วิธี คือการเรียนรู้แบบมีผู้สอน (Supervised Learning) และการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

2.2.1 การเรียนรู้แบบมีผู้สอน การเรียนรู้จะเกิดขึ้นได้โดยการผ่านชุดเวกเตอร์นำเข้า (Input Vector หรือ Training Vector) พร้อมด้วยเวกเตอร์ผลลัพธ์หรือค่าเป้าหมายที่ต้องการ (Target Output Vector) ให้กับเครือข่าย เพื่อเรียนรู้ความสัมพันธ์และปรับค่าน้ำหนักของการเชื่อมโยงตามกลไกการเรียนรู้ (Learning Algorithms) เพื่อให้เกิดค่าคลาดเคลื่อนระหว่างค่าผลลัพธ์กับค่าเป้าหมายน้อยที่สุด เครือข่ายประสาทเทียมที่ใช้การเรียนรู้แบบมีผู้สอน ได้แก่ เครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation) เรเดียลเบสฟังก์ชัน (Radial Basis Function) และเครือข่ายแบบแบ่งกลุ่ม (Association Nets)

2.2.2 การเรียนรู้แบบไม่มีผู้สอน เครือข่ายประสาทเทียมทำการเรียนรู้ด้วยตนเองโดยไม่มีการกำหนดเวกเตอร์ผลลัพธ์เป้าหมาย เครือข่ายประสาทเทียมจะปรับค่าน้ำหนัก เพื่อจัดกลุ่มเวกเตอร์นำเข้าที่เหมือนกันไว้ในกลุ่มเดียวกัน เครือข่ายประสาทเทียมที่ใช้การเรียนรู้ แบบไม่มีผู้สอน ได้แก่ โคโฮเนนซอม (Kohonen Self Organizing Map; Kohonen SOM) และอาร์ท (Adaptive Resonance Theory Network; ART)

2.3 ฟังก์ชันกระตุ้น เป็นฟังก์ชันที่ใช้ในการส่งผลรวมของน้ำหนักไปยังขั้นต่อไป ฟังก์ชันกระตุ้นที่นิยมใช้ ได้แก่

2.3.1 ฟังก์ชันปั้งซี่ (Identity Function) ใช้กับหน่วยนำเข้า สำหรับการรับและส่งผ่านสัญญาณเข้าสู่ระบบ ดังสมการ $f(x) = x$ (ภาพที่ 6)

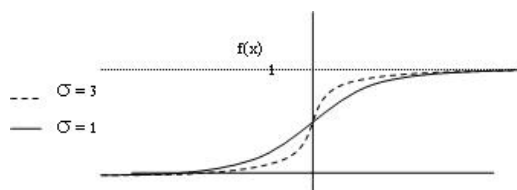


ภาพที่ 6 ฟังก์ชันปั้งซี่

ที่มา: Fausett (1994)

2.3.2 ฟังก์ชันซิกมอยด์ (Sigmoid Function)

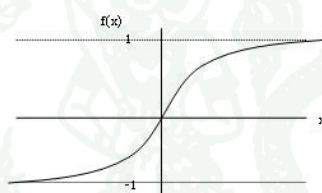
ก. ฟังก์ชันไบนารีซิกมอยด์ (Binary Sigmoid Function) มีค่าอยู่ระหว่าง 0 ถึง 1
 ดังสมการ $f(x) = \frac{1}{1 + e^{-\sigma x}}$ เมื่อ σ คือ ค่าพารามิเตอร์ของความชัน (ภาพที่ 7)



ภาพที่ 7 ฟังก์ชันซิกมอยด์

ที่มา: Fausett (1994)

ข. ฟังก์ชันไฮเพอร์โบลิคแทนเจน (Hyperbolic Tangent; tanH) มีค่าอยู่ระหว่าง -1 ถึง 1 และค่าเป็น 1 (ภาพที่ 8)



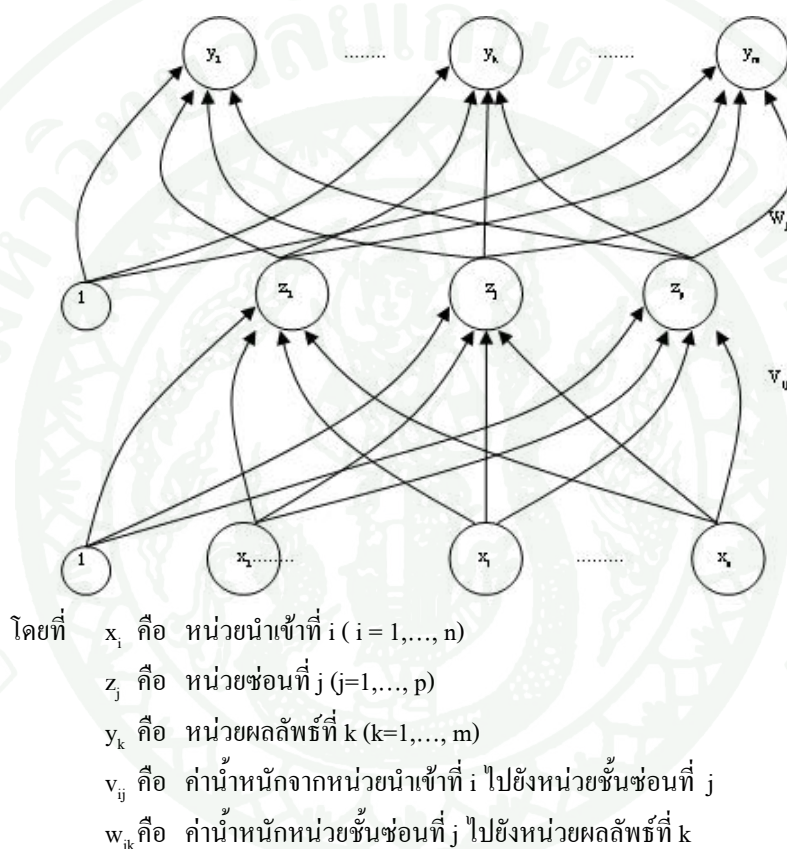
ภาพที่ 8 ฟังก์ชันไฮเพอร์โบลิคแทนเจน

ที่มา: Fausett (1994)

2.4 เครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับ (Back-Propagation Neural Network)

เครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับ มีลักษณะเป็นเครือข่ายหลายชั้นที่มีการส่งข้อมูลไปข้างหน้าแบบสมบูรณ์ ดังแสดงในภาพที่ 9 ในการเรียนรู้แบบมีผู้สอนกลไกการเรียนรู้จะส่งข้อมูลความคลาดเคลื่อนย้อนกลับ (Back-propagation of Error) เพื่อฝึกให้เครือข่ายตอบสนองได้อย่างถูกต้อง อย่างไรก็ตามเครือข่ายที่ดีจะต้องมีสมดุลระหว่างความสามารถในการตอบสนองรูปแบบที่ใช้สำหรับการเรียนรู้หรือการฝึกได้อย่างถูกต้อง (Memorization) และความสามารถในการตอบสนองต่อรูปแบบอื่นๆ ที่ไม่ใช้ในการฝึกหรือความสามารถในการใช้งานทั่วไปได้ดีเช่นกัน (Generalization) (Fausett, 1994) เครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับถือเป็นวิธีที่

นิยมใช้มากที่สุดและเป็นตัวประมาณค่าแบบสากลในทางทฤษฎี (Theoretic Universal Approximator) (Funahashi, 1989; Hornik *et al.*, 1989) สามารถใช้สำหรับการจำลองข้อมูล (Data Modeling) หรือประมาณค่าฟังก์ชัน (Function Approximation) การจำแนกกลุ่ม (Classification) การควบคุม (Control) การย่อข้อมูลและรูปภาพให้มีขนาดเล็กลง (Data and Image Compression) และการจดจำรูปแบบ (Pattern Recognition) (Hassoun, 1995)



ภาพที่ 9 โครงสร้างของเครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ

ที่มา: Fausett (1994)

การฝึกเครือข่ายโดยการส่งถ่ายข้อมูลย้อนกลับประกอบด้วย 3 ขั้นตอน ได้แก่ การป้อนข้อมูลชุดสำหรับการฝึกไปข้างหน้า การถ่ายทอดแบบย้อนกลับของข้อมูลความคลาดเคลื่อน และการปรับน้ำหนักของทั้งเครือข่าย ในระหว่างการส่งข้อมูลไปข้างหน้าแต่ละหน่วยนำเข้า (x_i) จะได้รับสัญญาณนำเข้าและส่งสัญญาณไปยังแต่ละหน่วยซ่อน (z_j) หน่วยซ่อนจะคำนวณสัญญาณกระตุ้น (z_j) และส่งไปยังหน่วยผลลัพธ์ (y_k) เพื่อคำนวณค่ากระตุ้น (y_k) เพื่อเป็นผลลัพธ์ของเครือข่าย

สำหรับรูปแบบการนำเข้าหนึ่งๆ ในระหว่างการเรียนรู้หน่วยผลลัพธ์แต่ละหน่วยจะเปรียบเทียบผลลัพธ์จากการคำนวณกับค่าเป้าหมาย (t_k) เพื่อคำนวณข้อมูลความคลาดเคลื่อน ε_k สำหรับกระจายความคลาดเคลื่อนที่หน่วยผลลัพธ์กลับไปยังหน่วยซ่อน เพื่อทำการปรับค่าน้ำหนักระหว่างชั้นผลลัพธ์และชั้นซ่อน ในทำนองเดียวกันจะคำนวณข้อมูลความคลาดเคลื่อน ε_{kj} เพื่อใช้ปรับน้ำหนักระหว่างชั้นซ่อนและชั้นนำเข้า ขั้นตอนการฝึกทั้ง 3 ขั้นตอนนี้ จะดำเนินการซ้ำไปซ้ำมาหลายรอบการฝึก จนกระทั่งได้ค่าความคลาดเคลื่อนต่ำตามที่ต้องการ หรือเมื่อค่าน้ำหนักไม่เปลี่ยนแปลงหรือถึงจำนวนรอบการฝึกที่กำหนดไว้ล่วงหน้า (Fausett, 1994)

ข้อดีของเครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ คือ สามารถใช้งานได้หลากหลายทั้งในการประมาณค่าฟังก์ชัน (Function Approximation) การทำนายค่าต่างๆ และการจำแนกกลุ่ม (Classification) เป็นตัวประมาณค่าแบบสากลในทางทฤษฎี (Theoretic Universal Approximator) ในการฝึกเครือข่ายและการปฏิบัติมีความน่าเชื่อถือตามทฤษฎีพื้นฐานที่สามารถเข้าใจได้ง่าย และมีโปรแกรมให้เลือกใช้ในการสร้างแบบจำลองที่หลากหลาย เช่น Neuralware , SPSS Clementine และ MATLAB เป็นต้น แต่ข้อเสียของเครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ คือ เรียนรู้ช้า เนื่องจากต้อง การจำนวนรอบในการเรียนรู้มาก การเลือกโครงสร้างที่เหมาะสมทำได้ยาก และมีความไวต่อการมีพารามิเตอร์ของแบบจำลอง (น้ำหนัก) มากเกินไป (Over-parameterization) และต่อการเกิดการเรียนรู้มากเกินไป (Overtraining) ทำให้เครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับเข้าสู่สภาวะการจดจำข้อมูลชุดที่เรียนรู้มากเกินไป ทำให้ความสามารถในการใช้งานกับข้อมูลชุดที่ไม่ได้ใช้ในการฝึกหรือความสามารถใช้งานทั่วไปลดลง (Hassoun, 1995)

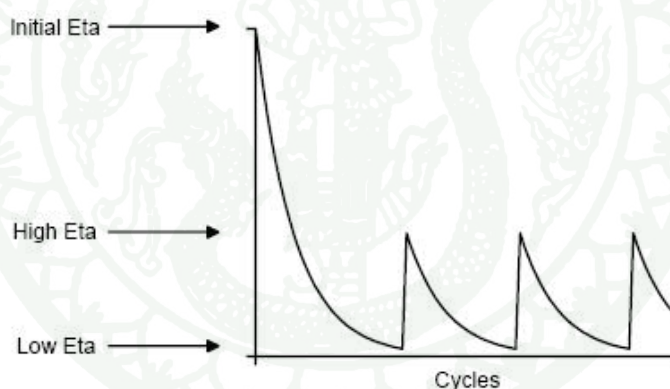
2.5 เครือข่ายประสาทเทียม ในโปรแกรม SPSS Clementine 12.0

โปรแกรม SPSS Clementine 12.0 สร้างเครือข่ายแบบป้อนสู่ด้านหน้าหลายชั้น (Multilayer Feed-Forward Layers) แบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation of Errors) โดยใช้การเรียนรู้ตามหลัก Delta Rule ฟังก์ชันกระตุ้นแบบซิกมอยด์ (Sigmoid or Logistic Transfer Function) ในขั้นตอนเริ่มต้นเรียนรู้ น้ำหนักทั้งหมดในเครือข่ายจะถูกกำหนดโดยกระบวนการสุ่ม กำหนดค่าอยู่ในช่วง -0.5 ถึง 0.5 ในแต่ละรอบการเรียนรู้เรียกว่า Cycles หรือ Epochs จะทำการเรียนรู้ผ่านข้อมูลชุดเรียนรู้ นำไปสู่การทำนายค่าชั้นผลลัพธ์ การทำนายค่าจะเปรียบเทียบระหว่างค่าคำนวณกับค่าเป้าหมาย เพื่อส่งค่าคลาดเคลื่อนไปปรับน้ำหนักของเครือข่าย โดยค่าน้ำหนักที่ถูกปรับจะคำนวณจากสมการที่ (9)

$$\Delta w_{ij}(n+1) = \eta \delta_{pj} o_{pi} + \alpha \Delta w_{ij}(n) \quad (9)$$

โดยที่ η คือ อัตราการเรียนรู้ (The Learning Rate Parameter)
 δ_{pj} คือ ค่าคลาดเคลื่อนที่ส่งย้อนกลับ (The Back-propagated Error Value)
 o_{pi} คือ ค่าผลลัพธ์ของนิวรอน i จากกระเปาะ (Record) p
 α คือ ค่าโมเมนตัม (The Momentum Parameter)
 $\Delta w_{ij}(n)$ คือ ค่าเปลี่ยนแปลงของน้ำหนักในรอบก่อนหน้า

ทั้งนี้ค่า α จะเป็นค่าคงที่ระหว่างการเรียนรู้เครือข่าย แต่ค่า η จะเปลี่ยนแปลงไประหว่างรอบการฝึก โดยจะเริ่มจากค่าที่กำหนดโดยผู้ใช้โปรแกรม เป็นค่า *initial eta* และจะมีค่าลดลงเป็นแบบลอการิทึมไปจนถึงค่า *low eta* และกลับมาสูงขึ้นเท่ากับค่า *high eta* และลดลงจนถึงค่า *low eta* อย่างนี้ไปเรื่อยๆ การเปลี่ยนแปลงค่าแสดงในภาพที่ 10



ภาพที่ 10 การเปลี่ยนแปลงอัตราการเรียนรู้ระหว่างการฝึกเครือข่าย

ที่มา: SPSS Inc. (2007)

รูปแบบวิธีการเรียนรู้ในการสร้างแบบจำลองเครือข่ายประสาทเทียม โดยโปรแกรม SPSS Clementine 12.0 กำหนดไว้ 5 รูปแบบ (SPSS Inc., 2007) ดังนี้

2.5.1 Quick Method เป็นวิธีที่ใช้กฎนิ้วหัวแม่มือ (Rules of Thumb) และคุณสมบัติของข้อมูลในการกำหนดโครงสร้างเครือข่าย (Topology) โดยเน้นการหาโครงสร้างที่มีหน่วยซ่อน

น้อยและเรียนรู้ชุดข้อมูลได้รวดเร็ว โดยจะการใช้การเรียนรู้เครือข่ายแบบส่งข้อมูลความคลาดเคลื่อนย้อนกลับ (Back-propagation Method) วิธีนี้อาจให้ค่าความถูกต้องของเครือข่ายน้อย ทั้งนี้ผู้ใช้สามารถเพิ่มขนาดของชั้นซ่อนได้ในส่วนผู้เชี่ยวชาญ (Expert Tab) หรือเลือกใช้รูปแบบวิธีการอื่นที่มีความซับซ้อนมากขึ้น

2.5.2 Dynamic Method ในวิธีนี้โครงสร้างเครือข่ายจะเปลี่ยนไประหว่างการฝึก โดยโปรแกรมจะเริ่มจากการกำหนดโครงสร้างเครือข่ายเริ่มต้นก่อน และทำการเพิ่มหรือลดจำนวนหน่วยชั้นซ่อนในระหว่างการเรียนรู้ เพื่อหาโครงสร้างที่ดีที่สุด ขั้นตอนประกอบไปด้วย 2 ขั้นตอนหลัก คือ การหาโครงสร้างที่เหมาะสม และการเรียนรู้ตามโครงสร้างที่เหมาะสม

2.5.3 Multiple Method เป็นวิธีที่สร้างหลายๆ โครงสร้างเครือข่ายเพื่อเปรียบเทียบค่า (จำนวนจะขึ้นอยู่กับข้อมูลชุดเรียนรู้) เครือข่ายจะถูกฝึกในแบบ Pseudo-parallel และในขั้นตอนสุดท้ายจะเลือกโครงสร้างเครือข่ายที่ให้ค่าความคลาดเคลื่อน RMSE ที่ต่ำที่สุดโดยผู้ใช้งานสามารถกำหนดจำนวนหน่วยในชั้นซ่อนที่ต้องการเปลี่ยนแปลงได้

2.5.4 Prune Method วิธีนี้จะตรงกันข้ามกับวิธี Dynamic โดยจะเริ่มจากการสร้างโครงสร้างเครือข่ายขนาดใหญ่และทำการตัดแต่งหรือลดหน่วยในชั้นซ่อนและชั้นนำเข้าที่มีความสำคัญน้อยออกไป วิธีนี้จะใช้เวลาในการสร้างเครือข่ายค่อนข้างนาน แต่มักจะให้ประสิทธิภาพที่ดีเมื่อเปรียบเทียบกับวิธีการอื่น มีขั้นตอนในการทำงาน 2 ขั้นตอนหลัก คือ การตัดแต่งนิวรอนในชั้นซ่อน และการตัดแต่งนิวรอนในชั้นนำเข้า

2.5.5 Exhaustive Prune Method วิธีนี้จะคล้ายกับวิธี Prune โดยจะเริ่มจากโครงสร้างเครือข่ายขนาดใหญ่ และลดจำนวนหน่วยในชั้นซ่อนและชั้นนำเข้าที่มีความสำคัญต่อแบบจำลองน้อยออกไป วิธีนี้จะมีการเลือกปัจจัยที่เกี่ยวข้องอย่างละเอียด ทำให้ได้โครงสร้างที่มั่นใจได้ว่าเป็นโครงสร้างที่ดีที่สุด ทั้งนี้วิธีนี้จะใช้เวลานานมาก โดยเฉพาะเมื่อชุดข้อมูลมีขนาดใหญ่

3. แบบจำลองต้นไม้การตัดสินใจ (Decision Trees)

ต้นไม้การตัดสินใจเป็นวิธีการเรียนรู้ที่ใช้มากที่สุดแบบหนึ่งและเป็นการจำแนกข้อมูลที่มีลักษณะคล้ายกันเป็นกลุ่มของข้อมูล (Class) โดยพิจารณาจากค่าคุณลักษณะข้อมูล (บุญเสริม, 2545)

3.1 องค์ประกอบของต้นไม้การตัดสินใจ

- (1) โหนดภายใน (Internal Node) คือคุณลักษณะของข้อมูล โหนดภายในที่เป็นจุดเริ่มต้นของต้นไม้เรียกว่าโหนดราก (Root Node)
- (2) กิ่ง (Branch) คือค่าที่เป็นไปได้ของแต่ละคุณลักษณะของข้อมูล ซึ่งขึ้นอยู่กับว่าคุณลักษณะใดหรือโหนดภายในใด
- (3) โหนดใบ (Leaf node) คือคำตอบของข้อมูลหรือผลของการจำแนกข้อมูล ส่วนมากมักเรียกว่า คลาส (Class)

การจำแนกข้อมูลเป็นกระบวนการสร้างตัวจัดการข้อมูล เพื่อแสดงให้เห็นความแตกต่างระหว่างคลาสหรือกลุ่มของข้อมูล เพื่อทำนายว่าข้อมูลนี้ควรจัดอยู่ในคลาสใด แบบจำลองที่ใช้การจำแนกข้อมูลออกเป็นกลุ่มตามที่ได้กำหนดไว้จะขึ้นอยู่กับการวิเคราะห์ข้อมูลชุดเรียนรู้ (Training data) ผลลัพธ์ที่ได้จากการเรียนรู้ คือ ตัวจำแนกข้อมูล (Classifier) สามารถแทนได้ในหลายรูปแบบ จากนั้นจะนำค่าที่ได้จากข้อมูลทดสอบมาเปรียบเทียบกับข้อมูลกลุ่มที่เรียนรู้ เพื่อทดสอบความถูกต้องในการจัดกลุ่มข้อมูล โดยจะปรับปรุงตัวจำแนกข้อมูล จนกว่าจะได้ค่าความถูกต้องในระดับที่น่าพอใจ หลังจากนั้นเมื่อมีข้อมูลใหม่เข้ามา จะนำข้อมูลผ่านตัวจำแนกเพื่อทำนายกลุ่มของข้อมูลนี้ได้ ส่วนของต้นไม้การตัดสินใจ ประกอบด้วยหน่วยแรกสุดที่เรียกว่าโหนดราก (Root Node) จากนั้นจะแตกกิ่งออกเป็นโหนดภายใน และจะแยกต่อไปเรื่อยๆ โดยโหนดในระดับสุดท้ายจะเรียกว่าโหนดใบ (Leaf Node) (Berry, 2004)

3.2 ข้อจำกัดของวิธีต้นไม้การตัดสินใจ (Decision Trees)

3.2.1 การแบ่งกลุ่มโดยวิธีต้นไม้การตัดสินใจ (Decision Trees) นิยมใช้กับข้อมูลเชิงกลุ่ม กรณีที่เป็นข้อมูลเชิงปริมาณ เช่น ข้อมูลรายได้ ข้อมูลราคา เป็นต้น จะจัดตัวเลขให้เป็นช่วงหรือกลุ่มก่อนนำข้อมูลไปใช้ในการแบ่งกลุ่ม เพื่อสร้างต้นไม้การตัดสินใจ

3.2.2 การจัดการกับข้อมูลที่ไม่มีทราบค่า อาจส่งผลกระทบต่อความแม่นยำในการทำนายหรือการจัดกลุ่มของต้นไม้การตัดสินใจ เนื่องจากจะสนใจเฉพาะตัวแปรที่ใช้เป็นตัวจำแนกกลุ่ม

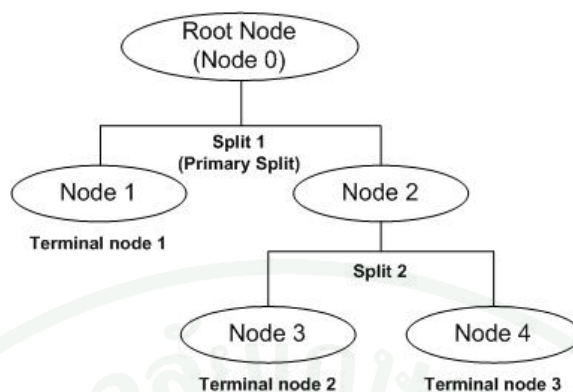
3.2.3 ต้นไม้การตัดสินใจที่มีระดับชั้นมากขึ้นไป จะทำให้ข้อมูลที่ผ่านโหนด มีความจำเพาะกับข้อมูลมากขึ้นไป ทำให้การวิเคราะห์เพื่อทำนายค่าหรือแบ่งกลุ่มมีความผิดพลาดได้

การสร้างต้นไม้การตัดสินใจมีวิธีที่ใช้ในการสร้างแบบจำลองหลายวิธี โดยวิธีที่นิยมและสามารถนำมาประยุกต์ใช้ในการทำนายค่าในงานวิจัยนี้ได้แก่ การสร้างแบบจำลองโดยวิธี Classification and Regression Trees (CART) และวิธี Chi-Square Automatic Iteration Detector (CHAID)

4. แบบจำลอง Classification and Regression Trees (CART)

การสร้างแบบจำลองการทำนาย (Predictive Modeling) ใช้ในการทำนายผลข้อมูล โดยจะเน้นการแบ่งข้อมูลออกเป็นกลุ่มตามคุณสมบัติของข้อมูล ถ้าข้อมูลตัวแปรเป็นค่าไม่ต่อเนื่องหรือเป็นตัวแปรเชิงกลุ่ม จะใช้กระบวนการแบ่งกลุ่มเป็นแบบ Classification Tree แต่ถ้าข้อมูลตัวแปรเป็นค่าต่อเนื่องหรือตัวแปรเชิงปริมาณ จะใช้กระบวนการแบ่งกลุ่มแบบ Regression Tree ทฤษฎีนี้ถูกสร้างขึ้นโดย Breiman, *et al.* (1984) โดยใช้หลักในการสร้างต้นไม้การตัดสินใจแบบทวิภาค (Binary Tree) คือการแตกข้อมูลจาก 1 เป็น 2 เสมอ ค่าทำนายที่ได้มาจากการจัดกลุ่มข้อมูลที่ทำให้ค่าตัวแปรตามมีค่าใกล้เคียงกันมากที่สุด

การสร้างต้นไม้การตัดสินใจด้วยแบบจำลอง CART แสดงในภาพที่ 11 เริ่มจากโหนดราก (Node 0) ซึ่งประกอบด้วยชุดข้อมูลทั้งหมดที่ใช้ในการสร้างแบบจำลอง จากนั้นแยกข้อมูลเป็น 2 กลุ่มย่อย (Node 1 และ Node 2) โดยการพิจารณาแบ่งกลุ่มข้อมูลที่มีลักษณะคล้ายกัน ในการแยกแต่ละครั้งจะคัดเลือกตัวแปรอิสระมาหนึ่งตัว กำหนดค่าของตัวแปรนั้นๆเพื่อใช้เป็นเกณฑ์ในการแยกข้อมูล (Split 1) จากนั้นจะพิจารณาโหนดที่ได้ว่าสามารถแบ่งกลุ่มได้อีกหรือไม่ กรณีที่สามารถแบ่งกลุ่มได้ โหนดนั้นจะเป็นโหนดแม่ (Parent Node) จากภาพ Node 2 แบ่งได้เป็นโหนดลูก (Child Nodes) 2 โหนด คือ Node 3 และ Node 4 ทำการแบ่งเช่นนี้ไปเรื่อยๆ จนได้โหนดสุดท้ายในแต่ละสาย (Nodes 1, 3 และ 4) ซึ่งหมายความว่าลักษณะข้อมูลในแต่ละโหนดมีความใกล้เคียงกันและไม่สามารถแยกคุณลักษณะต่างๆได้อีกแล้ว เรียกโหนดเหล่านี้ว่าโหนดใบ (Leaf Node หรือ Terminal Node)



ภาพที่ 11 โครงสร้างต้นไม้การตัดสินใจ แบบ CART

ที่มา: Breiman *et al.* (1984)

4.1 วิธีการสร้างต้นไม้การตัดสินใจ แบบ CART ประกอบไปด้วย 3 ขั้นตอน (Breiman *et al.*, 1984) คือ

4.1.1 ขั้นตอนการสร้างต้นไม้

กระบวนการสร้างต้นไม้จะเริ่มจากการแยกโหนด จากโหนดรากออกเป็น 2 โหนด คือ โหนดซ้ายและขวา โดย CART จะหาการแบ่งที่ดีที่สุดจากตัวแปรทั้งหมดที่เป็นไปได้ หลักการการกำหนดค่าการแบ่งที่ดีที่สุดนี้ จะใช้การหาจุดแบ่งที่ให้ค่าความต่างของค่าความไม่บริสุทธิ์ (Impurity Function) ระหว่างโหนดแม่กับโหนดลูก 2 โหนดที่มีค่าน้อยที่สุด โดยค่าความเหมาะสมในการแบ่งข้อมูล ($\Delta(s,t)$) หาได้จากสมการที่ (10)

$$\Delta(s,t) = i(t) - (p_L i(t_L) + p_R i(t_R)) \quad (10)$$

เมื่อ	$\Delta(s,t)$	คือ ค่าความต่างของค่า Impurity
	s	คือ ค่าแบ่งของตัวแปร v ที่โหนดแม่ t
	$i(t)$	คือ ค่าความไม่บริสุทธิ์ของโหนด t
	p_L, p_R	คือ สัดส่วนของโหนดลูกซ้ายและขวา
	$i(t_L), i(t_R)$	คือ ค่าความไม่บริสุทธิ์ของโหนดลูกซ้ายและขวา

ในการคำนวณค่าความไม่บริสุทธิ์ (Impurity Function) ของกลุ่มข้อมูล สามารถหาได้จากหลายกฎเกณฑ์ โดยค่าที่นิยมนำมาใช้เป็นเกณฑ์ในการหาค่า Impurity ได้แก่

(1) ค่าดัชนีเอนโทรปี (Entropy Index) หรือค่าสารสนเทศข้อมูล (Information) ขึ้นอยู่กับความน่าจะเป็นของข้อมูลสามารถคำนวณได้จากสมการที่ (11)

$$i(t) = - \sum_{j=1}^k p_j(t) \ln(p_j(t)) \quad (11)$$

เมื่อ $i(t)$ คือ ค่าความไม่บริสุทธิ์ของโหนด t
 $p_j(t)$ คือ ความน่าจะเป็นที่ข้อมูลโหนด t จะเกิดค่าที่คลาส j จากจำนวนคลาสทั้งหมด k กลุ่ม (Classes)

ลักษณะการวัดค่าความไม่บริสุทธิ์ (Impurity) หรือการกระจายของข้อมูลการเรียนรู้ โดยค่าดัชนีเอนโทรปี สามารถบอกประสิทธิภาพของคุณลักษณะข้อมูลที่ใช้จำแนกข้อมูลได้ ค่าเอนโทรปีที่มีค่าน้อยจะบ่งบอกว่า ชุดข้อมูลนั้นมีข้อมูลที่มีคลาสแตกต่างกันน้อยหรือเกือบจะเป็นคลาสเดียวกันหมด แต่ถ้าค่าเอนโทรปีมีค่ามาก แสดงว่ากลุ่มข้อมูลนั้นมีความแตกต่างกันมาก หรือประกอบด้วยข้อมูลหลายคลาส (Mitchell, 1997)

(2) ค่าดัชนีจีนิ (Gini Index) คือค่าสัมประสิทธิ์จีนิที่แสดงในรูปร้อยละของค่าสัมประสิทธิ์จีนิ เป็นการวัดการกระจายของข้อมูลทางสถิติที่นิยมใช้เป็นตัวบ่งชี้ความเหลื่อมล้ำของการกระจายรายได้หรือการกระจายความร่ำรวยของประชากร ซึ่งถูกพัฒนาขึ้นโดย Corrado Gini นักสถิติชาวอิตาลี ค่าสัมประสิทธิ์จีนิเป็นอัตราส่วนที่มีค่าระหว่าง 0 และ 1 ค่าสัมประสิทธิ์จีนิที่ต่ำจะแสดงถึงความเท่าเทียมกันในการกระจายรายได้ หากค่านี้สูงขึ้นจะบ่งชี้ถึงการกระจายรายได้มีการเหลื่อมล้ำกันมาก โดยค่าสัมประสิทธิ์จีนิที่เท่ากับ 0 หมายถึงความเท่าเทียมกันอย่างสมบูรณ์ (ทุกคนมีรายได้เท่ากัน) และ 1 หมายถึงความเหลื่อมล้ำอย่างสมบูรณ์ (มีคนที่มียาได้เพียงคนเดียว ส่วนคนที่เหลือไม่มีรายได้เลย) การคำนวณค่าสัมประสิทธิ์จีนิอยู่บนสมมติฐานว่าไม่มีใครมีรายได้ต่ำกว่าศูนย์ สามารถคำนวณได้จากสมการที่ (12)

$$i(t) = 1 - \sum_{j=1}^k (p_j(t))^2 \quad (12)$$

เมื่อ $i(t)$ คือ ค่าความไม่บริสุทธิ์ของโหนด t
 $p_j(t)$ คือ ความน่าจะเป็นของข้อมูลโหนด t ของการเกิดค่าที่คลาส j ของ
 จำนวนคลาสทั้งหมด k คลาส

(3) กฎสองส่วน (Twoing Rule) ใช้หลักการแยกข้อมูลเป็นสองคลาสที่มีขนาด
 และคุณสมบัติในแต่ละคลาสใกล้เคียงกันมากที่สุด แทนที่การจัดคลาสที่มีคุณสมบัติใกล้เคียงกันที่
 ละคลาส ซึ่งกฎนี้มีความสอดคล้องกับวิธีการหาต้นไม้ การแบ่งที่ดีที่สุดโดยกฎสองส่วนนี้จะ
 กำหนดให้โหนดนั้น มีการแบ่งแค่ 2 คลาสเท่านั้น การคัดเลือกจะเลือกจากการแบ่งที่ให้ค่าดัชนี
 วัดความต่างของความแปรปรวนมากที่สุด ดังสมการที่ (13)

$$\phi(s, t) = \frac{p_L p_R}{4} \left(\sum_{j=1}^k p_j(t_L) p_j(t_R) \right)^2 \quad (13)$$

เมื่อ $\phi(s, t)$ คือ ค่าดัชนีวัดความต่างของความแปรปรวน
 p_L, p_R คือ สัดส่วนของคุณลักษณะที่แบ่งไปยังโหนดลูกซ้ายและขวา
 ในโหนด t ตามลำดับ
 $p_j(t_L), p_j(t_R)$ คือ สัดส่วนของคุณลักษณะต่างๆภายในโหนดลูกซ้ายและขวา
 ที่เกิดค่าที่คลาส j จากจำนวนคลาสทั้งหมด k คลาส
 ตามลำดับ

4.1.2 ขั้นตอนการตัดแต่งกิ่ง (Pruning)

การสร้างต้นไม้ในขั้นตอนก่อนหน้านี้ เป็นการสร้างต้นไม้โดยใช้ข้อมูลชุดเรียนรู้
 เป็นต้นแบบ อาจทำให้ทำนายค่ากับชุดข้อมูลอื่นได้ไม่ดี ในขั้นตอนนี้จึงเป็นการตัดกิ่งที่มีขนาด
 หรือความสำคัญน้อยกว่าที่กำหนดออกไป เพื่อหลีกเลี่ยงการทำนายค่าที่จำเพาะกับชุดข้อมูลมาก
 เกินไป (Overfitting) การตัดแต่งกิ่งใช้ค่า Cost-Complexity (R_α) โดยค่าที่เหมาะสมที่จะนำไปใช้
 คือที่ค่าพารามิเตอร์เชิงซ้อน (α) น้อยที่สุด ความสัมพันธ์ระหว่างตัวแปรจะเป็นเชิงเส้นตรง ดัง
 สมการที่ (14)

$$R_\alpha = R(T) + \alpha |\tilde{T}| \Leftrightarrow \alpha = \frac{R_\alpha - R(T)}{|\tilde{T}|} \quad (14)$$

เมื่อ R_α คือ ค่า Cost-complexity

$R(T)$ คือ ค่าความคลาดเคลื่อนโดยประมาณจากการแทนค่า

$|\tilde{T}|$ คือ ขนาดของต้นไม้ (จำนวนโหนดใบ)

α คือ พารามิเตอร์เชิงซ้อน มีค่าระหว่าง 0 – 1

4.1.3 การเลือกขนาดต้นไม้ที่เหมาะสม

ในการเลือกขนาดต้นไม้จะเลือกขนาดต้นไม้ที่ให้ค่าความคลาดเคลื่อนน้อยที่สุดจากการทำนายค่าข้อมูลอื่นที่ไม่ได้ใช้ในการเรียนรู้ โดยสามารถทำได้ 2 วิธี คือ กรณีที่ข้อมูลมีขนาดใหญ่เพียงพอ ให้ทำการแบ่งชุดข้อมูล ออกเป็นชุดเรียนรู้ ซึ่งใช้ในการสร้างแบบจำลอง ชุดทดสอบ และชุดตรวจสอบ แต่ถ้าขนาดข้อมูลมีขนาดเล็ก จะใช้วิธี n-fold Cross Validation คือแบ่งข้อมูลออกเป็น n ส่วน โดยแต่ละส่วนมีขนาดเท่า ๆ กัน ใช้ 1 ส่วนเป็นชุดทดสอบ ส่วนที่เหลือ n-1 จะใช้เป็นข้อมูลชุดเรียนรู้ ทำการวิเคราะห์ค่าโดยเปลี่ยนข้อมูลชุดทดสอบจนครบ n ครั้ง จากนั้นหาค่าความคลาดเคลื่อนจากการจัดกลุ่มข้อมูลทั้งหมด แล้วเลือกต้นไม้ที่เหมาะสมโดยเลือกจากที่มีค่าความคลาดเคลื่อนในชุดทดสอบต่ำที่สุด

4.2 CART โดยโปรแกรม SPSS Clementine 12.0

การวัดความไม่บริสุทธิ์ (Impurity) โดยโปรแกรม SPSS Clementine 12.0 มี 3 วิธี ขึ้นกับชนิดของข้อมูลค่าเป้าหมาย (Target field) หรือตัวแปรตาม โดยเมื่อค่าเป้าหมายเป็นตัวแปรเชิงกลุ่ม จะสามารถเลือกได้ 2 วิธี คือ Gini และ Twoing (ดังที่ได้กล่าวใน 4.1.1) กรณีที่ค่าเป้าหมายเป็นตัวแปรที่เป็นค่าต่อเนื่อง จะใช้วิธีค่าเบี่ยงเบนกำลังสองน้อยที่สุด (Least-squared deviation: LSD)

วิธี LSD เป็นวิธีที่ใช้ในการวัดค่าความไม่บริสุทธิ์ โดยจะวัดค่า $R(t)$ หรือค่าความคลาดเคลื่อนโดยประมาณ จากการแทนค่าความแปรปรวนภายในโหนดที่คิดตามน้ำหนักของโหนด t

หรือเป็นค่าความแปรปรวนของโหนดนั้น จุดแบ่งโหนดที่กำหนดคือจุดที่มีค่าดัชนีวัดความต่างของความแปรปรวนมากที่สุด ($\Phi(s, t)$) กำหนดดังในสมการที่ (15) และ (16)

$$\Phi(s, t) = R(t) - p_L R(t_L) - p_R R(t_R) \quad (15)$$

เมื่อ $\Phi(s, t)$ คือ ค่าดัชนีวัดความต่างของความแปรปรวน

$R(t)$ คือ ค่าความแปรปรวนโดยประมาณ

p_L คือ สัดส่วนของค่าความแปรปรวนของโหนดลูกทางซ้ายมือ

p_R คือ สัดส่วนของค่าความแปรปรวนของโหนดลูกทางขวามือ

$$R(t) = \frac{1}{N_w(t)} \sum_{i \in t} w_i f_i (y_i - \bar{y}(t))^2 \quad (16)$$

เมื่อ $R(t)$ คือ ค่าความแปรปรวนโดยประมาณ

$N_w(t)$ คือ ค่าความคลาดเคลื่อนโดยประมาณ

$N_w(t)$ คือ จำนวนระเบียบที่คิณน้ำหนักในโหนด t (Weighted Number of Record)

w_i คือ ค่าน้ำหนักเขตข้อมูล (Field) ของระเบียบ i

f_i คือ ค่าความถี่เขตข้อมูล (Frequency Field)

y_i คือ ค่าของเขตข้อมูลเป้าหมาย (Target Field)

$\bar{y}(t)$ คือ ค่าเฉลี่ยของค่าเขตข้อมูลเป้าหมายของโหนด t

5. แบบจำลอง Chi-Square Automatic Iteration Detector (CHAID)

แบบจำลองนี้ถูกเผยแพร่โดย Gordon V. Kass ในปี ค.ศ. 1980 โดยนำมาใช้ในการคาดเดาผล (คล้ายกับการวิเคราะห์การถดถอย) หรือตรวจจับความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ โดยสามารถใช้ได้กับตัวแปรตามได้ทั้งตัวแปรเชิงกลุ่มและตัวแปรเชิงปริมาณ โดยถ้าตัวแปรตามเป็นตัวแปรเชิงกลุ่มจะใช้การวิเคราะห์ Chi-Square ในการหาความสัมพันธ์ระหว่างตัวแปรและทำการแบ่งกลุ่มข้อมูล แต่ถ้าตัวแปรเป้าหมายเป็นตัวแปรเชิงปริมาณ แบบจำลองจะใช้การวิเคราะห์ความแปรปรวน F-test (SPSS Inc., 2007)

หลักการทํางานของวิธี CHAID จะเป็นการแยกกลุ่มข้อมูล หรือสร้างกิ่งของต้นไม้แต่ละโหนดได้เท่ากับหรือมากกว่า 2 กิ่ง การแบ่งกลุ่มข้อมูลจะพิจารณาโดยการวิเคราะห์ทางสถิติเลือกตัวแปรที่มีค่าแตกต่างกันตามระดับนัยสำคัญมากที่สุด พิจารณาทีละตัวแปรจนครบทุกตัวแปรที่สามารถเป็นไปได้ สุดท้ายจะได้โหนดใบหรือโหนดสุดท้ายของต้นไม้ ซึ่งเป็นโหนดที่ไม่สามารถแบ่งกลุ่มต่อไปได้อีก โดยแผนภาพต้นไม้จะแสดงเป็นลำดับขั้น สามารถแสดงได้ทั้งแบบบนลงล่างแบบซ้ายไปขวาและแบบขวาไปซ้าย (SPSS Inc., 2007)

5.1 ขั้นตอนการสร้างต้นไม้ โดยวิธี CHAID

5.1.1 ใช้การรวมกลุ่มตัวแปรย่อยที่มีค่าใกล้เคียงกัน โดยพิจารณาตัวแปรทำนายค่าหรือตัวแปรอิสระทีละคู่ โดยถ้าตัวแปรตามเป็นตัวแปรเชิงกลุ่มจะใช้การทดสอบแบบ Chi-Square แต่ถ้าตัวแปรตามเป็นตัวแปรเชิงปริมาณ จะใช้การทดสอบแบบ F-test จากนั้นจะพิจารณาผลการทดสอบที่ได้ ถ้าให้ผลไม่แตกต่างอย่างมีนัยสำคัญ ที่ระดับนัยสำคัญ (α) ที่ใช้ในการรวมกลุ่มที่กำหนดจะทำการรวมกลุ่มข้อมูลคู่ นั้น และหาคู่ที่จะใช้ในการทดสอบใหม่ แต่ถ้าให้ผลการทดสอบแตกต่างอย่างมีนัยสำคัญ จะคำนวณค่า p-value ที่ถูกปรับค่า เพื่อกำหนดกลุ่มของชุดข้อมูลค่าทำนายต่อไป

5.1.2 เลือกจุดแบ่งในการแยกคุณลักษณะของตัวแปรที่ใช้ โดยเลือกจาก ค่า p-value ที่ถูกปรับค่าที่มีค่าน้อยที่สุดหรือมีค่าความแตกต่างระหว่างชุดข้อมูลมากที่สุด โดยมีค่ามากกว่าระดับนัยสำคัญที่กำหนด ในการแยกกลุ่มจะกำหนดให้จุดแบ่งคุณลักษณะนั้นเป็นลักษณะที่จะได้ในโหนดใบ (Terminal Node)

5.2 ข้อดีและข้อเสียของวิธี CHAID

วิธี CHAID สามารถที่จะสร้างผลลัพธ์ที่ง่ายต่อความเข้าใจ เพราะมีการใช้การแยกข้อมูลแบบหลายทาง (Multi-way Split) อีกทั้งประสิทธิภาพของการทำงานยังเป็นที่น่าพอใจทั้งข้อมูลขนาดเล็กและขนาดใหญ่ ในด้านของวิธีการทํางาน CHAID จะทำการตรวจจับปฏิสัมพันธ์ระหว่างตัวแปรในกลุ่มของข้อมูลในการสร้างความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variables) หลังจากที่ได้ความสัมพันธ์แล้ว CHAID จะทำการแยกข้อมูลออกเป็นกลุ่ม เพื่อใช้ในการอธิบายหรือทำนายค่าต่างๆ ที่จะเกิดขึ้นบนตัวแปรอิสระอื่นๆต่อไป ทั้งนี้มีข้อเสีย คือ ในการสร้าง

ต้น ไม่มีการตัดสินใจอาจมีความจำเพาะกับชุดข้อมูลที่ใช้ในการสร้างแบบจำลอง ในกรณีที่มีการสร้างระดับขั้นต้นไม่หลายชั้น ทำให้ความสามารถในการใช้งานทั่วไปของแบบจำลองลดลง อีกทั้งยังเลือกพิจารณาเฉพาะตัวแปรที่เป็นจุดแบ่งที่สำคัญเท่านั้น ทำให้การทำนายค่าข้อมูลที่อยู่ในกลุ่มที่ไม่ได้เป็นจุดแบ่งมีความแม่นยำในการทำนายที่ลดน้อยลง

5.3 วิธี CHAID ในโปรแกรม SPSS Clementine 12.0

วิธี CHAID สามารถทำการวิเคราะห์ข้อมูลได้กับตัวแปรเชิงปริมาณและเชิงกลุ่ม กรณีที่เป็นตัวแปรเชิงปริมาณ (Scale) โปรแกรมจะทำการแปลงค่าให้เป็นตัวแปรเชิงกลุ่มโดยอัตโนมัติ ทั้งนี้ผู้ใช้สามารถกำหนดค่าต่างๆในส่วนผู้เชี่ยวชาญ (Expert Option) รวมทั้งการเลือกวิธีการทดสอบต่างๆ ได้

วิธีการรวมกลุ่มข้อมูล (Merging Categories for Predictors) โดยวิธี CHAID ทำได้ 2 รูปแบบ คือ

ก. วิธี CHAID การพิจารณาแต่ละจุดแบ่ง โดยทำการรวมกลุ่มเขตข้อมูลตัวทำนายค่าที่ให้ค่าเขตข้อมูลเป้าหมายที่ไม่แตกต่างกันทางสถิติหรือเป็นกลุ่มข้อมูลที่คล้ายกัน แต่ละกลุ่มข้อมูลของเขตข้อมูลตัวทำนายค่า (X) ที่ใช้เป็นจุดแยกจะแสดงเป็นโหนดลูก โดยมีขั้นตอนดังนี้

- 1) ถ้า X มีหนึ่งหรือสองกลุ่ม จะเข้าสู่กระบวนการแบ่งโหนดทันที
- 2) ถ้า X มีมากกว่าสองกลุ่ม จะจับคู่กลุ่มของ X ที่มีความเหมือนกัน ถ้าให้ค่าการวิเคราะห์ทางสถิติไม่แตกต่างกันคือมีค่า p-value มากกว่าค่า α_{merge} จะทำการรวมกลุ่มนั้นเป็นกลุ่มเดียว ถ้าให้ค่า p-value น้อยกว่าหรือเท่ากับ $\alpha_{\text{split-merge}}$ จะทำการแบ่งกลุ่ม 2 กลุ่มนั้น
- 3) ถ้าจำนวนของข้อมูลในการแบ่งกลุ่มนั้นมีจำนวนน้อยกว่าที่กำหนดจะทำการรวมกับข้อมูลกลุ่มที่ให้ค่าใกล้เคียงกันมากที่สุด

ข. Exhaustive CHAID มีหลักการทำงานคล้าย CHAID แต่การรวมกลุ่มข้อมูลจะมีการทดสอบอย่างละเอียดเพื่อหาชุดของแต่ละเขตข้อมูลตัวทำนาย

ระบบสนับสนุนการตัดสินใจ

1. ระบบสารสนเทศและเทคโนโลยีสารสนเทศ

ระบบสารสนเทศ (Information System : IS) คือ ระบบที่รวบรวมองค์ประกอบต่างๆ เช่น คน ข้อมูล กระบวนการ การเชื่อมโยง เครื่องมือหรืออุปกรณ์ โดยมีการรวบรวมข้อมูลผ่านการประมวลผลหรือกระบวนการต่างๆ ได้เป็นสารสนเทศที่สามารถนำไปใช้เพื่อสนับสนุนการแก้ปัญหาและการตัดสินใจทางธุรกิจ โดยมีจะนำเทคโนโลยีทางคอมพิวเตอร์และเทคโนโลยีสื่อสารโทรคมนาคมมาช่วยในการจัดการและประยุกต์ใช้ในกระบวนการรวบรวมข้อมูล จัดหา จัดเก็บ สร้างและเผยแพร่ระบบสารสนเทศต่างๆ ที่มีความถูกต้อง แม่นยำและรวดเร็ว เป็นการใช้ประโยชน์อย่างมีประสิทธิภาพ

ระบบสารสนเทศจึงเป็นเครื่องมือที่ช่วยในการทำงานด้านต่างๆ เช่น ด้านประมวลผลข้อมูลทางธุรกิจที่เกิดขึ้นประจำวัน การนำเสนอข้อมูลด้วยรายงานชนิดต่าง ๆ หรือการนำไปช่วยสนับสนุนการตัดสินใจเพื่อแก้ปัญหาด้านธุรกิจ สามารถแบ่งได้เป็น 7 ชนิด ได้แก่ ระบบการประมวลผลด้วยรายการเปลี่ยนแปลง ระบบสารสนเทศเพื่อการจัดการ ระบบสนับสนุนการตัดสินใจ ระบบผู้เชี่ยวชาญ ระบบสารสนเทศสำหรับผู้บริหาร ระบบสารสนเทศสำนักงาน ระบบบนเว็บ (สกาวิรัตน์, 2550)

2. ระบบสนับสนุนการตัดสินใจ (Decision Support System หรือ DSS)

ระบบสนับสนุนการตัดสินใจเป็นระบบสารสนเทศที่จัดเตรียมข้อมูลไว้สำหรับผู้ใช้งานเพื่อช่วยในการตัดสินใจในการแก้ปัญหาต่างๆ ระบบนี้ได้ถูกออกแบบมาเพื่อสนับสนุนการตัดสินใจที่ไม่มีโครงสร้าง (Unstructured Decisions) หรือแบบกึ่งโครงสร้าง (Semi-Structured Decisions) ซึ่งยากต่อการหาแนวทางรองรับหรือแนวทางการแก้ปัญหาที่เกิดขึ้น โดย DSS นี้จะไม่ตัดสินใจแทนผู้บริหารหรือผู้ใช้ แต่จะเป็นเครื่องมือที่จัดหาและประมวลผลสารสนเทศต่างๆ ที่จำเป็นในการตัดสินใจของผู้บริหาร ผลลัพธ์ที่ได้จะขึ้นกับตัวแปรที่ผู้ใช้งานได้ป้อนเข้าสู่ระบบ แล้วผู้ใช้งานจะตัดสินใจเองภายใต้สถานการณ์ต่างๆ

เครื่องมือของระบบสนับสนุนการตัดสินใจ สามารถนำโปรแกรมต่างๆมาประยุกต์ใช้ในระบบได้ ตัวอย่างเช่น โปรแกรมตารางทำการ (Spreadsheet เช่น Microsoft Excel) โปรแกรมจัดการฐานข้อมูล (เช่น Microsoft Access) เครื่องมือสำหรับสร้างรายงาน (เช่น Crystal Report) และโปรแกรมวิเคราะห์งานทางสถิติ (เช่น SAS , SPSS) (สกวรัตน์, 2550)

3. ระบบจัดการฐานข้อมูล

ระบบจัดการฐานข้อมูล คือ ซอฟต์แวร์สำหรับบริหารและจัดการฐานข้อมูล เป็นเสมือนตัวกลางระหว่างผู้ใช้กับฐานข้อมูล ทำหน้าที่ในการสร้าง เรียกใช้ข้อมูล หรือปรับปรุงฐานข้อมูล ในการทำงานกับฐานข้อมูลจะต้องผ่านระบบจัดการฐานข้อมูล (Database Management System : DBMS) โดยผู้ใช้เป็นผู้ออกคำสั่งกับ DBMS เพื่อให้ไปจัดการตามคำสั่งนั้น ในขณะที่ผู้ใช้ระดับสุดท้ายจะติดต่อผ่านโปรแกรมประยุกต์สำหรับการใช้งาน ซึ่งจะแตกต่างกันไปตามชนิดของฐานข้อมูลนั้นๆ สุรวดีและตังจะ (2549) และ สกวรัตน์ (2550) ได้สรุปพื้นฐานความรู้ด้านฐานข้อมูลไว้ ดังนี้

3.1 ข้อดีของระบบฐานข้อมูล

3.1.1 ช่วยลดความซ้ำซ้อนของข้อมูล เป็นผลสืบเนื่องจากการจัดเก็บข้อมูลทั้งหมดไว้ที่เดียวกัน เป็นการประหยัดพื้นที่ในการจัดเก็บ ช่วยลดความขัดแย้งและความผิดพลาดของข้อมูลจากการแยกเก็บข้อมูลหลายที่ โดยที่ DBMS จะทำหน้าที่ในการจัดการระบบและติดตามผลการทำงานต่างๆ ภายในฐานข้อมูล ไม่ว่าจะเป็นการเพิ่มข้อมูล การปรับปรุงข้อมูลและการลบข้อมูลให้มีความครบถ้วนสมบูรณ์ในทุกแฟ้มข้อมูล

3.1.2 สามารถใช้ข้อมูลร่วมกันระหว่างข้อมูลที่อยู่ต่างแฟ้มข้อมูลกัน เพิ่มความสะดวกในการใช้งาน

3.1.3 สามารถกำหนดให้ข้อมูลมีรูปแบบที่เป็นมาตรฐานเดียวกันได้ ไม่ว่าจะเป็นรูปแบบข้อมูลก่อนการจัดเก็บ หรือรูปแบบข้อมูลที่ดึงขึ้นมาใช้งาน เป็นต้น ทำให้ผู้ใช้งานฐานข้อมูลสามารถเข้าใจและสื่อสารไปในทิศทางเดียวกัน

3.1.4 สามารถกำหนดระบบความปลอดภัยให้กับข้อมูลได้ โดยกำหนดระดับความสามารถในการเรียกใช้ข้อมูลของผู้ใช้แต่ละคนให้แตกต่างกันตามความรับผิดชอบ

3.1.5 สามารถควบคุมความคงสภาพของข้อมูลได้ โดย DBMS จะมีโปรแกรมสำหรับใช้ระบุกฎเกณฑ์ต่างๆ เพื่อประโยชน์ในการควบคุมความผิดพลาดที่อาจเกิดขึ้นจากการป้อนข้อมูลผิดพลาดของผู้ใช้งานเอง

3.1.6 ช่วยให้ข้อมูลเป็นอิสระจากโปรแกรมที่ใช้งานข้อมูลนั้นและสามารถสร้างสมดุลในการใช้ข้อมูลได้ เมื่อผู้ดูแลฐานข้อมูลทราบความต้องการใช้งานข้อมูลของผู้ใช้ว่าต้องการใช้ข้อมูลประเภทใดหรือความถี่ในการใช้งาน สามารถจัดระบบอำนวยความสะดวกไปยังผู้ใช้ได้สะดวกและรวดเร็วเป็นพิเศษ สามารถนำเสนอข้อมูลได้ง่าย และควบคุมการเข้าถึงข้อมูลในเวลาเดียวกัน

3.2 ข้อเสียของระบบฐานข้อมูล

3.2.1 การทำงานของฐานข้อมูลมีความซับซ้อนขึ้น หากต้องการฐานข้อมูลที่มีประสิทธิภาพ จำเป็นต้องพึ่งพาผู้ออกแบบฐานข้อมูลที่มีความชำนาญเท่านั้น

3.2.2 ผู้ใช้งานข้อมูลจำเป็นต้องได้รับการฝึกสอนการใช้งานฐานข้อมูลก่อน จึงจะสามารถทำงานร่วมกับฐานข้อมูลได้ ซึ่งบางองค์กรยังไม่คุ้นเคยกับระบบข้อมูลแบบนี้ ทำให้อาจเกิดการต่อต้านหรืออาจต้องมีการลงทุนเพิ่มในส่วนการฝึกอบรมพนักงานในด้านการใช้งาน

3.2.3 การสูญเสียข้อมูลอาจเกิดขึ้นได้ เนื่องจากข้อมูลถูกเก็บไว้ที่เดียว หากฐานข้อมูลเกิดมีปัญหาอาจทำให้เสียข้อมูลบางส่วนไปได้

3.2.4 ค่าใช้จ่ายสูง เนื่องจากราคาของซอฟต์แวร์ DBMS ค่อนข้างแพงและมีความจำเป็นต้องใช้คอมพิวเตอร์ที่มีประสิทธิภาพค่อนข้างสูง

3.3 องค์ประกอบของระบบฐานข้อมูล

3.3.1 User คือ ผู้ใช้งานฐานข้อมูล โดยผู้ใช้งานไม่จำเป็นต้องมีความรู้เกี่ยวกับฐานข้อมูลก็ได้ แต่จะต้องทราบว่ามีความต้องการข้อมูลอะไรหรือใช้งานอย่างไร ในปัจจุบันคำว่าผู้ใช้ อาจะหมายถึงตัวโปรแกรมหรือระบบงานอื่น ๆ ที่เข้ามาดึงเอาข้อมูลจากฐานข้อมูลไปใช้

3.3.2 Data คือ ข้อมูลที่เก็บอยู่ในฐานข้อมูล เป็นส่วนที่ถูกเรียกนำมาใช้งาน ถูกเรียกมาแก้ไขปรับปรุงเพิ่มเติม โดยปกติมักจะเก็บอยู่ในสื่อเก็บข้อมูลที่เป็นอิเล็กทรอนิกส์ แต่ในมุมมองของการจัดการฐานข้อมูล เราจะมองเห็นในรูปของตารางหรือไฟล์ต่างๆ

3.3.3 DBMS (Database Management System) คือ ซอฟต์แวร์ที่ทำหน้าที่จัดการดูแลฐานข้อมูลให้สามารถใช้งานได้ง่าย มีประสิทธิภาพ และรักษาข้อมูลที่เก็บอยู่ให้เชื่อถือได้เสมอ

3.3.4 Database Server เป็นระบบคอมพิวเตอร์รองรับโปรแกรม DBMS และเป็นระบบคอมพิวเตอร์ที่ใช้เก็บข้อมูลของฐานข้อมูล ดังนั้น จึงมักจะเป็นระบบที่มีประสิทธิภาพสูง เชื่อถือได้ เพราะใช้รองรับการใช้งานจาก User

3.3.5 DBA (Database Administrator) คือ เจ้าหน้าที่ที่มีหน้าที่รับผิดชอบดูแล และจัดการงานต่างๆ ที่เกี่ยวกับฐานข้อมูล เช่น การสร้างฐานข้อมูล การสำรองข้อมูล การเรียกคืนข้อมูล การจัดการด้านความปลอดภัยของข้อมูล ซึ่งจะใช้งาน DBMS เป็นหลัก คอยแก้ไขปัญหาให้กับผู้ใช้งาน กรณีที่เกิดปัญหาขึ้นกับฐานข้อมูล

4. โปรแกรมจัดการด้านฐานข้อมูล Microsoft Access

เป็นโปรแกรมตัวหนึ่งในชุดโปรแกรม Microsoft Office สามารถสร้างฐานข้อมูลเพื่อจัดเก็บข้อมูล รายงานผลข้อมูล เรียกใช้งานข้อมูลตามรูปแบบที่กำหนด มีความสามารถในการจัดการข้อมูลให้ปลอดภัย และใช้งานข้อมูลร่วมกับข้อมูลจากแหล่งอื่นๆ ได้

โปรแกรมจัดการด้านฐานข้อมูล Microsoft Access ประกอบด้วยการออกแบบฐานข้อมูล การสร้างตารางเพื่อจัดเก็บข้อมูลได้หลายรูปแบบ ทั้งข้อความ ตัวเลข รูปภาพ วิดีโอ ฯลฯ ซึ่งจะช่วย

ลดความผิดพลาดในการจัดเก็บข้อมูลที่ซ้ำซ้อน ในกรณีที่มีข้อมูลอยู่แล้ว สามารถปรับปรุงข้อมูลเดิม หรือลบข้อมูลที่มีอยู่ได้อย่างถูกต้องปลอดภัย นอกจากนี้ยังใช้ในการค้นข้อมูลจากฐานข้อมูลได้อย่างมีประสิทธิภาพตามเงื่อนไขที่กำหนด โดยผู้ใช้งานสามารถออกแบบการค้นหาข้อมูลได้ง่ายด้วยเครื่องมือสำเร็จรูปที่โปรแกรมกำหนดไว้ หรืออาจจะใช้ภาษา SQL (Structured Query Language) ซึ่งเป็นภาษาที่ใช้ในการเรียกค้นข้อมูลโดยเฉพาะมาร่วมใช้งานด้วยก็ได้

ในส่วนการใช้งานเพื่อการรายงานผลข้อมูล สามารถทำได้ทั้งในรูปแบบของตาราง กราฟ หรือแผนภูมิชนิดต่างๆ และยังสามารถนำไปประยุกต์ใช้งานกับโปรแกรมอื่นๆ ได้สะดวก และสามารถปรับแต่งให้มีรูปแบบสวยงามตามที่ต้องการได้ โดยสามารถนำข้อมูลเข้าและออกจากฐานข้อมูลระหว่างโปรแกรม Microsoft Access กับโปรแกรมอื่นๆ ได้ ไม่ว่าจะเป็น Excel, Word, PowerPoint หรือจะเป็นโปรแกรมจัดการฐานข้อมูลอื่นๆ เช่น dBase, Foxbase, Paradox, SQL Server ฯลฯ นอกจากนี้ Microsoft Access สามารถสำรองข้อมูล (Backup) ในกรณีที่ระบบเกิดปัญหา สามารถนำข้อมูลสำรองมาใช้งานหรือกู้คืนข้อมูลสำคัญได้ รวมทั้งสามารถจัดการด้านความปลอดภัยในการเข้าถึงข้อมูลได้อีกด้วย (นันท์, 2548)

4.1 องค์ประกอบของโปรแกรมจัดการฐานข้อมูล Microsoft Access

โปรแกรมการจัดการฐานข้อมูลประกอบไปด้วยส่วนประกอบต่างๆ ทำงานร่วมกันภายในโปรแกรม เพื่อสร้างระบบฐานข้อมูล นันท์ (2548) ได้อธิบายแต่ละส่วนของฐานข้อมูลดังนี้

4.1.1 ตาราง (Table) เป็นส่วนประกอบที่ใช้เก็บข้อมูลจริงโดยแต่ละตารางจะเก็บข้อมูลที่เป็นเรื่องเดียวกันเท่านั้น ตารางจึงเป็นส่วนที่มีความสำคัญมากที่สุดเพราะจะถูกนำไปใช้เป็นแหล่งข้อมูลของส่วนประกอบอื่นๆ เช่น คิวรี ฟอรัม และรายงานต่างๆ

4.1.2 คิวรี (Query) เป็นส่วนประกอบที่ช่วยในการสอบถามและประมวลผลข้อมูล โดยแสดงผลลัพธ์ที่ได้ในรูปแบบของ Datasheet ซึ่งอาจเป็นข้อมูลที่ดึงมาทั้งหมดหรือเพียงบางส่วนจาก 1 ตารางหรือหลายๆ ตารางพร้อมกันหรือเป็นผลลัพธ์ที่ได้จากการประมวลผลข้อมูลภายใต้เงื่อนไขที่กำหนดคนนอกจากนี้ยังใช้เป็นเครื่องมือในการแก้ไขเปลี่ยนแปลงข้อมูลในตารางได้ด้วย เช่น การเพิ่ม/ลบระเบียนหรือปรับปรุงข้อมูลในแต่ละเขตข้อมูล และเนื่องจากคิวรีมีคุณสมบัติที่อาจเรียกได้ว่าเป็นตารางเสมือนจึงใช้เป็นแหล่งข้อมูลของคิวรีเอง ฟอรัม และรายงาน ได้เช่นเดียวกับตาราง

4.1.3 ฟอรม์ (Form) เป็นส่วนประกอบช่วยในการใช้งาน มีการนำเครื่องมือต่างๆ เช่น ปุ่มคำสั่ง ปุ่มตัวเลือก ช่องเลือกรายการ เป็นต้น มาประกอบกันเป็นรูปแบบตามความต้องการของผู้ใช้ เพื่อช่วยให้การทำงานร่วมกับข้อมูลบนจอภาพมีความถูกต้อง รัดกุม ปลอดภัยและตรงกับวัตถุประสงค์ในการใช้งานมากที่สุด สามารถประยุกต์ใช้ฟอรม์กับการทำงานในลักษณะต่าง ๆ ที่หลากหลาย ทั้งยังสามารถปรับแต่งรูปแบบให้สวยงามน่าใช้งานได้อีกด้วย

4.1.4 รายงาน (Report) เป็นส่วนประกอบที่นำเสนอข้อมูลให้อยู่ในรูปเอกสารรายงาน โดยผู้ใช้งานสามารถกำหนดรูปแบบให้ตรงกับความต้องการได้หลายวิธี อาจนำส่วนควบคุมต่างๆ มาประกอบกันเป็นรูปแบบที่ต้องการได้เช่นเดียวกับฟอรม์ โดยสามารถแสดงผลได้ทั้งทางจอภาพหรือพิมพ์ลงแผ่นกระดาษก็ได้ โดยลักษณะเฉพาะของรายงานก็คือ ใช้สำหรับรายงานผลเพียงอย่างเดียวเท่านั้น จะใช้เพื่อแก้ไขแหล่งข้อมูลใด ๆ ของรายงานนั้นไม่ได้ ทำให้งานมีความแตกต่างจากฟอรม์

งานวิจัยที่เกี่ยวข้อง

จากการศึกษางานวิจัยด้านการทำนายค่ามีการนำเทคนิคทางคณิตศาสตร์และสถิติมาใช้หลากหลายวิธีขึ้นกับลักษณะของปัจจัยและข้อมูลที่สร้างแบบจำลอง โดยการทำนายค่าที่ให้ผลตัวแปรตามเป็นค่าต่อเนื่อง (Continuous Variables) นิยมใช้แบบจำลองเครือข่ายประสาทเทียม (Artificial Neural Network) เปรียบเทียบกับแบบจำลองวิเคราะห์การถดถอย (Regression Analysis) เป็นการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variables) สร้างเป็นสมการการทำนายค่า ดังงานวิจัยต่อไปนี้

อัมรินทร์ (2548) ศึกษาการทำนายระดับความชอบของผู้บริโภคต่อเบียร์ และนำระดับความชอบของผู้บริโภคมาใช้บ่งชี้คุณลักษณะสำคัญที่มีอิทธิพลต่อความชอบของเบียร์ โดยจำลองความสัมพันธ์ระหว่างคุณลักษณะของเบียร์ในรูปตัวแปรอิสระ 35 ตัวแปรหรืออยู่ในรูปของกลุ่มปัจจัย 10 กลุ่มที่ได้จากการวิเคราะห์ปัจจัย เพื่อลดจำนวนตัวแปรอิสระกับคะแนนความชอบของผู้บริโภคที่ได้จากการทดสอบทางประสาทสัมผัส ศึกษาเปรียบเทียบระหว่างการใช้แบบจำลองการถดถอยแบบโพลิโนเมียลรูปแบบต่างๆ แบบจำลองเครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับ และแบบจำลองเครือข่ายประสาทเทียมแบบเรเดียลเบสฟังก์ชัน จากข้อมูลทั้งหมด 111 ชุด

อัมรินทร์ (2548) แบ่งข้อมูลออกเป็น 3 ชุด ได้แก่ ชุดเรียนรู้ (Training sets) จำนวน 69 ชุด ชุดทดสอบ (Testing sets) 22 ชุด และชุดทวนสอบ (Validating sets) จำนวน 20 ชุด จากนั้นเปรียบเทียบแบบจำลองต่างๆ โดยการคำนวณความคลาดเคลื่อนในรูปของ RMSE และ MAE ผลการวิจัยพบว่า การลดจำนวนตัวแปรอิสระโดยวิเคราะห์ปัจจัยจะช่วยปรับปรุงความสามารถในการทำนายความชอบของผู้บริโภคในแบบจำลองการถดถอย แต่ไม่มีส่วนช่วยในแบบจำลองเครือข่ายประสาทเทียม โดยแบบจำลองเครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับสามารถทำนายความชอบผู้บริโภคได้ดีที่สุดใกล้เคียงกับค่าจริงและมีความแม่นยำมากกว่าแบบจำลองอื่นๆ รวมทั้งมีความสามารถในการใช้งานทั่วไปดีกว่า โดยมีโครงสร้างที่ดีที่สุดประกอบด้วยหน่วยข้อมูลรับเข้า ซึ่งแสดงถึงข้อมูลตัวแปรคุณลักษณะทางเคมี 35 หน่วย หน่วยซ่อน 4 หน่วย และหน่วยข้อมูลผลลัพธ์ ที่แสดงถึงระดับความชอบ 1 หน่วย ใช้อัตราการเรียนรู้ที่ 0.5 โมเมนตัม 0.5 ใช้กฎการเรียนรู้แบบ Delta Rule และใช้ฟังก์ชันกระตุ้นแบบซิกมอยด์ (Sigmoid Transfer Functions)

สิริยาภรณ์ (2552) ได้ศึกษาและเปรียบเทียบแบบจำลองเพื่อทำนายคุณภาพของผลิตภัณฑ์กาแฟแก้วบดก่อนการทดสอบโดยผู้เชี่ยวชาญในห้องปฏิบัติการ โดยใช้แบบจำลองการถดถอยแบบพหุ การถดถอยแบบโพลิโนเมียล แบบจำลองเครือข่ายประสาทเทียม และการทำเทคนิคพื้นผิวตอบสนองมาประยุกต์ใช้สร้างแบบจำลองเครือข่ายประสาทเทียม โดยมีข้อมูลตัวแปรอิสระทั้งหมด 21 ตัวแปร ประกอบด้วย ระยะเวลาในการเก็บรักษาเมล็ดกาแฟก่อนนำมาคั่ว สี ความมันวาว กลิ่น และขนาดของเมล็ดกาแฟ แหล่งเพาะปลูกของกาแฟโรบัสต้าจำนวน 7 แหล่ง และแหล่งเพาะปลูกของกาแฟอาราบิก้าจำนวน 9 แหล่ง และข้อมูลตัวแปรตาม 5 ตัวแปร ได้แก่ คะแนนจากการทดสอบคุณภาพเชิงประสาทสัมผัสของสี กลิ่น ความขม รสชาติ และความชอบโดยรวมของน้ำกาแฟ สิริยาภรณ์ (2552) แบ่งข้อมูลออกเป็น 3 ชุด ได้แก่ ชุดเรียนรู้ (Training sets) ร้อยละ 70 ของข้อมูลทั้งหมด ชุดทดสอบ (Testing sets) ร้อยละ 20 ของข้อมูลทั้งหมด และชุดทวนสอบ (Validating sets) ร้อยละ 10 ของข้อมูลทั้งหมด จากนั้นเปรียบเทียบแบบจำลองต่างๆ โดยการคำนวณความคลาดเคลื่อนในรูปของ RMSE และ MAE ผลจากการศึกษาพบว่าแบบจำลองเครือข่ายประสาทเทียมมีความสามารถในการทำนายค่าคะแนนทางประสาทสัมผัสของกาแฟได้ถูกต้องใกล้เคียงกับค่าจริงมากกว่าแบบจำลองการถดถอย แต่ทั้งนี้ค่าคลาดเคลื่อนที่ได้มีค่าไม่แตกต่างกันมากนัก

ธัญนันท์ (2549) ได้ศึกษาการเปลี่ยนแปลงทางคุณภาพและการสูญเสียหลังการเก็บเกี่ยวของส้มสายน้ำผึ้งหลังการขนส่งโดยรถบรรทุกจากโรงคัดบรรจุจังหวัดเชียงใหม่ไปยังตลาดกลางค้าส่งกรุงเทพมหานคร โดยทำการวัดค่าการเปลี่ยนแปลงของปริมาณวิตามินซี ปริมาณกรดที่วิเคราะห์

ได้ ค่าความเป็นกรดค่า การเปลี่ยนแปลงของปริมาณของแข็งที่ละลายน้ำได้ อัตราส่วนระหว่างปริมาณของแข็งที่ละลายน้ำได้ต่อปริมาณกรดที่วิเคราะห์ได้ ร้อยละการสูญเสียน้ำหนักสด ร้อยละความเสียหายทางกล ร้อยละการเน่าเสีย และการเปลี่ยนแปลงของสีผิวเปลือกส้ม เพื่อเป็นดัชนีบ่งชี้การเปลี่ยนแปลงทางด้านคุณภาพ ใช้การวางแผนการทดลองแบบ 32 Factorial และทำการทดลองทั้งหมด 3 ชั่วโมง (เกี่ยวข้องของการขนส่ง) โดยมีอุณหภูมิและความชื้นสัมพัทธ์เป็นปัจจัยที่ไม่สามารถควบคุมได้ รัชุนันท์ (2549) ใช้แบบจำลองเครือข่ายประสาทเทียม (Artificial Neural Network) ในการทำนายคุณภาพและการสูญเสียของส้มสายน้ำผึ้งหลังการขนส่ง และนำมาเปรียบเทียบกับการทำนายโดยวิธีแบบจำลองการถดถอยเชิงพหุ (Multiple Linear Regression) จากการวิเคราะห์ผลเปรียบเทียบจากชุดข้อมูลทั้งหมด 27 ชุด แบ่งเป็นชุดเรียนรู้ (Training set) จำนวน 22 ชุด และชุดทดสอบ (Testing set) จำนวน 5 ชุด เพื่อทำนายการเปลี่ยนแปลงทางด้านคุณภาพของส้มสายน้ำผึ้ง พบว่าการทำนายคุณภาพของส้มสายน้ำผึ้งหลังการขนส่งทางรถบรรทุก โดยแบบจำลองเครือข่ายประสาทเทียมให้ค่าความผิดพลาดในการทำนายที่น้อยกว่า และมีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) ที่เพิ่มขึ้นมากกว่าการทำนายโดยแบบจำลองการถดถอยเชิงพหุ (Multiple Linear Regression)

Rai *et al* (2005) ได้สร้างแบบจำลองการทำนายค่าความหนืดของน้ำผลไม้ โดยแบบจำลองเครือข่ายประสาทเทียม ภายได้ปัจจัยเรื่องความเข้มข้นและอุณหภูมิ โดยศึกษาในผลไม้ 5 ชนิด ได้แก่ ส้ม ลูกพีช ลูกแพร์ Malus Floribunda และ แบลคเคอเรนด (Black Current) ข้อมูลค่าความหนืดที่ศึกษาอยู่ในช่วง 1.53-3300 mPas อันเป็นผลจากการใช้ความเข้มข้นในช่วง 5 - 70 องศาบริกซ์ และ อุณหภูมิในช่วง 30.7 - 71.7 องศาเซลเซียส ผู้วิจัยทำการพัฒนาแบบจำลองเครือข่ายประสาทเทียมและเลือกรูปแบบที่ดีที่สุดในการสร้างแบบจำลอง โดยกำหนดรูปแบบการเชื่อมโยงเป็นแบบส่งสัญญาณไปข้างหน้าอย่างเดียว (Feed Forward) ชี้นำเข้ามีตัวแปร 2 ตัว คือ ความเข้มข้นและอุณหภูมิ ชั้นผลลัพธ์ คือ ค่าความหนืด ฟังก์ชันกระตุ้นแบบ Hyperbolic Tangent Function เนื่องจากฟังก์ชันดังกล่าวสามารถแยกความแปรปรวนของตัวแปรนำเข้าแต่ละตัวได้ดี และสามารถใช้งานได้ดีกับข้อมูลที่มีความซับซ้อน ผลการศึกษา พบว่าอัตราการเรียนรู้ที่เหมาะสมเท่ากับ 0.2 แบบจำลองเป็นแบบมีผู้ฝึกสอน (Supervise Learning) โดยเลือกใช้การเรียนรู้แบบส่งถ่ายข้อมูลย้อนกลับ วัดผลการทำนายโดยค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์ (MAE) เท่ากับ 3.78 mPas ข้อมูลที่นำมาใช้ศึกษามีทั้งหมด 322 ข้อมูล แบ่งข้อมูลเป็น 2 ชุด คือชุดเรียนรู้และชุดทดสอบ (Training/Testing Set) จำนวน 270 ข้อมูล และชุดทดสอบ (Validation Set) มีจำนวน 52 ข้อมูล เมื่อพิจารณาจากค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์เฉลี่ย และจำนวนรอบการเรียนรู้ พบว่าแบบจำลองที่เหมาะสมมีโครงสร้าง คือ ประกอบด้วยชั้นซ่อน 2 ชั้น และมีจำนวนหน่วยในชั้นซ่อนเท่า

กับ 2 หน่วย เนื่องจากให้ค่าความคลาดเคลื่อนเฉลี่ยสมบูรณ์ และจำนวนรอบการเรียนรู้ต่ำที่สุด พบว่าแบบจำลองดังกล่าวสามารถทำนายข้อมูลได้อย่างแม่นยำและมีประสิทธิภาพ

Uno *et al.* (2005) ได้ศึกษาการทำนายผลผลิตของข้าวโพดในประเทศแคนาดา เปรียบเทียบระหว่างแบบจำลองเครือข่ายประสาทเทียม และแบบจำลองการถดถอย พิจารณาปัจจัยการผลิตข้าวโพด ได้แก่ ระดับปริมาณการให้ปุ๋ยไนโตรเจน 3 ระดับ (60, 120, 250 กิโลกรัมต่อเฮกเตอร์) และเทคนิคการควบคุมวัชพืช 3 เทคนิค ได้แก่ การควบคุมวัชพืชประเภทหญ้า การควบคุมวัชพืชประเภทใบกว้าง และการควบคุมวัชพืชทุกประเภท เปรียบเทียบกับการถดถอย (ไม่มีการควบคุมวัชพืช) โดยเก็บรวบรวมข้อมูล 3 ช่วงเวลา ได้แก่ ระยะเริ่มการเจริญเติบโต (30 วันหลังจากเริ่มปลูก) ระยะการออกพุ่ม (66 วันหลังจากเริ่มปลูก) และระยะโตเต็มที่ (86 วันหลังจากเริ่มปลูก) แบบจำลองเครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation Learning Algorithm) โครงสร้างประกอบด้วยหน่วยนำเข้า 17 หน่วย วิเคราะห์ปัจจัยโดยวิธีองค์ประกอบหลัก (Principal Component Analysis : PCA) คือจัดให้ปัจจัยที่มีความสัมพันธ์ต่อกันรวมอยู่ในกลุ่มเดียวกัน จำนวนหน่วยชั้นซ่อน 1 ชั้น และหน่วยชั้นผลลัพธ์ คือ ปริมาณผลผลิตข้าวโพด จากข้อมูลเบื้องต้น 192 ชุด แบ่งเป็นข้อมูลชุดเรียนรู้ 72 ชุด ข้อมูลชุดทดสอบ 72 ชุด และชุดทวนสอบ 48 ชุด ใช้การประเมินแบบจำลองด้วยวิธี 10-fold Cross-Validation โดยกำหนดจำนวนปัจจัยการผลิตทั้งหมด 71 ปัจจัยและตัวแปรเชิงกลุ่มที่ได้จากการวิเคราะห์ปัจจัย โดยวิธีวิเคราะห์องค์ประกอบหลัก (PCA) 5 กลุ่ม โดยทดสอบประสิทธิภาพการทำนายด้วยค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ค่าความคลาดเคลื่อนเฉลี่ย (Root Mean Square Error, RMSE) และค่าความแตกต่างเฉลี่ย (Average Difference, AVDIF) ผลการวิจัยพบว่าแบบจำลองที่ใช้ข้อมูลจากการวิเคราะห์ปัจจัยด้วยวิธี PCA ให้ผลการทำนายที่มีความแม่นยำต่ำกว่า โดยผู้วิจัยได้วิเคราะห์ว่าเกิดจากข้อมูลปัจจัยนำเข้าไม่มีความสัมพันธ์กันเพียงพอ พบว่าการทำนายด้วยแบบจำลองเครือข่ายประสาทเทียม และแบบจำลองการถดถอยให้ผลการทำนายใกล้เคียงกัน

Kaul *et al.* (2005) ได้ศึกษาการทำนายปริมาณผลผลิตข้าวโพดและถั่วเหลือง โดยใช้แบบจำลองเครือข่ายประสาทเทียมและแบบจำลองวิเคราะห์การถดถอย โดยใช้ข้อมูลตั้งแต่ปี 1978-1998 ได้จากแหล่งเพาะปลูก 5 แหล่ง ชนิดดินที่แตกต่างกัน 9 ชนิด กำหนดปัจจัยที่ใช้ในการทำนายค่าปริมาณผลผลิตข้าวโพดและถั่วเหลือง คือ อัตราการใช้ดินของพืช (Soil Rating for Plant Growth ; SRPG) ปริมาณน้ำฝนเฉลี่ยรายเดือน (ช่วงเดือนเมษายน - สิงหาคม) และปริมาณน้ำฝนเฉลี่ยรายสัปดาห์ (ช่วงเดือนมิถุนายน - สิงหาคม) วิเคราะห์หาความสัมพันธ์กับปริมาณผลผลิตที่ 3 ระดับ

ได้แก่ ระดับฟาร์ม ระดับเมือง และระดับรัฐ การพัฒนาแบบจำลองเครือข่ายประสาทเทียมแบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation Learning Algorithm) โครงสร้างประกอบด้วยหน่วยนำเข้า (Input Neuron) ได้แก่ ปริมาณน้ำฝนและอัตราการใช้น้ำของพืช หน่วยผลลัพธ์ (Output Neuron) คือ ปริมาณผลผลิต รวบรวมข้อมูลส่วนข้าวโพดมีทั้งหมด 108 ชุด แบ่งเป็นชุดเรียนรู้ 67 ชุด และชุดทดสอบ 41 ชุด และส่วนถั่วเหลืองมีทั้งหมด 109 ชุด แบ่งเป็นชุดเรียนรู้ 60 ชุด และชุดทดสอบ 49 ชุด ทำการเปรียบเทียบแบบจำลองเครือข่ายประสาทเทียมและแบบจำลองการถดถอย โดยการประเมินความสามารถแบบจำลอง โดยการหาค่าความคลาดเคลื่อนจากค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (RMSE) และสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient ; R^2) ผลการวิจัย พบว่า การทำนายค่าผลผลิตทั้งข้าวโพดและถั่วเหลืองโดยใช้แบบจำลองเครือข่ายประสาทเทียมจากปริมาณน้ำฝนรายสัปดาห์ที่ระดับเมืองให้ผลการทำนายค่าใกล้เคียงกับค่าจริงมากที่สุด

สำหรับเทคนิคการแบ่งกลุ่มข้อมูล โดยวิธีการสร้างต้นไม้การตัดสินใจ (Decision Trees) นั้น มีการนำเทคนิคดังกล่าวมาประยุกต์ใช้ในการทำนายค่าหลายด้าน ทั้งในด้านสังคมศาสตร์(เช่น Scheffer, 2002) ด้านวิทยาศาสตร์การแพทย์ (เช่น Petitti, 2000) และแพร่หลายในด้านสิ่งแวดล้อม (เช่น Vayssieres *et al.*, 2000 และ Yang *et al.*, 2003) เนื่องจากเทคนิคนี้มีข้อดี เมื่อเปรียบเทียบกับแบบจำลองการถดถอย กล่าวคือ ไม่มีข้อจำกัดในด้านข้อสมมติการกระจายตัวของข้อมูลตัวแปรตาม (Dependent Variables) ไม่มีปัญหาเรื่องความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) แสดงผลด้านความสัมพันธ์ของตัวแปรอิสระที่มีต่อค่าทำนายหรือตัวแปรตามได้ชัดเจน และสามารถใช้กับตัวแปรได้ทุกประเภท ทั้งแบบมาตรนามบัญญัติ (Nominal Scale) (เช่น ชนิดของดิน) แบบมาตรอันดับ (Ordinal Scale) (เช่น การจัดอันดับ ต่ำ ปานกลาง สูง) และแบบมาตรอันตรภาค (Interval Scale) (เช่น ข้อมูลด้าน biomass) (Thuraisingham, 1999; Vayssieres *et al.*, 2000) งานวิจัยที่ใช้วิธีการสร้างต้นไม้การตัดสินใจในการทำนายค่ามีดังนี้

Mahjoobi and Etemad-Shahidi (2008) ได้นำเสนอการประยุกต์ใช้ Classification and Regression Trees (CART) ในการทำนายค่าความสูงของคลื่นในเขตนํ้าลึกใน Lake Michigan โดยใช้ข้อมูลความเร็วลมและทิศทางลม 5 ปี (มีนาคม ปีค.ศ. 2000 ถึง ธันวาคม ปีค.ศ. 2004) และเปรียบเทียบผลระหว่างแบบจำลองการสร้างต้นไม้ตัดสินใจ CART กับแบบจำลองเครือข่ายประสาทเทียม โดยใช้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (Root mean square error, RMSE) ค่าความลำเอียง (Bias) และค่าดัชนีการกระจาย (Scatter Index; SI) โครงสร้างเครือข่ายประสาทเทียมที่เหมาะสมเป็นแบบ 3 ชั้น แบบส่งสัญญาณไปข้างหน้าอย่างเดียว (Feed forward) ประกอบด้วย ชั้นนำเข้า 2

หน่วย ชั้นซ้อน 5 หน่วย และชั้นผลลัพธ์ 1 หน่วย และเป็นเครือข่ายแบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation Network) จากผลการวิจัยให้ผลความถูกต้องในการทำนายไม่แตกต่างกัน ค่าความคลาดเคลื่อนอยู่ในช่วงที่ยอมรับได้ ซึ่งผู้วิจัยได้เสนอถึงข้อดีของวิธีการสร้างต้นไม้การตัดสินใจว่าวิธีนี้ใช้เวลาในการวิเคราะห์ไม่นาน อีกทั้งยังแสดงผลเป็นแผนภาพลำดับการแบ่งกลุ่มเพื่อทำนายค่า ทำให้ง่ายต่อการทำความเข้าใจความสัมพันธ์ของปัจจัยต่างๆ กว่าการใช้เครือข่ายประสาทเทียม ซึ่งเปรียบเสมือนกล่องดำ (Black Box) สำหรับการทำนายค่า

Tso and Yau (2007) ได้เปรียบเทียบเทคนิคการทำนายค่าด้วยแบบจำลองการวิเคราะห์การถดถอย (Regression Analysis) แบบจำลองเครือข่ายประสาทเทียม และการสร้างต้นไม้การตัดสินใจ (Decision Trees) โดยศึกษาการใช้พลังงานไฟฟ้าในฮ่องกงช่วงฤดูร้อน ในปีค.ศ. 1999 -2000 ข้อมูลที่ใช้ในการทำนายประกอบด้วยข้อมูลปริมาณการใช้ไฟฟ้ารายเดือนและข้อมูลทั่วไปที่รวบรวมจากแบบสอบถาม เช่น จำนวนอุปกรณ์ไฟฟ้าที่ใช้ ข้อมูลพื้นฐานของสมาชิกภายในครัวเรือน เช่น อายุ เงินเดือน เป็นต้น ในงานวิจัยนี้มีชุดข้อมูลครัวเรือนที่สำรวจจำนวน 1,166 ชุด วิเคราะห์โดยสร้างแบบจำลองการถดถอย แบบ Stepwise ที่ระดับนัยสำคัญ 0.05 แบบจำลองเครือข่ายประสาทเทียมเป็นแบบส่งสัญญาณไปข้างหน้าอย่างเดียว (Feed forward) แบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation Network) และใช้ฟังก์ชันกระตุ้นแบบซิกมอยด์ (Sigmoid Transfer Functions) และแบบจำลองการสร้างต้นไม้การตัดสินใจ โดยใช้ค่า F-test เปรียบเทียบผลการทำนายโดยใช้รากที่สองของค่าคลาดเคลื่อนเฉลี่ยกำลังสอง (Root of Mean Squared Error; RMSE) ผลการวิจัยพบว่าวิธีสร้างต้นไม้การตัดสินใจให้ค่าคลาดเคลื่อนในการทำนายน้อยที่สุด เมื่อพิจารณาปัจจัยที่สำคัญที่ใช้ในการแบ่งระดับต้นไม้ พบว่ามีปัจจัยที่สำคัญ 3 อย่าง ได้แก่ พื้นที่ที่พัก จำนวนสมาชิกที่อาศัย และจำนวนเครื่องปรับอากาศ ซึ่งทั้ง 3 ปัจจัยมีระดับความสำคัญสูงเช่นเดียวกับในแบบจำลองการถดถอยและวิธีเครือข่ายประสาทเทียม

จากงานวิจัยที่เกี่ยวข้องทั้งหมด ผู้วิจัยจึงได้กำหนดวิธีทางคณิตศาสตร์และสถิติเพื่อใช้ในการสร้างแบบจำลองการทำนายค่าความหวานวัตถุคิบสับปะรดในกระบวนการผลิตสับปะรดกระป๋องได้ 3 วิธี ได้แก่ วิธีวิเคราะห์การถดถอย (Regression Analysis) วิธีเครือข่ายประสาทเทียม (Artificial Neural Network, ANN) และวิธีต้นไม้การตัดสินใจ (Decision Trees) และนำผลที่ได้จากการพัฒนาแบบจำลองนี้ไปสร้างระบบสนับสนุนการตัดสินใจอย่างง่ายบน Microsoft Access 2007 เพื่อประโยชน์ในการนำไปใช้งานในสายการผลิตของโรงงานกรณีศึกษา

อุปกรณ์และวิธีการ

อุปกรณ์

- คอมพิวเตอร์ระบบปฏิบัติการ Window XP
- โปรแกรมวิเคราะห์ทางสถิติ MINITAB 15.0
- โปรแกรม SPSS Clementine 12.0
- โปรแกรมจัดการฐานข้อมูล Microsoft Access 2007

วิธีการ

1. การเก็บรวบรวมและเตรียมข้อมูล

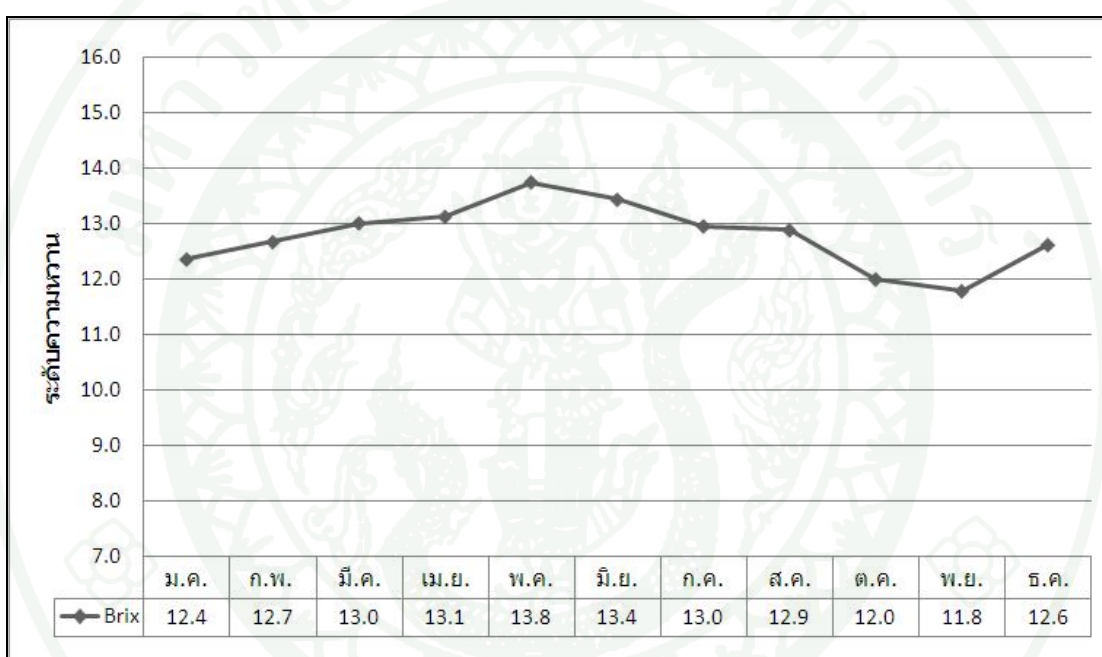
1.1 การเก็บรวบรวมข้อมูลและกำหนดปัจจัยที่เกี่ยวข้อง

ข้อมูลที่ใช้ในการศึกษา เป็นข้อมูลทุติยภูมิ ซึ่งรวบรวมจากโรงงานของบริษัทมาลีสามพราน จำกัด (มหาชน) ซึ่งเป็นบริษัทผู้ผลิตผลไม้บรรจุกระป๋องตามฤดูกาลแห่งหนึ่งในประเทศไทย โดยนำข้อมูลในส่วนการตรวจรับวัตถุดิบและการสุ่มตรวจคุณลักษณะของวัตถุดิบ ในช่วงการผลิตในเดือนมกราคม พ.ศ. 2550 จนถึงเดือนธันวาคม พ.ศ. 2551 ครอบคลุมระยะเวลา 2 ปี ประกอบด้วยข้อมูลตัวแปรอิสระ 4 ตัวแปร ได้แก่ เดือนที่รับวัตถุดิบ (Month) เกรดสีวัตถุดิบ (Color) อำเภอที่ปลูก (Amphoe) และผู้รวบรวมวัตถุดิบ (Quota หรือ Supplier) และข้อมูลตัวแปรตาม 1 ตัวแปร คือ ค่าความหวานของวัตถุดิบ มีหน่วยเป็นองศาบริกซ์ ($^{\circ}\text{Brix}$) ซึ่งเป็นค่าที่ได้จากห้องปฏิบัติการในการวิเคราะห์คุณลักษณะของวัตถุดิบก่อนเข้าสู่กระบวนการผลิต รายละเอียดของตัวแปรแต่ละตัวที่นำมาใช้ในการวิเคราะห์เป็นดังนี้

1.1.1 ระดับความหวานของวัตถุดิบ ($^{\circ}\text{Brix}$) เป็นข้อมูลคุณลักษณะของวัตถุดิบที่ได้จากห้องปฏิบัติการในแต่ละวันที่มีการผลิตสับปะรดบรรจุกระป๋อง เมื่อโรงงานได้รับสับปะรดเข้ามาแล้ว ก่อนเข้าสู่ส่วนการผลิตพนักงานจะทำการสุ่มวัดระดับความหวานของวัตถุดิบสับปะรดที่รับ

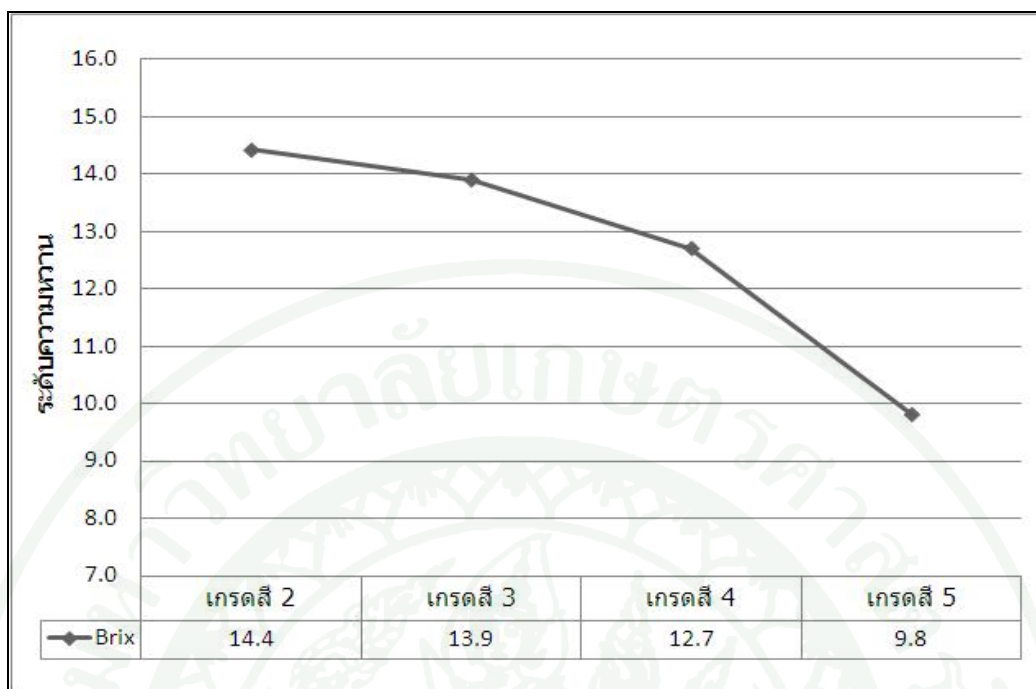
เข้ามา และใช้ข้อมูลนี้ในการปรับคุณลักษณะของสารบรรจุ (Packing Media) ที่ได้เตรียมไว้ให้เหมาะสม เพื่อให้ได้ระดับความหวานสุดท้ายของผลิตภัณฑ์ตรงตามความต้องการ

1.1.2 วันที่รับวัตถุดิบ พิจารณาตัวแปรเดือนที่รับวัตถุดิบ (Month) แสดงถึงฤดูกาล เพาะปลูก และสภาวะในการเพาะปลูก (ได้แก่ อุณหภูมิ ความชื้นในอากาศ ปริมาณน้ำฝน เป็นต้น) ปัจจัยในด้านฤดูกาลนี้มีผลต่อระดับความหวานของวัตถุดิบ ทำให้ระดับความหวานของวัตถุดิบในแต่ละเดือนที่รับวัตถุดิบมีค่าแตกต่างกัน ดังแสดงในภาพที่ 12



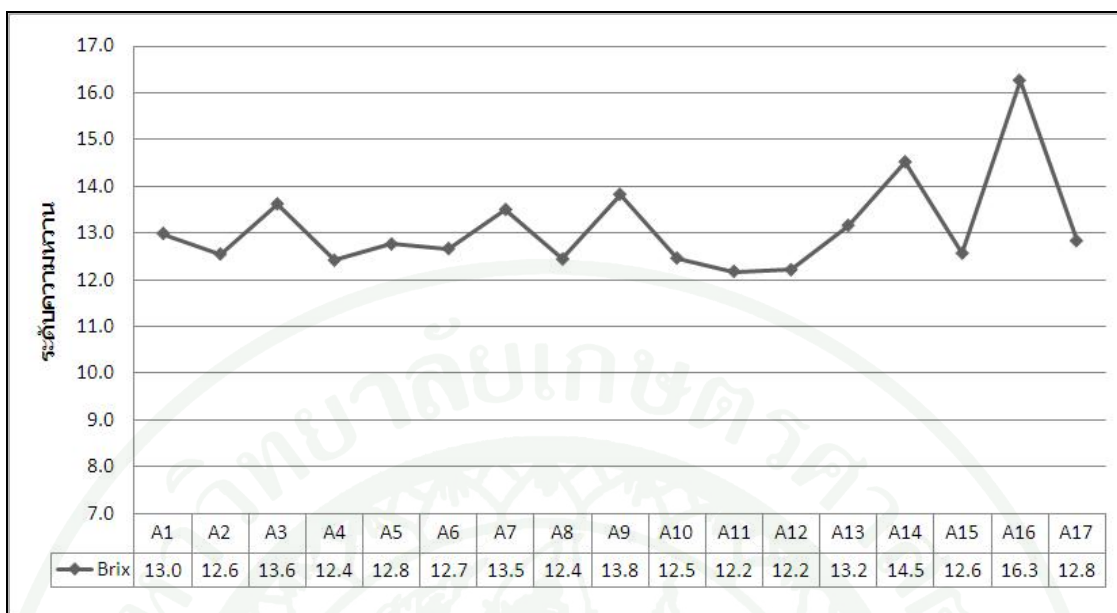
ภาพที่ 12 ค่าเฉลี่ยระดับความหวานวัตถุดิบสับปะรดในแต่ละเดือนที่รับวัตถุดิบ

1.1.3 เกรดสีวัตถุดิบ (Color) แสดงถึงระดับความสุกของเนื้อสับปะรด โดยโรงงาน ภูมิศึกษาได้แบ่งเกรดสีของวัตถุดิบออกเป็น 5 ระดับ คือ ระดับ 1-5 โดยระดับ 1 หมายถึง สีเหลืองจัด ขณะที่ระดับ 5 หมายถึง สีเหลืองซีดเกือบขาว ซึ่งสับปะรดแต่ละลูกจะมีสีของเนื้อสับปะรดมากกว่า 1 ระดับสี ส่วนที่มีสีเหลืองจัด มักจะมีความหวานมากและค่าความเป็นกรดต่ำ แสดงค่าเฉลี่ยของระดับความหวานของวัตถุดิบในแต่ละเกรดสีวัตถุดิบในภาพที่ 13 แต่ทั้งนี้ มีวัตถุดิบจากบางแหล่งปลูกที่มีลักษณะพิเศษแตกต่างออกไป เช่น มีสีเหลืองค่อนข้างมาก แต่มีระดับความหวานต่ำกว่าวัตถุดิบจากแหล่งอื่น เป็นต้น

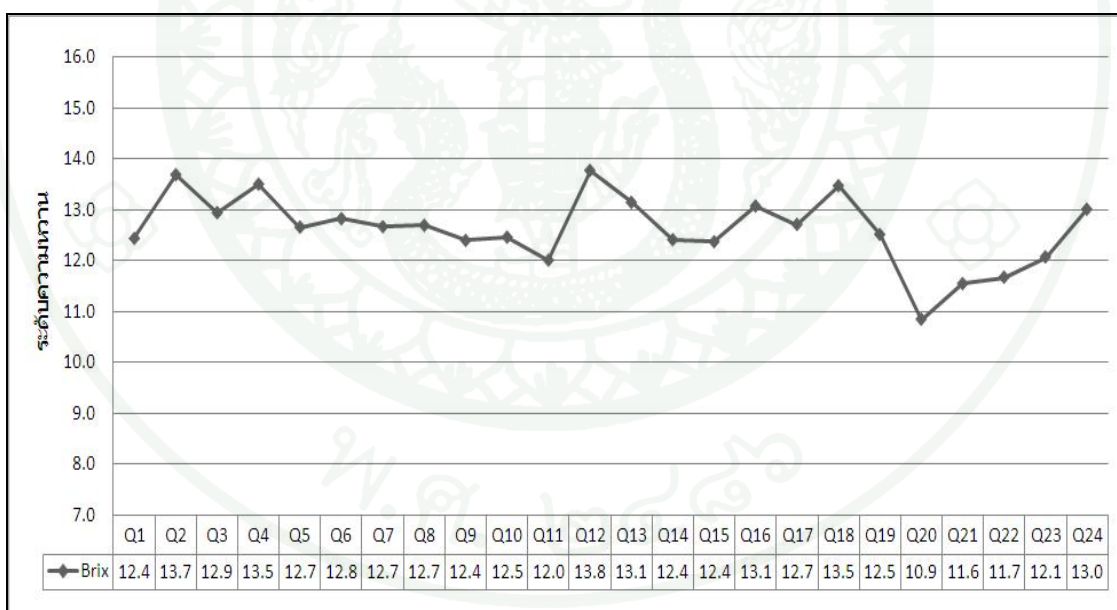


ภาพที่ 13 ค่าเฉลี่ยระดับความหวานวัตถุดิบสับปะรดในแต่ละเกรดสีวัตถุดิบ

1.1.4 แหล่งของวัตถุดิบ ประกอบด้วยอำเภอที่ปลูก (Amphoe) แสดงถึงสถานะทางภูมิศาสตร์ ภูมิอากาศ ในด้านการเพาะปลูก ในที่มีทั้งหมด 17 อำเภอ ตั้งอยู่ในพื้นที่จังหวัดต่าง ๆ รวม 11 จังหวัด แสดงค่าเฉลี่ยของระดับความหวานวัตถุดิบในแต่ละอำเภอที่ปลูกในภาพที่ 14 ซึ่งในกระบวนการจัดซื้อวัตถุดิบของทางโรงงานเป็นระบบทำสัญญาล่วงหน้า โดยผ่านผู้รวบรวมวัตถุดิบ ซึ่งผู้รวบรวมวัตถุดิบ อาจเป็นเกษตรกรที่ทำการปลูกสับปะรดเพื่อจัดส่งให้กับโรงงานหรือเป็นพ่อค้าคนกลางที่ทำหน้าที่รวบรวมวัตถุดิบจากเกษตรกรเฉพาะรายในพื้นที่หนึ่งๆ หรือหลายพื้นที่ ดังนั้นจึงได้กำหนดให้ผู้รวบรวมวัตถุดิบ (Quota) ซึ่งมีจำนวนทั้งหมด 24 ราย เป็นอีกปัจจัยในด้านแหล่งของวัตถุดิบที่นำมาศึกษา โดยค่าเฉลี่ยของระดับความหวานวัตถุดิบในแต่ละผู้รวบรวมวัตถุดิบ แสดงในภาพที่ 15 เนื่องจากผู้วิจัยสนใจที่จะศึกษาถึงความแตกต่างของวัตถุดิบจากผู้รวบรวมรายหนึ่งๆ ในเขตพื้นที่ต่างๆ จึงได้สร้างตัวแปรอิสระเพิ่มขึ้น คือ ปัจจัยด้านแหล่งที่มาของวัตถุดิบ (Source) ซึ่งเป็นการแยกกลุ่มเป็นหน่วยย่อยตามอำเภอที่เป็นแหล่งเพาะปลูกกับผู้รวบรวมวัตถุดิบ มีทั้งหมด 48 แหล่ง แสดงค่าเฉลี่ยของระดับความหวานวัตถุดิบในแต่ละแหล่งที่มีของวัตถุดิบ ดังภาพที่ 16 และรายละเอียดของแหล่งวัตถุดิบสับปะรดในงานวิจัยนี้ แสดงในภาคผนวก ก



ภาพที่ 14 ค่าเฉลี่ยระดับความหวานวัตถุดิบสับปะรดในแต่ละอำเภอที่ปลูก



ภาพที่ 15 ค่าเฉลี่ยระดับความหวานวัตถุดิบสับปะรดในแต่ละผู้รวบรวมวัตถุดิบ



ภาพที่ 16 ค่าเฉลี่ยระดับความหวานวัดทุติยสัปดาห์ในแต่ละแหล่งที่มาของวัดทุติย

ในที่นี้ผู้วิจัยได้พิจารณาปัจจัยในด้านสถานะในการเพาะปลูก เช่น อุณหภูมิในช่วงการปลูกและเก็บเกี่ยว ปริมาณน้ำฝน สภาพพื้นที่เพาะปลูก ลักษณะดิน ฯลฯ โดยตรง ทั้งนี้เนื่องจากต้องการเลือกใช้ตัวแปรอิสระที่สอดคล้องกับโรงงานกรณีศึกษาในการทำนายค่าให้มากที่สุด ซึ่งเป็นไปตามความเห็นชอบจากฝ่ายผลิตของโรงงานกรณีศึกษา ตารางที่ 1 แสดงรายละเอียดของตัวแปรที่ใช้ในการสร้างแบบจำลองทำนายค่าความหวานของวัดทุติยสัปดาห์ในงานวิจัยนี้

ตารางที่ 1 รายละเอียดของตัวแปรอิสระและตัวแปรตามที่ใช้ในการสร้างแบบจำลองการทำนายค่า

ตัวแปรอิสระ	รายละเอียด	ตัวแปรตาม	รายละเอียด
X1	เดือนที่รับวัตถุดิบ (Month)	Y	ระดับความหวานของวัตถุดิบ
X2	เกรดสีวัตถุดิบ (Color)		(°Brix)
X3	อำเภอที่ปลูก (Amphoe)		
X4	ผู้รวบรวมวัตถุดิบ (Quota)		
X5	แหล่งที่มาวัตถุดิบ (Source)		

1.2 การเตรียมข้อมูล

ข้อมูลที่ได้จากโรงงานกรณีศึกษาเป็นข้อมูลที่จัดเก็บอยู่ในรูปแฟ้มเอกสาร ผู้วิจัยจึงได้ป้อนข้อมูลที่ได้ให้อยู่ในรูปอิเล็กทรอนิกส์ เพื่อความสะดวกในการนำข้อมูลดังกล่าวไปใช้ในการวิเคราะห์ข้อมูลด้วยโปรแกรมทางสถิติต่อไป ข้อมูลที่เก็บรวบรวมจากโรงงานกรณีศึกษามีทั้งหมด 2,202 ชุด แบ่งชุดข้อมูลออกเป็น 3 ส่วน โดยใช้วิธีการแบ่งข้อมูลแบบสุ่ม แบ่งได้เป็นดังนี้

ก. ชุดเรียนรู้ (Training set) มีจำนวนข้อมูล 70% ของจำนวนข้อมูลทั้งหมด เพื่อใช้ในการสร้างแบบจำลองความสัมพันธ์ระหว่างตัวแปรที่ต้องการศึกษาและค่าความหวานมีจำนวน 1,542 ชุด

ข. ชุดทดสอบ (Testing set) มีจำนวนข้อมูล 20% ของจำนวนข้อมูลทั้งหมด เพื่อสร้างแบบจำลองและช่วยในการเลือกโครงสร้างที่เหมาะสมของแบบจำลอง จำนวน 440 ชุด

ค. ชุดทดสอบ (Validating set) มีจำนวนข้อมูล 10% ของจำนวนข้อมูลทั้งหมด เพื่อทดสอบความสามารถในการใช้งานทั่วไป ในการทำนายค่ากับชุดข้อมูลที่ไม่เคยเรียนรู้ จำนวน 220 ชุด

2. การกำหนดตัวแปรที่ต้องการศึกษาความสัมพันธ์ ในแบบจำลองการทำนายค่าความหวานที่จะนำไปใช้ในการสร้างแบบจำลองทางสถิติต่าง ๆ

กำหนดแบบจำลองความสัมพันธ์ระหว่างตัวแปรอิสระที่ต้องการศึกษากับค่าความหวานที่จะใช้ในการวิจัยทางสถิติต่าง ๆ แบ่งออกเป็น 5 แบบจำลอง เพื่อแยกการพิจารณาตัวแปรอิสระที่เกี่ยวข้องกับแหล่งปลูก คือ อำเภอที่ปลูก (Amphoe) ผู้รวบรวมวัตถุดิบ (Quota หรือ Supplier) และแหล่งที่มาของวัตถุดิบ (Source) ในแบบจำลอง สำหรับตัวแปรอิสระเดือนที่รับวัตถุดิบ (Month) และเกรดสีวัตถุดิบ (Color) จะถูกพิจารณาในทุกแบบจำลอง การกำหนดตัวแปรอิสระในแบบจำลอง สามารถสรุปได้ดังนี้ (ตารางที่ 2)

2.1 แบบจำลอง MC (Month/Color) เป็นแบบจำลองที่พิจารณาปัจจัยในด้านฤดูกาลการปลูกร่วมกับความสุขของสับปะรด ส่งผลต่อค่าความหวานของสับปะรดรับเข้า แบบจำลองนี้จะไม่พิจารณาแหล่งปลูกวัตถุดิบมาใช้ในการทำนายค่า กำหนดตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (Month) และเกรดสีวัตถุดิบ (Color)

ตารางที่ 2 ตัวแปรอิสระที่พิจารณาในแบบจำลองการทำนายค่าความหวานวัตถุดิบ

แบบจำลอง	เกรดสี (Color)	เดือน (Month)	อำเภอ (Amphoe)	ผู้รวบรวม (Supplier)	แหล่งที่มา (Source)
	มาตราอันดับ (Ordinal Scale)	มาตรานามบัญญัติ (Nominal Scale)			
Model MC	X	X			
Model A	X	X	X		
Model Q	X	X		X	
Model AQ	X	X	X	X	
Model S	X	X			X

2.2 แบบจำลอง A (Month/Color/Amphoe) เป็นแบบจำลองที่นอกจากจะพิจารณาปัจจัยในด้านฤดูกาลปลูกและความสุขของสับปะรดแล้ว ยังพิจารณาแหล่งปลูกวัตถุดิบ คืออำเภอที่ปลูกเป็นปัจจัยที่แสดงถึงลักษณะทางภูมิศาสตร์และภูมิอากาศในพื้นที่เพาะปลูก มีผลต่อค่าความหวานของ

สับปะรดรับเข้า กำหนดตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (Month) เกรดสีวัตถุดิบ (Color) และ อำเภอที่ปลูก (Amphoe)

2.3 แบบจำลอง Q (Month/Color/Quota) เป็นแบบจำลองที่พิจารณาปัจจัยในด้านฤดูกาล ปลูกและความสุขของสับปะรด ร่วมกับผู้รวบรวมวัตถุดิบ เนื่องจากการรับวัตถุดิบของโรงงานกรณีศึกษาใช้ระบบการรับวัตถุดิบเป็นการซื้อขายในระบบสัญญา (Contract Farming) จึงนำปัจจัยผู้รวบรวมมาเป็นตัวแทน ปัจจัยด้านที่มาของวัตถุดิบ กำหนดตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (Month) เกรดสีวัตถุดิบ (Color) และ ผู้รวบรวมวัตถุดิบ (Quota)

2.4 แบบจำลอง AQ (Month/Color/Amphoe/Quota) เป็นแบบจำลองที่พิจารณาปัจจัยในด้านฤดูกาล ความสุขของสับปะรด รวมทั้งปัจจัยในด้านที่มาของวัตถุดิบทั้งหมด ได้แก่ อำเภอที่ปลูกและผู้รวบรวมวัตถุดิบ กำหนดตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (Month) เกรดสีวัตถุดิบ (Color) อำเภอที่ปลูก (Amphoe) และผู้รวบรวมวัตถุดิบ (Quota)

2.5 แบบจำลอง S (Month/Color/Source) เป็นแบบจำลองที่พิจารณาปัจจัยในด้านฤดูกาล ความสุขของสับปะรด และแหล่งที่มาของวัตถุดิบคล้ายกับแบบจำลอง AQ แต่ในแบบจำลองนี้ ผู้วิจัยได้สร้างตัวแปรใหม่ ซึ่งเป็นหน่วยย่อยจำเพาะระหว่างอำเภอและผู้รวบรวมวัตถุดิบ กำหนดตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (Month) เกรดสีวัตถุดิบ (Color) และ แหล่งที่มาของวัตถุดิบ (Source)

3. ศึกษาความเป็นไปได้ในการจำลองความสัมพันธ์ระหว่างตัวแปรทำนายค่ากับค่าความหวาน เพื่อสร้างแบบจำลองทางสถิติต่างๆ

ในขั้นตอนนี้ผู้วิจัยได้ทำการสร้างและทวนสอบแบบจำลองจากข้อมูลชุดเรียนรู้ ชุดทดสอบ และชุดทวนสอบ ที่ได้ทำการแบ่งจากข้อ 1.2.2 ดังนี้

3.1 แบบจำลองการถดถอย

ดำเนินการสร้างแบบจำลองการถดถอยด้วยข้อมูลที่จัดเตรียมไว้โดยโปรแกรมวิเคราะห์ทางสถิติ MINITAB 15.0 มีขั้นตอน ดังนี้

3.1.1 พิจารณาตัวแปรที่เป็นตัวแปรเชิงกลุ่ม ได้แก่ เดือนที่รับวัตถุดิบ อำเภอที่ปลูก ผู้รวบรวมวัตถุดิบ และแหล่งที่มาของวัตถุดิบ ทำการแปลงข้อมูลตัวแปรเชิงกลุ่มดังกล่าวให้อยู่ในรูปตัวแปรหุ่น (Dummy Variable) หรือการกำหนดรหัสตัวบ่งชี้ (Indicator Coding) ได้จำนวนตัวแปรเท่ากับ k ค่า เมื่อตัวแปรมีค่าที่เป็นไปได้ $k-1$ ค่า เพื่อให้สามารถนำตัวแปรที่ได้ไปใช้ในการวิเคราะห์เพื่อสร้างแบบจำลองต่อไป เช่น หากตัวแปรมีค่าที่เป็นไปได้จำนวน 3 ค่า ได้แก่ A B และ C ทำการแปลงข้อมูลเป็นตัวแปรหุ่นได้จำนวน 2 ตัวแปร

ลำดับชุดข้อมูล	X	X_1	X_2
1	A	1	0
2	B	0	1
3	C	0	0

3.1.2 สร้างแบบจำลองการถดถอยเชิงพหุ (Multiple Regression Model) ซึ่งเป็นแบบจำลองการถดถอยกำลังหนึ่ง (First-Order Regression Model) เพื่อหารูปแบบที่เหมาะสมในแบบจำลองความสัมพันธ์ระหว่างตัวแปรที่ต้องการศึกษาจากข้อมูลชุดเรียนรู้ โดยใช้วิธี Stepwise เพื่อคัดเลือกปัจจัยที่มีความสำคัญเข้าสู่แบบจำลอง และกำหนดให้ความน่าจะเป็นในการเลือกตัวแปรเข้าและตัวแปรออกเท่ากับ 0.05

3.1.3 ตรวจสอบความถูกต้องของข้อสมมติและเงื่อนไขของแบบจำลองการถดถอยที่ได้ ดังนี้

ก. ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) โดยการคำนวณค่า Variance Inflation Factor (VIF) ตัวแปรอิสระที่ใช้ในการสร้างแบบจำลองทุกตัว ควรเป็นอิสระต่อกัน นั่นคือค่า VIF ที่ได้ไม่ควรเกิน 10 (Kutner et al.,2005)

ข. ตรวจสอบวิเคราะห์ความคลาดเคลื่อนสุ่ม (Residual Analysis) มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 โดยตรวจสอบจากการสร้างแผนภาพการแจกแจงแบบปกติ (Normal Probability Plot) มีการเรียงตัวเป็นเส้นตรง และมีความแปรปรวนคงที่ โดยตรวจสอบจากการสร้างแผนภาพระหว่างค่าคลาดเคลื่อนและค่าผลลัพธ์จากการพยากรณ์ มีการกระจายตัวอย่างสุ่มไม่มีรูปแบบในการกระจายหรือไม่พบการกระจายออกนอกตอนปลายใดปลายหนึ่ง

3.1.4 นำแบบจำลองที่สร้างได้และผ่านการตรวจสอบความถูกต้องของข้อสมมติและเงื่อนไขต่างๆในการวิเคราะห์จากข้อ 3.1.3 มาทดสอบการทำนายค่าความหวาน ในข้อมูลชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2x เพื่อหารูปแบบที่เหมาะสมจากค่าคลาดเคลื่อนในรูปรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Squared Error, RMSE) และค่าเฉลี่ยของค่าความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error, MAE) ของชุดเรียนรู้และชุดทดสอบ ซึ่งคำนวณได้จากสมการที่ (17)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (17)$$

เมื่อ y_i แทนค่าจริงของอนุกรมเวลาที่เวลา i
 $\hat{y}_i$ แทนค่าของพยากรณ์อนุกรมเวลาที่เวลา i
 $y_i - \hat{y}_i$ แทนค่าความคลาดเคลื่อนของค่าพยากรณ์ (Forecast Error)
 n แทนจำนวนที่ทดสอบในชุดข้อมูลนั้น

3.1.5 ทดสอบความสามารถในการใช้งานทั่วไปของแบบจำลองที่มีรูปแบบที่เหมาะสม โดยทำการทดสอบการทำนายค่าความหวาน ในข้อมูลชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ค เพื่อหาความคลาดเคลื่อนในรูป RMSE และ MAE เปรียบเทียบค่าคลาดเคลื่อนที่ได้กับค่าคลาดเคลื่อนของข้อมูลชุดเรียนรู้และชุดทดสอบ

3.2 แบบจำลองเครือข่ายประสาทเทียม

ดำเนินการสร้างแบบจำลองเครือข่ายประสาทเทียมด้วยข้อมูลที่จัดเตรียมไว้ในชุดเรียนรู้และชุดทดสอบ ด้วยโปรแกรมสำเร็จรูป SPSS Clementine 12.0 แบบจำลองที่สร้างจากโปรแกรมเป็นแบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation Neural Network หรือ BPN) มีขั้นตอนดังนี้

3.2.1 พิจารณาตัวแปรที่เป็นตัวแปรเชิงกลุ่ม ได้แก่ เดือนที่รับวัตถุดิบ อำเภอที่ปลูก ผู้รวบรวมวัตถุดิบ และแหล่งที่มาของวัตถุดิบ ทำการแปลงข้อมูลตัวแปรเชิงกลุ่มให้อยู่ในรูปทวิภาค (Binary) คือ มีค่าเป็น 1 กับ 0 เมื่อตัวแปรมีค่าเป็นไปได้ k ค่า จะได้จำนวนตัวแปรเท่ากับ $\log_2(k)$ ค่า

(ปิดค่าขึ้น เป็นจำนวนเต็ม) เพื่อนำตัวแปรทวิภาคที่ได้ไปใช้ในการวิเคราะห์เพื่อสร้างแบบจำลองต่อไป เช่น มีค่าเป็นไปได้จำนวน 3 ค่า ได้แก่ A B และ C จะต้องทำการกำหนดรหัสได้จำนวนตัวแปรเท่ากับ $\log_2(3) \approx 1.58$ ได้จำนวน 2 ตัวแปร

ลำดับชุดข้อมูล	X	X ₁	X ₂
1	A	0	1
2	B	1	0
3	C	1	1

3.2.2 เลือกรูปแบบการเรียนรู้ในการสร้างแบบจำลองที่เหมาะสมในแต่ละแบบจำลองในตารางที่ 2 ด้วยโปรแกรมสำเร็จรูป SPSS Clementine 12.0 กำหนดวิธีการเรียนรู้แบบ Exhaustive Prune ทำการเปลี่ยนแปลงค่าฟังก์ชันต่างๆ เพื่อหาแบบจำลองที่ให้เครือข่ายที่เหมาะสม

3.2.3 ทดสอบความสามารถในการทำนายค่าความหวานในข้อมูลชุดเรียนรู้และชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ก และ 1.2.2ข เพื่อทำการเลือกวิธีการที่ดีที่สุดของแบบจำลองการทำนายค่าที่มีค่าความคลาดเคลื่อนในรูป RMSE และ MAE

3.2.4 ทดสอบความสามารถใช้งานทั่วไปของแบบจำลอง โดยใช้แบบจำลองจากวิธีการเรียนรู้ที่เหมาะสมจากข้อ 3.2.2 กับข้อมูลชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ค จำนวนความคลาดเคลื่อนในรูปของ RMSE และ MAE และเปรียบเทียบกับค่าคลาดเคลื่อนของชุดเรียนรู้และทดสอบ

3.3 แบบจำลองต้นไม้การตัดสินใจ CART

3.3.1 สร้างแบบจำลองต้นไม้การตัดสินใจ CART ที่เหมาะสม ในแต่ละแบบจำลองในตารางที่ 2 ด้วยโปรแกรมสำเร็จรูป SPSS Clementine 12.0

3.3.2 ทดสอบความสามารถในการทำนายค่าความหวานในข้อมูลชุดเรียนรู้และชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ก และ 1.2.2ข ในรูป RMSE และ MAE

3.3.3 ทดสอบความสามารถใช้งานทั่วไปของแบบจำลองในข้อมูลชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ก ในรูป RMSE และ MAE

3.4 แบบจำลองต้นไม้การตัดสินใจ CHAID

3.4.1 สร้างแบบจำลองต้นไม้การตัดสินใจ CHAID ที่เหมาะสม ในแต่ละแบบจำลองในตารางที่ 2 ด้วยโปรแกรม SPSS Clementine 12.0

3.4.2 ทดสอบความสามารถในการทำนายค่าความหวานในข้อมูลชุดเรียนรู้และชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ก และ 1.2.2ข ในรูป RMSE และ MAE

3.4.3 ทดสอบความสามารถใช้งานทั่วไปของแบบจำลองในข้อมูลชุดทดสอบที่ทำการแบ่งจากข้อ 1.2.2ค ในรูป RMSE และ MAE

4. สร้างระบบสนับสนุนการตัดสินใจ ในการใช้งานฐานข้อมูลการทำนายค่าความหวานวัตถุสิบประการ

4.1 กำหนดชุดข้อมูลทั้งหมดที่ต้องการในการทำนายค่าความหวาน เพื่อสร้างฐานข้อมูลให้กับผู้ใช้ (Users) ในการทำนายค่าความหวานวัตถุสิบประการ

4.2 คำนวณค่าทำนายจากแบบจำลองที่ได้จากการสร้างแบบจำลองในข้อ 3 ในแต่ละวิธีการโดยวิธีวิเคราะห์การถดถอยจะใช้การแทนค่าตามสมการที่ได้จากการสร้างแบบจำลองการถดถอย ขณะที่วิธีเครือข่ายประสาทเทียมและวิธีต้นไม้การตัดสินใจหาค่าทำนายจากโปรแกรม SPSS Clementine 12.0 โดยใช้โครงสร้างและวิธีการเรียนรู้ที่เหมาะสมที่คัดเลือกได้จากข้อ 3 จากนั้นเก็บค่าทำนายทั้งหมดที่ได้ในรูปแบบตาราง

4.3 นำข้อมูลที่เก็บเป็นฐานข้อมูลจากข้อ 4.2 เข้าโปรแกรม Microsoft Access 2007 สร้างฟอร์มการใช้งานตามความเหมาะสม เพื่อความสะดวกในการใช้งานของผู้ใช้ (Users) ในการทำนายค่าความหวานวัตถุสิบประการ เพื่อเป็นข้อมูลประกอบในการตัดสินใจเลือกค่าความหวานวัตถุสิบที่เหมาะสมเพื่อนำไปใช้ในการเตรียมสารบรรจุ (Packing Media) ในกระบวนการผลิตต่อไป

ผลและวิจารณ์

1. ผลการจำลองความสัมพันธ์ระหว่างตัวแปรทำนายค่ากับค่าความหวานของวัตถุดิบสับปะรด เพื่อสร้างแบบจำลองทางสถิติต่างๆ

1.1 แบบจำลองการถดถอย

1.1.1 การเตรียมข้อมูล ได้ทำการแปลงข้อมูลจากตัวแปรเชิงกลุ่มที่มีค่าเป็นไปได้อีก k ค่า กำหนดเป็นตัวแปรหุ่น (Dummy Variable) จำนวน $k - 1$ ตัว และกำหนดค่าในตัวแปรหุ่นเป็นค่า 0 และ 1 (กัลยา, 2549) หลังจากที่ได้กำหนดตัวแปรหุ่นที่ใช้ในการวิเคราะห์แล้วได้จำนวนตัวแปรใหม่ และกำหนดตัวแปรที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรอิสระแต่ละตัวแปรในตารางที่ 3

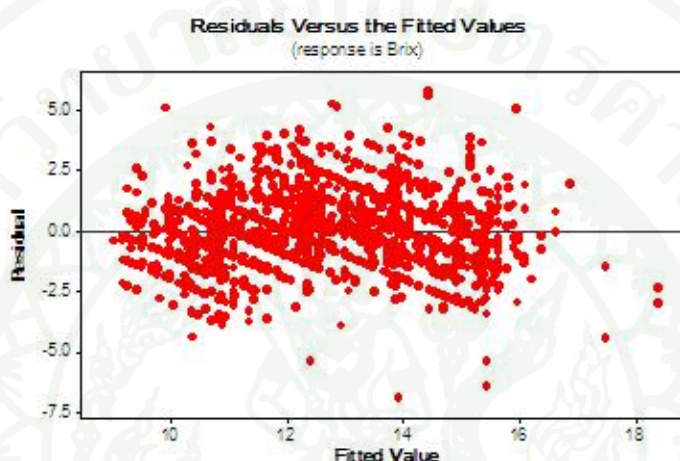
ตารางที่ 3 การกำหนดตัวแปรหุ่นที่ใช้ในการสร้างแบบจำลองการถดถอย

ตัวแปร	ตัวแปรหุ่น	รายละเอียดตัวแปร	ตัวแปรอ้างอิง
เดือนที่รับวัตถุดิบ (Month)	10 ตัวแปร (ยกเว้นเดือนกันยายน ซึ่งไม่มีการผลิต)	Feb – Dec	Jan
อำเภอที่ปลูก (Amphoe)	16 ตัวแปร	A01 - A16	A17
ผู้รวบรวมวัตถุดิบ (Quota)	23 ตัวแปร	Q01 - Q23	Q24
แหล่งที่มาของวัตถุดิบ (Source)	47 ตัวแปร	S01 - S47	S48

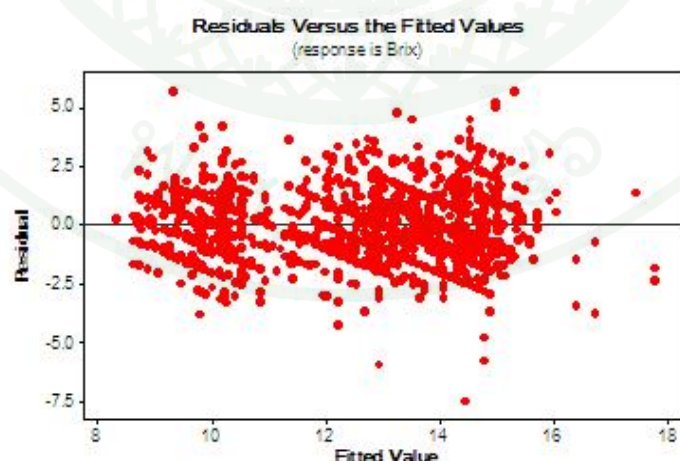
ตารางผนวกที่ ก1 – ก8 แสดงรายละเอียดของตัวแปรอิสระต่างๆที่เป็นตัวแปรหุ่นในแบบจำลองการถดถอย

1.1.2 วิเคราะห์การถดถอยเบื้องต้น เพื่อตรวจสอบความถูกต้องของข้อสมมติและเงื่อนไขของแบบจำลองการถดถอยที่ได้ โดยเมื่อทดสอบสร้างแบบจำลองความสัมพันธ์ระหว่างค่าตัวแปรอิสระกับค่าความหวานด้วยข้อมูลชุดสร้างแบบจำลอง และทำการตรวจสอบค่าคลาดเคลื่อนพบว่าแผนภาพการกระจาย (Residual Plots) มีการกระจายตัวเป็นแนวเส้นโค้ง ดังแสดงในภาพที่ 17 จึงทำการทดสอบสร้างแบบจำลองโดยเพิ่มค่ายกกำลังสอง (Quadratic Term) เข้าไปในแบบจำลอง

โดยกำหนดค่ายกกำลังสองให้กับตัวแปรเกรดสีวัตถุบ (Color) พบว่าแผนภาพการกระจายมีการกระจายตัวที่ดี ไม่มีรูปแบบความสัมพันธ์ที่ชัดเจนปรากฏ ดังแสดงในภาพที่ 18 ผู้วิจัยจึงเลือกสร้างแบบจำลองด้วยการวิเคราะห์การถดถอยหลายตัวแปร (Multiple Regression Analysis) โดยรวมค่ายกกำลังสองของตัวแปรเกรดสีวัตถุบ (Color Squared – ColorSq) ในแบบจำลองด้วย กำหนดระดับนัยสำคัญทางสถิติที่ใช้ในการสร้างแบบจำลอง เท่ากับ 0.05



ภาพที่ 17 แผนภาพการกระจายของค่าคลาดเคลื่อนในการวิเคราะห์การถดถอยเบื้องต้น
เมื่อไม่มีเทอมยกกำลังสองของตัวแปรเกรดสี



ภาพที่ 18 แผนภาพการกระจายของค่าคลาดเคลื่อนในการวิเคราะห์การถดถอยเบื้องต้น
เมื่อมีเทอมยกกำลังสองของตัวแปรเกรดสี

1.1.3 การกำหนดตัวแปรอิสระที่ใช้ในการสร้างแบบจำลองทางสถิติต่าง ๆ ในการทำนายค่าความหวานของวัตถุดิบสับปะรด โดยวิธีการวิเคราะห์การถดถอย ผู้วิจัยได้ทำการเพิ่มตัวแปรกำลังสองของตัวแปรเกรดสีวัตถุดิบเข้าไปในทุกแบบจำลอง ดังที่ได้แสดงการกำหนดตัวแปรอิสระในแต่ละแบบจำลองในตารางที่ 4

ตารางที่ 4 ตัวแปรอิสระที่พิจารณาในแบบจำลองการทำนายค่าความหวาน โดยการวิเคราะห์การถดถอย

แบบจำลอง (Model)	เกรดสี (Color)	เกรดสีกำลังสอง (ColorSq)	เดือน (Month)	อำเภอ (Amphoe)	ผู้รวบรวม (Quota)	แหล่งที่มา (Source)
	มาตราอันดับ (Ordinal Scale)		มาตรานามบัญญัติ (Nominal Scale)			
Model MC	X	X	X			
Model A	X	X	X	X		
Model Q	X	X	X		X	
Model AQ	X	X	X	X	X	
Model S	X	X	X			X

จากการตรวจสอบความถูกต้องของข้อสมมติและเงื่อนไขของแบบจำลอง พบว่า ข้อมูลมีการแจกแจงแบบปกติ (Normal Probability Plot) มีการเรียงตัวเป็นเส้นตรง และให้ค่า Variance Inflation Factor (VIF) ตัวแปรอิสระที่ใช้ในการสร้างแบบจำลองมีค่าน้อยกว่า 10 แสดงว่า แบบจำลองสามารถนำไปประยุกต์ใช้ในการทำนายค่าในงานวิจัยนี้ได้

1.1.4 ผลการสร้างแบบจำลองการทำนายค่าโดยการวิเคราะห์การถดถอย แบบ Stepwise Regression กำหนดระดับนัยสำคัญเท่ากับ 0.05 ได้สมการดังนี้

ก. การวิเคราะห์การถดถอยเชิงพหุ แบบจำลอง MC (Month/Color)

$$\begin{aligned} \text{Brix} = & 11.4 + 2.56\text{Color} - 0.582\text{ColorSq} + 0.52\text{Mar} + 0.62\text{Apr} + 1.15\text{May} \\ & + 1.0\text{Jun} + 0.44\text{Jul} - 0.58\text{Oct} - 0.74\text{Nov} \end{aligned}$$

แบบจำลองการถดถอยเชิงพหุมีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) เท่ากับ 0.6181 และค่าสัมประสิทธิ์สหสัมพันธ์ที่ปรับค่าแล้ว (Adjusted R^2) เท่ากับ 0.6158 ตัวแปรที่มีผลต่อค่าความหวานในแบบจำลองนี้ เมื่อกำหนดตัวแปรกลุ่มที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรเดือนที่รับวัตถุดิบ คือ เดือนมกราคม ได้แก่

Color	เกรดสีวัตถุดิบ
ColorSq	เทอมกำลังสองของเกรดสีวัตถุดิบ
Mar	วัตถุดิบที่รับในเดือนมีนาคม
Apr	วัตถุดิบที่รับในเดือนเมษายน
May	วัตถุดิบที่รับในเดือนพฤษภาคม
Jun	วัตถุดิบที่รับในเดือนมิถุนายน
Jul	วัตถุดิบที่รับในเดือนกรกฎาคม
Oct	วัตถุดิบที่รับในเดือนตุลาคม
Nov	วัตถุดิบที่รับในเดือนพฤศจิกายน

ข. การวิเคราะห์การถดถอยเชิงพหุ แบบจำลอง A (Month/Color/Amphoe)

$$\begin{aligned} \text{Brix} = & 11.4 + 2.56\text{Color} - 0.581\text{ColorSq} + 0.51\text{Mar} + 0.65\text{Apr} + 1.14\text{May} \\ & + 0.79\text{Jun} + 0.46\text{Jul} - 0.55\text{Oct} - 0.84\text{Nov} + 1.75\text{A03} - 0.64\text{A11} \\ & + 2.81\text{A16} \end{aligned}$$

แบบจำลองการถดถอยเชิงพหุมีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) เท่ากับ 0.6307 และค่าสัมประสิทธิ์สหสัมพันธ์ที่ปรับค่าแล้ว (Adjusted R^2) เท่ากับ 0.6278 ตัวแปรที่มีผลต่อค่าความหวานในแบบจำลองนี้ เมื่อกำหนดตัวแปรกลุ่มที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรเดือนที่รับวัตถุดิบ คือ เดือนมกราคม และในตัวแปรอำเภอที่ปลูก คือ อำเภอบ้านคา จังหวัดราชบุรี ได้แก่

Color	เกรดสีวัตถุดิบ
ColorSq	เทอมกำลังสองของเกรดสีวัตถุดิบ
Mar	วัตถุดิบที่รับในเดือนมีนาคม

Apr	วัตถุดิบที่รับในเดือนเมษายน
May	วัตถุดิบที่รับในเดือนพฤษภาคม
Jun	วัตถุดิบที่รับในเดือนมิถุนายน
Jul	วัตถุดิบที่รับในเดือนกรกฎาคม
Oct	วัตถุดิบที่รับในเดือนตุลาคม
Nov	วัตถุดิบที่รับในเดือนพฤศจิกายน
A03	วัตถุดิบที่มาจากอำเภอวังจันทน์ จังหวัดระยอง
A11	วัตถุดิบที่มาจากอำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์
A16	วัตถุดิบที่มาจากอำเภอห้วยมุ่น จังหวัดอุดรธานี

ก. การวิเคราะห์การถดถอยเชิงพหุ แบบจำลอง Q (Month/Color/Quota)

$$\begin{aligned} \text{Brix} = & 11.35 + 2.57\text{Color} - 0.583\text{ColorSq} + 0.63\text{Mar} + 0.72\text{Apr} + 1.26\text{May} \\ & + 1.02\text{Jun} + 0.42\text{Jul} - 0.57\text{Oct} - 0.74\text{Nov} + 0.89\text{Q13} - 0.6\text{Q14} \\ & + 0.35\text{Q17} + 0.73\text{Q18} - 1.38\text{Q22} \end{aligned}$$

แบบจำลองการถดถอยเชิงพหุมีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) เท่ากับ 0.6269 และค่าสัมประสิทธิ์สหสัมพันธ์ที่ปรับค่าแล้ว (Adjusted R^2) เท่ากับ 0.6235 ตัวแปรที่มีผลต่อค่าความหวานในแบบจำลองนี้ เมื่อกำหนดตัวแปรกลุ่มที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรเดือนที่รับวัตถุดิบ คือ เดือนมกราคม และในตัวแปรผู้รวบรวมวัตถุดิบ คือ ผู้รวบรวมที่ 24 ได้แก่

Color	เกรดสีวัตถุดิบ
ColorSq	เทอมกำลังสองของเกรดสีวัตถุดิบ
Mar	วัตถุดิบที่รับในเดือนมีนาคม
Apr	วัตถุดิบที่รับในเดือนเมษายน
May	วัตถุดิบที่รับในเดือนพฤษภาคม
Jun	วัตถุดิบที่รับในเดือนมิถุนายน
Jul	วัตถุดิบที่รับในเดือนกรกฎาคม
Oct	วัตถุดิบที่รับในเดือนตุลาคม
Nov	วัตถุดิบที่รับในเดือนพฤศจิกายน

Q13	วัตถุดิบที่มาจากผู้รวบรวมที่ 13
Q14	วัตถุดิบที่มาจากผู้รวบรวมที่ 14
Q17	วัตถุดิบที่มาจากผู้รวบรวมที่ 17
Q18	วัตถุดิบที่มาจากผู้รวบรวมที่ 18
Q22	วัตถุดิบที่มาจากผู้รวบรวมที่ 22

จ. การวิเคราะห์การถดถอยเชิงพหุ แบบจำลอง AQ (Month/Color/Amphoe
/Quota)

$$\begin{aligned} \text{Brix} = & 11.36 + 2.58\text{Color} - 0.59\text{ColorSq} + 0.59\text{Mar} + 0.71\text{Apr} + 1.19\text{May} \\ & + 0.73\text{Jun} + 0.42\text{Jul} - 0.57\text{Oct} - 0.89\text{Nov} + 1.81\text{A03} - 0.43\text{A10} \\ & - 0.63\text{A11} + 2.87\text{A16} + 0.63\text{Q02} + 1.3\text{Q12} - 0.59\text{Q14} + 0.38\text{Q17} \\ & + 0.69\text{Q18} - 1.42\text{Q21} \end{aligned}$$

แบบจำลองการถดถอยเชิงพหุมีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) เท่ากับ 0.6417 และค่าสัมประสิทธิ์สหสัมพันธ์ที่ปรับค่าแล้ว (Adjusted R^2) เท่ากับ 0.6372 ตัวแปรที่มีผลต่อค่าความหวานในแบบจำลองนี้ โดยกำหนดตัวแปรกลุ่มที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรเดือนที่รับวัตถุดิบ คือ เดือนมกราคม ในตัวแปรอำเภอที่ปลูก คือ อำเภอบ้านคา จังหวัดราชบุรี และในตัวแปรผู้รวบรวมวัตถุดิบ คือ ผู้รวบรวมที่ 24 ได้แก่

Color	เกรดสีวัตถุดิบ
ColorSq	เทอมกำลังสองของเกรดสีวัตถุดิบ
Mar	วัตถุดิบที่รับในเดือนมีนาคม
Apr	วัตถุดิบที่รับในเดือนเมษายน
May	วัตถุดิบที่รับในเดือนพฤษภาคม
Jun	วัตถุดิบที่รับในเดือนมิถุนายน
Jul	วัตถุดิบที่รับในเดือนกรกฎาคม
Oct	วัตถุดิบที่รับในเดือนตุลาคม
Nov	วัตถุดิบที่รับในเดือนพฤศจิกายน
A03	วัตถุดิบที่มาจากอำเภอวังจั่น จังหวัดระยอง

A10	วัตถุดิบที่มาจากอำเภอหัวหิน จังหวัดประจวบคีรีขันธ์
A11	วัตถุดิบที่มาจากอำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์
A16	วัตถุดิบที่มาจากอำเภอหัวหมื่น จังหวัดอุดรธานี
Q02	วัตถุดิบที่มาจากผู้รวบรวมที่ 2
Q12	วัตถุดิบที่มาจากผู้รวบรวมที่ 12
Q14	วัตถุดิบที่มาจากผู้รวบรวมที่ 14
Q17	วัตถุดิบที่มาจากผู้รวบรวมที่ 17
Q18	วัตถุดิบที่มาจากผู้รวบรวมที่ 18
Q21	วัตถุดิบที่มาจากผู้รวบรวมที่ 21

จ. การวิเคราะห์การถดถอยเชิงพหุ แบบจำลอง S (Month/Color/Source)

$$\begin{aligned} \text{Brix} = & 11.60 + 2.59\text{Color} - 0.586\text{ColorSq} + 0.59\text{Mar} + 0.76\text{Apr} + 1.22\text{May} \\ & + 0.52\text{Jun} + 0.52\text{Jul} + 0.87\text{Aug} - 0.61\text{Oct} - 0.85\text{Nov} - 0.44\text{S01} \\ & - 0.47\text{S02} - 0.63\text{S04} + 1.63\text{S05} - 0.86\text{S09} - 1.68\text{S11} - 0.48\text{S13} \\ & - 1.23\text{S17} + 3.49\text{S20} - 0.68\text{S22} - 0.36\text{S23} - 0.82\text{S33} - 0.77\text{S36} \\ & - 2.94\text{S38} - 0.84\text{S42} - 0.65\text{S46} + 2.83\text{S47} \end{aligned}$$

แบบจำลองการถดถอยเชิงพหุ มีค่าสัมประสิทธิ์สหสัมพันธ์ (R^2) เท่ากับ 0.6569 และค่าสัมประสิทธิ์สหสัมพันธ์ที่ปรับค่าแล้ว (Adjusted R^2) เท่ากับ 0.6507 ตัวแปรที่มีผลต่อค่าความหวานในแบบจำลองนี้ โดยกำหนดตัวแปรกลุ่มที่ใช้เป็นตัวแปรอ้างอิง (Reference Variable) ในตัวแปรเดือนที่รับวัตถุดิบ คือ เดือนมกราคม และในตัวแปรแหล่งที่มาของวัตถุดิบ คือ อำเภอบ้านคา จังหวัดราชบุรี ที่รับจากผู้รวบรวมวัตถุดิบ คือ ผู้รวบรวมที่ 17 ได้แก่

Color	เกรดสีวัตถุดิบ
ColorSq	เทอมกำลังสองของเกรดสีวัตถุดิบ
Mar	วัตถุดิบที่รับในเดือนมีนาคม
Apr	วัตถุดิบที่รับในเดือนเมษายน
May	วัตถุดิบที่รับในเดือนพฤษภาคม
Jun	วัตถุดิบที่รับในเดือนมิถุนายน

Jul	วัตถุดิบที่รับในเดือนกรกฎาคม
Aug	วัตถุดิบที่รับในเดือนสิงหาคม
Oct	วัตถุดิบที่รับในเดือนตุลาคม
Nov	วัตถุดิบที่รับในเดือนพฤศจิกายน
S01	วัตถุดิบที่มาจากอำเภอเขาชะเมา จังหวัดระยอง จากผู้รวบรวมที่ 24
S02	วัตถุดิบที่มาจากอำเภอปลวกแดง จังหวัดระยอง จากผู้รวบรวมที่ 1
S04	วัตถุดิบที่มาจากอำเภอปลวกแดง จังหวัดระยอง จากผู้รวบรวมที่ 9
S05	วัตถุดิบที่มาจากอำเภอวังจันทร์ จังหวัดระยอง จากผู้รวบรวมที่ 1
S09	วัตถุดิบที่มาจากอำเภอบ้านคา จังหวัดราชบุรี จากผู้รวบรวมที่ 14
S11	วัตถุดิบที่มาจากอำเภอบ้านคา จังหวัดราชบุรี จากผู้รวบรวมที่ 21
S13	วัตถุดิบที่มาจากอำเภอปากท่อ จังหวัดราชบุรี จากผู้รวบรวมที่ 15
S17	วัตถุดิบที่มาจากอำเภอเมือง จังหวัดชัยภูมิ จากผู้รวบรวมที่ 24
S20	วัตถุดิบที่มาจากอำเภอนครไทย จังหวัดพิษณุโลก จากผู้รวบรวมที่ 24
S22	วัตถุดิบที่มาจากอำเภอบ่อพลอย จังหวัดกาญจนบุรี จากผู้รวบรวมที่ 1
S23	วัตถุดิบที่มาจากอำเภอบ่อพลอย จังหวัดกาญจนบุรี จากผู้รวบรวมที่ 6
S33	วัตถุดิบที่มาจากอำเภอหัวหิน จังหวัดประจวบคีรีขันธ์ จากผู้รวบรวมที่ 11
S36	วัตถุดิบที่มาจากอำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์ จากผู้รวบรวมที่ 24
S38	วัตถุดิบที่มาจากอำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์ จากผู้รวบรวมที่ 13
S42	วัตถุดิบที่มาจากอำเภอแม่สาย จังหวัดเชียงราย จากผู้รวบรวมที่ 24
S46	วัตถุดิบที่มาจากอำเภอเขาสมิง จังหวัดตราด จากผู้รวบรวมที่ 19
S47	วัตถุดิบที่มาจากอำเภอห้วยมุ่น จังหวัดอุดรธานี จากผู้รวบรวมที่ 24

1.1.5 ทดสอบการทำนายค่าความหวานวัตถุดิบสับปะรดด้วยข้อมูลชุดเรียนรู้ ชุด

ทดสอบและชุดทวนสอบของแบบจำลองการถดถอยเชิงพหุแบบต่างๆ ในรูปของค่าความคลาดเคลื่อน RMSE และ MAE ได้ผลดังตารางที่ 5

แบบจำลองการถดถอยเชิงพหุในแบบจำลองต่างๆ ในชุดเรียนรู้ ชุดทดสอบ และชุดทวนสอบ ให้ค่าความคลาดเคลื่อน RMSE และ MAE ใกล้เคียงกัน กล่าวคือ ผู้ใช้สามารถใช้แบบจำลองได้ทุกแบบจำลองในการทำนายค่าความหวานวัตถุดิบได้ ค่าทำนายของทุกแบบจำลองจะใช้ในการสร้างฐานข้อมูลการทำนายค่าต่อไป

ตารางที่ 5 ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่าง ๆ โดยวิธีวิเคราะห์การถดถอยเชิงพหุ

แบบจำลอง	ชุดเรียนรู้		ชุดทดสอบ		ชุดทวนสอบ	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
MR_MC	1.47	1.14	1.45	1.11	1.39	1.13
MR_A	1.45	1.13	1.47	1.11	1.36	1.10
MR_Q	1.46	1.13	1.44	1.11	1.38	1.12
MR_AQ	1.43	1.11	1.44	1.10	1.35	1.09
MR_S	1.40	1.09	1.46	1.12	1.35	1.08

1.2 แบบจำลองเครือข่ายประสาทเทียม

1.2.1 การเตรียมข้อมูล ได้ทำการแปลงข้อมูลตัวแปรเชิงกลุ่มให้อยู่ในรูปทวิภาค (Binary) คือ มีค่าเป็น 1 กับ 0 หลังจากที่ได้กำหนดตัวแปรหุ่นที่ใช้ในการวิเคราะห์แล้วได้จำนวนตัวแปรใหม่ ดังนี้

เดือนที่รับวัตถุดิบ (Month)	กำหนดตัวแปรนำเข้า 4 ตัวแปร ได้แก่ M01-M04
เกรดสีวัตถุดิบ (Color)	กำหนดตัวแปรนำเข้า 4 ตัวแปร ได้แก่ G01-G04
อำเภอที่เป็นแหล่งปลูก (Amphoe)	กำหนดตัวแปรนำเข้า 5 ตัวแปร ได้แก่ A01-A05
ผู้รวบรวมวัตถุดิบ (Supplier)	กำหนดตัวแปรนำเข้า 5 ตัวแปร ได้แก่ Q01-Q05
แหล่งที่มาของวัตถุดิบ (Source)	กำหนดตัวแปรนำเข้า 6 ตัวแปร ได้แก่ S01-S06

ตารางผนวกที่ ก1 – ก3 แสดงรายละเอียดของตัวแปรอิสระต่างๆ และตารางที่ ก4 - ก8 แสดงค่าของตัวแปรหุ่นที่ใช้ในแบบจำลองเครือข่ายประสาทเทียม

1.2.2 ผลการหาโครงสร้างที่เหมาะสมของแบบจำลองเครือข่ายประสาทเทียมของแบบจำลองต่างๆ โดยโปรแกรม SPSS Clementine 12.0 เลือกวิธีการเรียนรู้เครือข่ายเป็นแบบ Exhaustive Prune สร้างเครือข่ายแบบป้อนสู่ด้านหน้าหลายชั้น (Multilayer Feed-forward Layers) แบบส่งถ่ายข้อมูลย้อนกลับ (Back-propagation of Error) ใช้การเรียนรู้ตามหลัก Delta Rule ฟังก์ชันกระตุ้นแบบซิกมอยด์ (Sigmoid Transfer Function) ได้โครงสร้างที่เหมาะสมดังตารางที่ 6 และตารางที่ 7 เป็นการแสดงค่าความคลาดเคลื่อนในรูปแบบ RMSE และ MAE ของแบบจำลองเครือข่ายประสาทเทียมทุกแบบจำลอง

ตารางที่ 6 โครงสร้างที่เหมาะสมของแบบจำลองต่างๆ โดยวิธีเครือข่ายประสาทเทียม

แบบจำลอง	หน่วยชั้นนำเข้า	โครงสร้าง			
		I	H1	H2	O
ANN_MC	M1, M2, M3, M4, C1, C2, C3, C4	8	6	3	1
ANN_A	M1, M2, M3, M4, C1, C2, C3, C4, A1, A2, A3, A4, A5	13	30	18	1
ANN_Q	M1, M2, M3, M4, C1, C2, C3, C4, Q1, Q2, Q3, Q5	12	30	16	1
ANN_AQ	M1, M2, M3, M4, C1, C2, C3, C4, A1, A2, A3, A4, A5, Q1, Q2, Q3, Q4, Q5	18	30	18	1
ANN_S	M1, M2, M3, M4, C1, C2, C3, C4, S1, S2, S3, S4, S5, S6	14	7	2	1

หมายเหตุ I จำนวนหน่วยในชั้นนำเข้า (Input Neurons)
H1 จำนวนหน่วยในชั้นซ่อนชั้นที่ 1 (Hidden 1 Neurons)
H2 จำนวนหน่วยในชั้นซ่อนชั้นที่ 2 (Hidden 2 Neurons)
O จำนวนหน่วยในชั้นผลลัพธ์ (Output Neurons)

แบบจำลองเครือข่ายประสาทเทียมให้ค่าความคลาดเคลื่อน RMSE และ MAE ดังแสดงในตารางที่ 7 มีค่าใกล้เคียงกัน ทั้งในชุดเรียนรู้ ชุดทดสอบ และชุดทวนสอบ ดังนั้น ผู้ใช้สามารถนำทุกแบบจำลองในตารางที่ 6 ไปใช้ในการทำนายค่าความหวานวัตถุดิบได้ ซึ่งค่าทำนายของทุกแบบจำลองที่ได้พัฒนาขึ้นนี้จะนำไปใช้ในการสร้างฐานข้อมูลการทำนายค่าต่อไป

ตารางที่ 7 ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธีเครือข่ายประสาทเทียม

แบบจำลอง	ชุดเรียนรู้		ชุดทดสอบ		ชุดทวนสอบ	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
ANN_MC	1.45	1.12	1.42	1.11	1.42	1.15
ANN_A	1.30	1.02	1.35	1.04	1.38	1.11
ANN_Q	1.35	1.05	1.34	1.05	1.48	1.17
ANN_AQ	1.26	0.99	1.32	1.04	1.40	1.12
ANN_S	1.30	1.02	1.34	1.06	1.33	1.07

1.3 แบบจำลอง Classification And Regression Trees (CART)

การวัดค่า Impurity ใช้วิธีค่าเบี่ยงเบนกำลังสองน้อยที่สุด (Least-Squared Deviation: LSD) กำหนดระดับสูงสุดในการแบ่งเท่ากับ 5 ระดับ แสดงผลการแบ่งกลุ่มตามคุณลักษณะของข้อมูลได้ ดังในภาพผนวก ข ในตารางผนวกที่ ข1 – ข5

ตารางที่ 8 ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธี CART

แบบจำลอง	ชุดเรียนรู้		ชุดทดสอบ		ชุดทวนสอบ	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CART_MC	1.45	1.12	1.41	1.10	1.41	1.14
CART_A	1.45	1.13	1.43	1.11	1.46	1.20
CART_Q	1.40	1.08	1.38	1.05	1.45	1.18
CART_AQ	1.43	1.12	1.40	1.10	1.50	1.24
CART_S	1.35	1.03	1.37	1.05	1.43	1.16

จากตารางที่ 8 แบบจำลอง CART ต่างๆ ให้ค่าความคลาดเคลื่อน RMSE และ MAE ใกล้เคียงกัน ทั้งในชุดเรียนรู้ ชุดทดสอบ และชุดทวนสอบ ดังนั้น ผู้ใช้สามารถใช้แบบจำลองทุกแบบจำลอง CART ในตารางที่ 8 ในการทำนายค่าความหวานวัตถุดิบได้ ซึ่งค่าทำนายของทุกแบบจำลองนี้ จะใช้ในการสร้างฐานข้อมูลการทำนายค่าต่อไป

1.4 แบบจำลอง Chi-Square Automatic Iteration Detector (CHAID)

ใช้วิธีการรวมกลุ่มข้อมูลแบบ Exhaustive CHAID กำหนดระดับสูงสุดในการแบ่งเท่ากับ 5 ระดับ แสดงผลการแบ่งกลุ่มตามคุณลักษณะของข้อมูลได้ ดังในภาคผนวก ข ตารางผนวกที่ ข6 – ข10

ตารางที่ 9 แสดงค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลอง CHAID ต่างๆ ซึ่งให้ค่าความคลาดเคลื่อนใกล้เคียงกัน กล่าวคือ ผู้ใช้สามารถใช้แบบจำลอง CHAID ในตารางที่ 9 ได้ทุกแบบจำลองในการทำนายค่าความหวานวัตถุดิบได้ ซึ่งค่าทำนายของทุกแบบจำลองจะใช้ในการสร้างฐานข้อมูลการทำนายค่าต่อไป

ตารางที่ 9 ค่าความคลาดเคลื่อน RMSE และ MAE ของแบบจำลองต่างๆ โดยวิธี CHAID

แบบจำลอง	ชุดเรียนรู้		ชุดทดสอบ		ชุดทวนสอบ	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CHAID_MC	1.45	1.12	1.41	1.09	1.42	1.15
CHAID_A	1.40	1.08	1.39	1.09	1.43	1.16
CHAID_Q	1.38	1.06	1.38	1.07	1.45	1.18
CHAID_AQ	1.37	1.05	1.36	1.06	1.42	1.17
CHAID_S	1.33	1.03	1.37	1.06	1.32	1.07

1.5 การเปรียบเทียบผลการทำนายที่ได้จากแบบจำลองต่างๆ

จากผลการสร้างแบบจำลองทางคณิตศาสตร์และสถิติ พบว่า แบบจำลองที่ให้ค่าการทำนายใกล้เคียงกับค่าความหวานจริงมากที่สุดหรือให้ค่าความคลาดเคลื่อนโดยรวมต่ำที่สุด คือแบบจำลองเครือข่ายประสาทเทียม อย่างไรก็ตามในการหาโครงสร้างเครือข่ายที่เหมาะสมทำได้ยาก และแบบจำลองมีการเก็บค่าน้ำหนักของเครือข่ายในรูปกล่องดำ (Black Box) ทำให้ไม่สามารถที่จะนำข้อมูลดังกล่าวมาใช้ในการวิเคราะห์ตัวแปรอิสระที่นำเข้าไปในแบบจำลองการทำนายได้ อีกทั้งจำเป็นต้องหาค่าทำนายผ่านการใช้โปรแกรมสำเร็จรูปที่รองรับการทำงานเครือข่ายประสาทเทียม ทำให้การนำวิธีเครือข่ายประสาทเทียมไปประยุกต์ใช้มีความซับซ้อน โดยภาคผนวก ค แสดงการตรวจสอบค่าความคลาดเคลื่อนของแบบจำลองต่างๆ

เมื่อพิจารณาแบบจำลองการถดถอยและแบบจำลองต้นไม้ตัดสินใจ พบว่า เดือนที่รับ วัสดุคิบมีผลต่อค่าความหวาน โดยเดือนที่ให้ค่าความหวานวัสดุคิบที่ระดับเกรดเดียวกันมีค่าความหวานมาก คือ วัสดุคิบที่รับในเดือนพฤษภาคมและเดือนมิถุนายน และเดือนที่ให้ค่าความหวานน้อย คือ วัสดุคิบที่รับในเดือนตุลาคมและเดือนพฤศจิกายน เนื่องจากสภาพแวดล้อมที่เหมาะสมในการ เพาะปลูกสับปะรดคือที่อากาศค่อนข้างร้อน อุณหภูมิที่เหมาะสมอยู่ระหว่าง 23.9-29.4 องศาเซลเซียส ปริมาณน้ำฝนอยู่ในช่วง 1,000-1,500 มิลลิเมตรต่อปี มีปริมาณน้ำฝนสม่ำเสมอตลอดปี และมีความชื้นในอากาศสูง (เกตุอร, 2536) ทำให้ปัจจัยเดือนที่รับสับปะรดส่งผลต่อค่าความหวานที่ระดับ เกรดเดียวกัน ปัจจัยเกรดมีผลต่อค่าความหวานจากแบบจำลองการถดถอย พบว่า เกรดที่ 2 ซึ่งเป็น เกรดที่มีความเหลืองมากที่สุดจะให้ค่าความหวานมากที่สุด และเกรดที่ 3, 4 และ 5 ให้ค่าความ หวานน้อยลง ตามลำดับ สำหรับเกรดที่ 1 นั้น พบว่า เป็นเกรดที่หายากและไม่ค่อยพบวัสดุคิบ สับปะรดที่มีเกรดนี้ แต่จากแบบจำลองต้นไม้ตัดสินใจ มีการจัดกลุ่มโดยรวมเกรดที่ 2 และ 3 เข้าเป็นกลุ่มที่ให้ค่าความหวานมาก และ เกรดที่ 4 และ 5 ให้ค่าความหวานน้อยลงตามลำดับ

นอกจากนี้บางแหล่งปลูกจะให้ค่าความหวานที่แตกต่างออกไป ซึ่งจากแบบจำลองการ ถดถอย แสดงรายละเอียดได้ว่า วัสดุคิบจากอำเภอห้วยมุ่น จังหวัดอุดรดิตถ์ และอำเภอวังจันทน์ จังหวัด ระยอง ให้ค่าความหวานที่มากกว่าวัสดุคิบจากอำเภออื่น และวัสดุคิบจากอำเภอสามร้อยยอด จังหวัด ประจวบคีรีขันธ์ ให้ค่าความหวานที่น้อยกว่าวัสดุคิบจากอำเภออื่นที่เกรดเดียวกัน เนื่อง จากความแตกต่างระหว่างอุณหภูมิเพาะปลูก ปริมาณน้ำฝน ลักษณะดินที่ใช้ปลูก ตลอดจนการใส่ ปุ๋ยและการดูแลรักษา

เมื่อพิจารณาจากแบบจำลองการถดถอยที่ศึกษาปัจจัยในด้านแหล่งที่มาวัสดุคิบ ซึ่งเป็น หน่วยย่อยที่เฉพาะระหว่างอำเภอที่ปลูกและผู้รวบรวมวัสดุคิบ พบว่า

1. แหล่งที่มาของวัสดุคิบที่ให้ค่าความหวานมากกว่าวัสดุคิบจากแหล่งอื่นที่เกรดที่ ระดับเดียวกัน คือ วัสดุคิบที่รับจากผู้รวบรวมวัสดุคิบที่ 24 ที่ปลูกในอำเภอนครไทย จังหวัดพิษณุโลก และอำเภอห้วยมุ่น จังหวัดอุดรดิตถ์ ทั้งนี้ผู้รวบรวมที่ 24 ได้รับวัสดุคิบที่ปลูกในอำเภออื่นๆด้วย แต่ให้ค่าความหวานไม่แตกต่างจากแหล่งที่มาอื่นที่เกรดในระดับเดียวกัน ได้แก่ อำเภอเขาชะเมา จังหวัดระยอง อำเภอบ้านคา จังหวัดราชบุรี อำเภอเมืองชัยภูมิ จังหวัดชัยภูมิ อำเภอปอพลอย จังหวัดกาญจนบุรี อำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์ อำเภอบ้านไร่ จังหวัดอุทัยธานี อำเภอแม่สาย จังหวัดเชียงราย และอำเภอแม่จัน จังหวัดเชียงราย ทั้งนี้วัสดุคิบจากอำเภอห้วยมุ่น

รับผ่านผู้รวบรวมที่ 24 เพียงรายเดียว ขณะที่วัตถุดิบที่ปลูกในอำเภอ นครไทจะผ่านผู้รวบรวมที่ 1 และ 24 โดยที่วัตถุดิบที่ปลูกในอำเภอ นครไทที่รับผ่านผู้รวบรวมที่ 1 จะมีระดับความหวานใกล้เคียงกับแหล่งที่มาอื่น

2. แหล่งที่มาของวัตถุดิบที่ให้ค่าความหวานน้อยกว่าวัตถุดิบจากแหล่งอื่นที่เกรดสีระดับเดียวกัน คือ วัตถุดิบที่รับจากผู้รวบรวมวัตถุดิบที่ 13 ซึ่งมาจากอำเภอสามร้อยยอด จังหวัดประจวบคีรีขันธ์ ทั้งนี้ผู้รวบรวมที่ 13 ได้รับวัตถุดิบจากอำเภอบ้านคา จังหวัดราชบุรี ด้วยแต่ให้ค่าความหวานไม่แตกต่างจากแหล่งที่มาอื่นที่เกรดสีในระดับเดียวกัน และวัตถุดิบที่ปลูกในอำเภอสามร้อยยอดจะผ่านผู้รวบรวมที่ 9, 13 และ 24 โดยเฉพาะผู้รวบรวมที่ 13 ให้ระดับความหวานน้อยกว่าวัตถุดิบจากแหล่งอื่น

ดังนั้นในการรับวัตถุดิบเข้ากระบวนการผลิต ทางโรงงานกรณีศึกษาควรมีการคัดแยกวัตถุดิบจากแหล่งต่างๆเหล่านี้ออกจากแหล่งวัตถุดิบอื่นๆ เพื่อให้กระบวนการผลิตมีความต่อเนื่องและให้ค่าความหวานของผลิตภัณฑ์ที่เป็นไปตามมาตรฐานที่กำหนด

2. ผลการสร้างระบบสนับสนุนการตัดสินใจ ในการสร้างฐานข้อมูลการทำนายค่าความหวาน โดยแบบจำลองทั้งหมดที่ได้ทำการศึกษา

2.1 การเก็บรวบรวมข้อมูลค่าทำนายค่าความหวานของวัตถุดิบสับปะรด

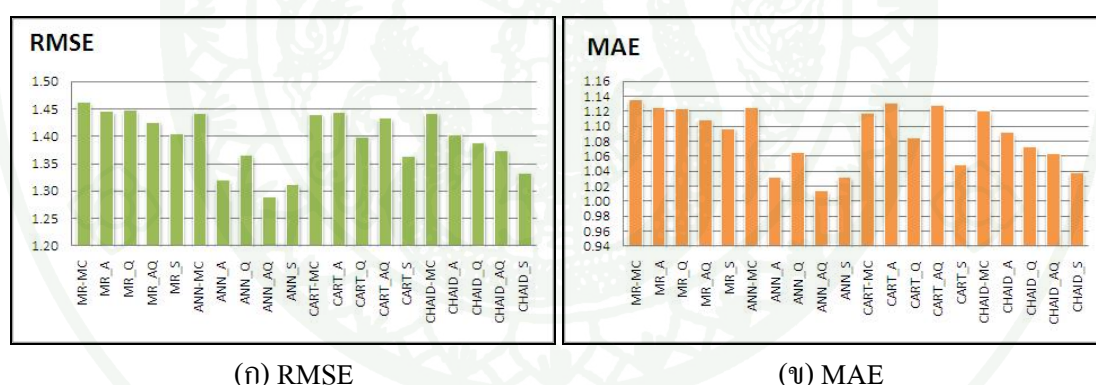
ในการทำนายค่าความหวาน สามารถแบ่งระดับความน่าเชื่อถือในการทำนายค่าความหวานได้ตามชุดข้อมูล ซึ่งเป็นรายละเอียดของวัตถุดิบสับปะรดรับเข้าที่ใช้เป็นตัวแปรนำเข้า เมื่อเปรียบเทียบกับชุดข้อมูลที่ใช้ในการสร้างแบบจำลอง ผู้ใช้จำเป็นต้องทราบระดับความน่าเชื่อถือของการทำนายค่าของชุดข้อมูล เพื่อใช้ประกอบการตัดสินใจในการนำค่าทำนายความหวานที่ได้ไปใช้ในกระบวนการผลิตต่อไป โดยแบ่งระดับความน่าเชื่อถือได้ดังนี้

2.1.1 ระดับสูงมาก (Very High)

ที่ระดับความน่าเชื่อถือสูงมาก ชุดข้อมูลวัตถุดิบทุกตัวแปร ได้แก่ เดือนที่รับวัตถุดิบ อำเภอที่ปลูก ผู้รวบรวมวัตถุดิบ และเกรดสีวัตถุดิบที่ต้องการทำนาย เป็นชุดข้อมูลที่ตรงกับ

ชุดข้อมูลที่ใช้เป็นข้อมูลในการวิเคราะห์ทางสถิติเพื่อสร้างแบบจำลอง ทำให้มีความแม่นยำในการทำนายค่าได้มากที่สุด เมื่อผู้ใช้ทำการทำนายค่าความหวานโดยชุดข้อมูลในระดับนี้ จะมีค่าความหวานจริงในอดีต ซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองปรากฏในฐานข้อมูลด้วย

ค่าทำนายที่ได้จะเป็นค่าทำนายที่มาจากแบบจำลองการทำนายค่าทั้งหมด ได้แก่ แบบจำลอง MC (Month/Color) แบบจำลอง A (Month/Color/Amphoe) แบบจำลอง Q (Month/Color/Quota) แบบจำลอง AQ (Month/Color/Amphoe/Quota) และแบบจำลอง S (Month/Color/Source) ซึ่งได้จากการทำนายโดยวิธีแบบจำลองการถดถอย (MR) แบบจำลองเครือข่ายประสาทเทียม (ANN) แบบจำลอง CART และแบบจำลอง CHAID ดังนั้นในการทำนายค่าของชุดข้อมูลรายละเอียดของสัปดาห์รับเข้าแต่ละชุดในระดับสูงมากนี้ ข้อมูลจะทำนายค่าความหวานได้ทั้งหมด 20 ค่า ภาพที่ 19 แสดงการเปรียบเทียบความคลาดเคลื่อน RMSE และ MAE โดยรวมจากแบบจำลองต่างๆที่สามารถทำนายได้ของชุดข้อมูลที่มีระดับความน่าเชื่อถือในการทำนายค่าความหวานในระดับสูงมาก



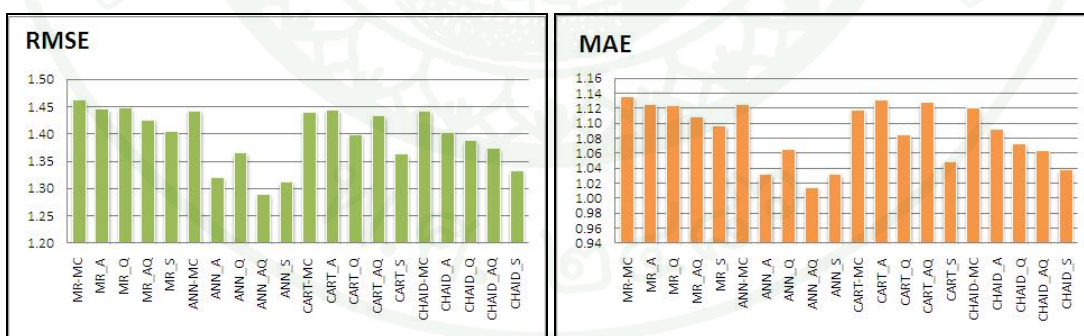
ภาพที่ 19 ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลในระดับสูงมาก

จากภาพที่ 19 แบบจำลองที่ดีที่สุดที่ให้ค่า MAE และ RMSE ต่ำที่สุด คือแบบจำลองเครือข่ายประสาทเทียม (ANN) ที่กำหนดตัวแปรนำเข้า เป็นแบบจำลอง AQ (Month/Color/Amphoe/Quota) ได้แก่ เดือนที่รับวัตถุดิบ เกรดสีวัตถุดิบ อำเภอที่ปลูกและผู้รวบรวมวัตถุดิบ

2.1.2 ระดับสูง (High)

ที่ระดับความน่าเชื่อถือสูง ชุดข้อมูลด้านแหล่งที่มาของวัตถุดิบ ได้แก่ อำเภอที่ปลูกและผู้รวบรวมวัตถุดิบที่ต้องการทำนายค่าเป็นชุดข้อมูลที่ตรงกับฐานข้อมูลที่ใช้สร้างแบบจำลอง แต่ไม่ตรงกับเดือนที่ต้องการทำนายค่า ระดับความน่าเชื่อถือของข้อมูลในระดับนี้จะน้อยกว่าในระดับสูงมาก กล่าวคือ มีค่าความคลาดเคลื่อนจากการทำนายที่มากกว่า เนื่องจากเป็นชุดข้อมูลที่ตัวแปรนำเข้ามีค่าไม่ตรงกับข้อมูลที่ใช้ในการสร้างแบบจำลองทั้งหมด

ค่าทำนายที่ได้สามารถทำนายได้จากแบบจำลองทั้งหมด ได้แก่ แบบจำลอง MC (Month/Color) แบบจำลอง A (Month/Color/Amphoe) แบบจำลอง Q (Month/Color/Quota) แบบจำลอง AQ (Month/Color/Amphoe/Quota) และแบบจำลอง S (Month/Color/Source) ซึ่งได้จากการทำนายโดยวิธีแบบจำลองการถดถอย (MR) แบบจำลองเครือข่ายประสาทเทียม (ANN) แบบจำลอง CART และแบบจำลอง CHAID ดังนั้นในการทำนายค่าของชุดข้อมูลรายละเอียดของลัทธิประดริบเข้าแต่ละชุดข้อมูลจะทำนายค่าความหวานได้ทั้งหมด 20 ค่า ภาพที่ 20 แสดงการเปรียบเทียบความคลาดเคลื่อน RMSE และ MAE โดยรวม จากแบบจำลองต่างๆที่สามารถทำนายได้ของชุดข้อมูลที่มีระดับความน่าเชื่อถือในการทำนายค่าความหวานในระดับสูง



(ก) RMSE

(ข) MAE

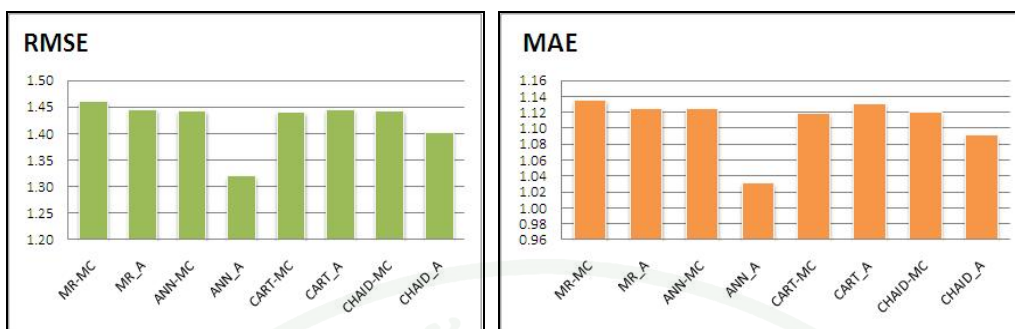
ภาพที่ 20 ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่มีระดับสูง

จากภาพที่ 20 แบบจำลองที่ดีที่สุดที่ให้ค่า RMSE และ MAE ต่ำที่สุด คือ แบบจำลองเครือข่ายประสาทเทียม (ANN) ที่กำหนดตัวแปรนำเข้า เป็นแบบจำลอง AQ (Month/Color/Amphoe/Quota) ได้แก่ เดือนที่รับวัตถุดิบ เกรดสีวัตถุดิบ อำเภอที่ปลูกและผู้รวบรวมวัตถุดิบ

2.1.3 ระดับปานกลาง (Medium)

ที่ระดับความน่าเชื่อถือปานกลาง ชุดข้อมูลที่ใช้การทำนายค่าชุดข้อมูล ซึ่งอำเภอที่ปลูกและผู้รวบรวมวัตถุดิบตรงกับข้อมูลที่ใช้สร้างแบบจำลองการทำนายค่า แต่เป็นการจับคู่ใหม่ระหว่างอำเภอที่ปลูกและผู้รวบรวมวัตถุดิบที่เป็นข้อมูลที่ใช้สร้างแบบจำลอง ทำให้มีระดับความน่าเชื่อถือลดน้อยลงมากกว่าระดับสูง หรือให้ค่าความคลาดเคลื่อนจากการทำนายค่าที่เพิ่มมากขึ้น เนื่องจากตัวแปรที่ใช้ในการทำนายค่ามีค่าไม่ตรงกับข้อมูลที่ใช้ในแบบจำลอง เช่น เป็นวัตถุดิบที่มาจากอำเภอ A แต่ผู้ส่งมอบเป็น B ซึ่งข้อมูลในฐานข้อมูลโรงงานไม่เคยรับวัตถุดิบ โดยผู้ส่งมอบ B ที่มาจากอำเภอ A ทั้งนี้แบบจำลองจะทำนายค่าความหวาน โดยอาศัยตัวแปรเดือนที่รับวัตถุดิบ ระดับสีของสับปะรด (ระดับความสุก) และอำเภอที่รับวัตถุดิบ (ตัวแปรอิสระ) เป็นตัวแปรนำเข้า

ค่าทำนายที่ได้สามารถทำนายได้จากแบบจำลอง ได้แก่ แบบจำลอง MC (Month/Color) และแบบจำลอง A (Month/Color/Amphoe) ซึ่งได้จากการทำนายโดยวิธีแบบจำลองการถดถอย (MR) แบบจำลองเครือข่ายประสาทเทียม (ANN) แบบจำลอง CART และแบบจำลอง CHAID ดังนั้นในการทำนายค่าของชุดข้อมูลรายละเอียดของสับปะรดรับเข้าแต่ละชุดข้อมูลจะทำนายค่าความหวานได้ทั้งหมด 8 ค่า ภาพที่ 21 แสดงการเปรียบเทียบความคลาดเคลื่อน RMSE และ MAE โดยรวม จากแบบจำลองต่างๆที่สามารถทำนายได้ของชุดข้อมูลที่มีระดับความน่าเชื่อถือในการทำนายค่าความหวานในระดับปานกลาง แบบจำลองที่ดีที่สุดที่ให้ค่า RMSE และ MAE ต่ำที่สุด คือ แบบจำลองเครือข่ายประสาทเทียม (ANN) ที่กำหนดตัวแปรนำเข้าเป็นแบบจำลอง A (Month/Color/Amphoe) ได้แก่ เดือนที่รับวัตถุดิบ เกรดสีวัตถุดิบ และอำเภอที่รับวัตถุดิบ



(ก) RMSE

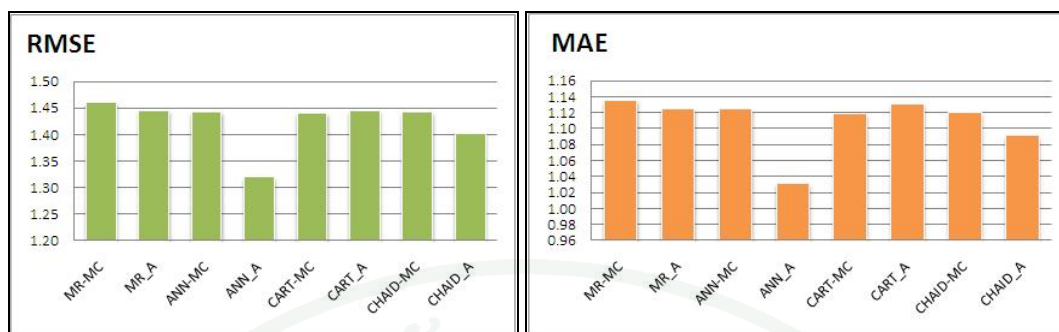
(ข) MAE

ภาพที่ 21 ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับปานกลาง

2.1.4 ระดับต่ำ (Low)

ในกรณีนี้ชุดข้อมูลที่มีการรับวัตถุดิบจากผู้รวบรวมใหม่ แต่มาจากอำเภอที่ปลูก ซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลอง ซึ่งเป็นชุดข้อมูลที่ไม่ได้อยู่ในฐานข้อมูลที่ใช้ในการวิเคราะห์เพื่อสร้างแบบจำลอง ทำให้มีความคลาดเคลื่อนจากการทำนายค่ามากกว่าในระดับปานกลาง การทำนายค่าจะใช้ตัวแปรเดือน (ตัวแปรด้านฤดูกาล) ระดับ สีของสับปะรด (ระดับความสุก) และอำเภอที่รับวัตถุดิบเป็นตัวแปรนำเข้า (ตัวแปรอิสระ)

ค่าทำนายที่ได้สามารถทำนายได้จากแบบจำลอง ได้แก่ แบบจำลอง MC (Month/Color) และแบบจำลอง A (Month/Color/Amphoe) ซึ่งได้จากการทำนายโดยวิธีแบบจำลองการถดถอย (MR) แบบจำลองเครือข่ายประสาทเทียม (ANN) แบบจำลอง CART และแบบจำลอง CHAID ดังนั้นในการทำนายค่าของชุดข้อมูลรายละเอียดของสับปะรดรับเข้าแต่ละชุดข้อมูลจะทำนายค่าความหวานได้ทั้งหมด 8 ค่า ภาพที่ 22 แสดงการเปรียบเทียบความคลาดเคลื่อน RMSE และ MAE โดยรวมจากแบบจำลองต่างๆที่สามารถทำนายได้ของชุดข้อมูลที่มีระดับความน่าเชื่อถือในการทำนายค่าความหวานในระดับสูงต่ำ



(ก) RMSE

(ข) MAE

ภาพที่ 22 ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับต่ำ

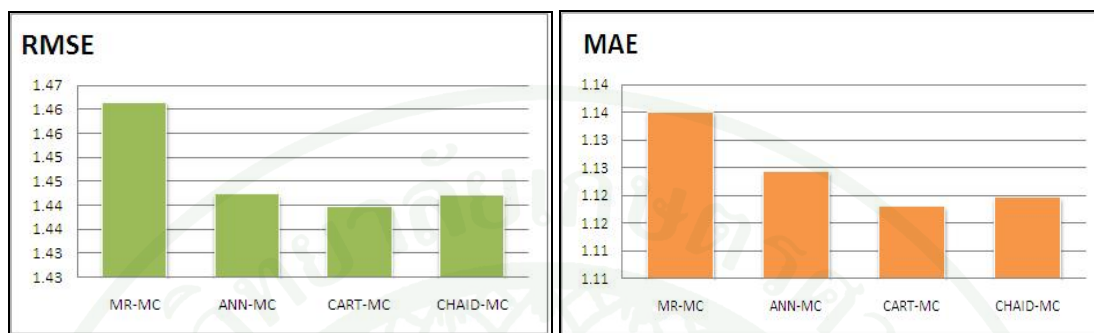
จากภาพที่ 22 แบบจำลองที่ดีที่สุดที่ให้ค่า RMSE และ MAE ต่ำที่สุด คือแบบจำลองเครือข่ายประสาทเทียม (ANN) ที่กำหนดตัวแปรนำเข้า เป็นแบบจำลอง A (Month/Color/Amphoe) ได้แก่ เดือนที่รับวัตถุดิบ เกรดสีวัตถุดิบ และอำเภอที่ปลูก

2.1.5 ระดับต่ำมาก (Very Low)

ในกรณีนี้ชุดข้อมูลที่มีการรับวัตถุดิบจากอำเภอที่ปลูกใหม่หรือเป็นอำเภอที่ไม่ได้เป็นข้อมูลที่ใช้ในการสร้างแบบจำลอง หรือรับจากแหล่งปลูกใหม่ (ได้แก่ อำเภอที่ปลูกใหม่ซึ่งรับมาจากผู้รวบรวมใหม่ด้วย) ซึ่งเป็นชุดข้อมูลที่ไม่ได้เป็นข้อมูลที่ใช้ในการสร้างแบบจำลอง ทำให้มีความคลาดเคลื่อนจากการทำนายค่ามากกว่าระดับอื่นๆ การทำนายค่าความหวานใช้ตัวแปรเดือนที่รับวัตถุดิบ (ตัวแปรด้านฤดูกาล) และระดับสีของสับปะรด (ระดับความสุก) เป็นตัวแปรนำเข้าเท่านั้น

ค่าทำนายที่ได้สามารถทำนายได้จากแบบจำลอง MC (Month/Color) เท่านั้น ซึ่งได้จากการทำนายโดยวิธีแบบจำลองการถดถอย (MR) แบบจำลองเครือข่ายประสาทเทียม (ANN) แบบจำลอง CART และแบบจำลอง CHAID ดังนั้นในการทำนายค่าของชุดข้อมูลรายละเอียดของสับปะรดรับเข้าแต่ละชุดข้อมูลจะทำนายค่าความหวานได้ทั้งหมด 4 ค่า ภาพที่ 23 แสดงการเปรียบเทียบความคลาดเคลื่อน RMSE และ MAE โดยรวมจากแบบจำลองต่างๆที่สามารถทำนายได้ของชุดข้อมูลที่มีระดับความน่าเชื่อถือในการทำนายค่าความหวานในระดับต่ำมาก แบบจำลองที่ดีที่สุด

ที่ให้ค่า RMSE และ MAE ต่ำที่สุด คือแบบจำลอง CART ที่กำหนดตัวแปรนำเข้า เป็นแบบจำลอง MC (Month/Color) ได้แก่ เดือนที่รับวัตถุดิบ และเกรดสีวัตถุดิบ



(ก) RMSE

(ข) MAE

ภาพที่ 23 ค่าความคลาดเคลื่อน RMSE และ MAE ของชุดข้อมูลที่ระดับต่ำมาก

2.2 การสร้างตารางเก็บข้อมูลที่ใช้ในฐานข้อมูลการทำนายค่าความหวาน ประกอบด้วย 3 ตาราง ได้แก่ ตารางค่าคุณลักษณะวัตถุดิบจริง (จากห้องปฏิบัติการ) ในช่วงปีพ.ศ. 2550 - 2551 ตารางค่าทำนายความหวานที่ได้จากแบบจำลองต่าง ๆ และตารางรายละเอียดระดับความน่าเชื่อถือในการทำนายค่า

2.2.1 ตารางค่าความหวานจริงในอดีต (Historical Brix Table) เป็นตารางแสดงรายละเอียดของข้อมูลจริงที่ได้จากห้องปฏิบัติการ โรงงาน ซึ่งใช้เป็นข้อมูลในการสร้างแบบจำลองการทำนายค่าทางสถิติ (ตารางที่ 10)

ตารางที่ 10 รายละเอียดของตารางค่าความหวานจริงในอดีต

ชื่อเขตข้อมูล	คำอธิบาย	ชนิดของข้อมูล	ค่าที่เป็นไปได้ทั้งหมด
ID (PK)	หมายเลขของชุดข้อมูล	ตัวเลขอัตโนมัติ (Auto Number)	จำนวนเต็ม
Month	เดือนที่รับวัตถุดิบ	ข้อความ (Text)	ชื่อเดือนในภาษาไทย
Amphoe	อำเภอที่ปลูก	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Province	จังหวัดที่ปลูก	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Quota	ผู้รวบรวมวัตถุดิบ	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Color	เกรดสีวัตถุดิบ	ตัวเลข (Number)	2 , 3 , 4 , 5
Date	วันเดือนปีที่รับวัตถุดิบในอดีต	วันที่/เวลา (Date/Time)	
Brix	ค่าความหวานวัตถุดิบ	ตัวเลข (Number)	ทศนิยม 1 ตำแหน่ง
%Acid	ค่าความเป็นกรดวัตถุดิบ	ตัวเลข (Number)	ทศนิยม 2 ตำแหน่ง
pH	ค่าพีเอชวัตถุดิบ	ตัวเลข (Number)	ทศนิยม 2 ตำแหน่ง

2.2.2 ตารางค่าทำนายความหวานวัตถุดิบ (Predicted Brix Table) เป็นตารางแสดงรายละเอียดของค่าทำนายจากแบบจำลองต่าง ๆ และค่าทางสถิติต่างๆ ของค่าทำนายที่ได้ เพื่อเป็นประโยชน์ต่อการนำค่าไปใช้ในกระบวนการผลิตต่อไป (ตารางที่ 11)

ตารางที่ 11 รายละเอียดของตารางค่าทำนายความหวานวัตถุคิบ

ชื่อเขตข้อมูล	คำอธิบาย	ชนิดของข้อมูล	ค่าที่เป็นไปได้ทั้งหมด
Month (CK)	เดือนที่รับวัตถุคิบ	ข้อความ (Text)	ชื่อเดือนในภาษาไทย
Amphoe (CK)	อำเภอที่ปลูก	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Quota (CK)	ผู้รวบรวมวัตถุคิบ	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Color (CK)	เกรดสีวัตถุคิบ	ตัวเลข (Number)	2 , 3 , 4 , 5
Province	จังหวัดที่ปลูก	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Level	ระดับความน่าเชื่อถือ	ข้อความ (Text)	ระดับสูงมาก สูง ปานกลาง และต่ำ
MR_MC	ค่าทำนายโดยวิธี MR จากแบบจำลอง MC	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
ANN_MC	ค่าทำนายโดยวิธี ANN จากแบบจำลอง MC	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CART_MC	ค่าทำนายโดยวิธี CART จากแบบจำลอง MC	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CHAID_MC	ค่าทำนายโดยวิธี CHAID จากแบบจำลอง MC	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
MR_A	ค่าทำนายโดยวิธี MR จากแบบจำลอง A	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
ANN_A	ค่าทำนายโดยวิธี ANN จากแบบจำลอง A	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CART_A	ค่าทำนายโดยวิธี CART จากแบบจำลอง A	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CHAID_A	ค่าทำนายโดยวิธี CHAID จากแบบจำลอง A	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง

ตารางที่ 11 (ต่อ)

ชื่อเขตข้อมูล	คำอธิบาย	ชนิดของข้อมูล	ค่าที่เป็นไปได้ทั้งหมด
MR_Q	ค่าทำนายโดยวิธี MR จากแบบจำลอง Q	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
ANN_Q	ค่าทำนายโดยวิธี ANN จากแบบจำลอง Q	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CART_Q	ค่าทำนายโดยวิธี CART จากแบบจำลอง Q	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CHAID_Q	ค่าทำนายโดยวิธี CHAID จากแบบจำลอง Q	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
MR_AQ	ค่าทำนายโดยวิธี MR จากแบบจำลอง AQ	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
ANN_AQ	ค่าทำนายโดยวิธี ANN จากแบบจำลอง AQ	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CART_AQ	ค่าทำนายโดยวิธี CART จากแบบจำลอง AQ	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CHAID_AQ	ค่าทำนายโดยวิธี CHAID จากแบบจำลอง AQ	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
MR_S	ค่าทำนายโดยวิธี MR จากแบบจำลอง S	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
ANN_S	ค่าทำนายโดยวิธี ANN จากแบบจำลอง S	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CART_S	ค่าทำนายโดยวิธี CART จากแบบจำลอง S	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
CHAID_S	ค่าทำนายโดยวิธี CHAID จากแบบจำลอง S	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
Average	ค่าเฉลี่ยค่าทำนายที่สามารถ ทำนายได้	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง

ตารางที่ 11 (ต่อ)

ชื่อเขตข้อมูล	คำอธิบาย	ชนิดของข้อมูล	ค่าที่เป็นไปได้ทั้งหมด
Min	ค่าน้อยสุดของค่าทำนายที่สามารถทำนายได้	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
Max	ค่ามากที่สุดของค่าทำนายที่สามารถทำนายได้	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
Best Method	แบบจำลองที่ดีที่สุด	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร
Best Predicted	ค่าทำนายจากแบบจำลองที่ดีที่สุด	ตัวเลข (Number)	ตัวเลขทศนิยม 1 ตำแหน่ง
MAE	ค่า MAE ของแบบจำลองที่ดีที่สุด	ตัวเลข (Number)	ตัวเลขทศนิยม 2 ตำแหน่ง
MAPE	ค่า MAPE ของแบบจำลองที่ดีที่สุด	ตัวเลข (Number)	ตัวเลขทศนิยม 2 ตำแหน่ง

2.2.3 ตารางระดับความน่าเชื่อถือในการทำนายค่า (Reliability Level Table) เป็นตารางอธิบายความหมายของระดับความน่าเชื่อถือต่างๆของชุดข้อมูลที่ต้องการทำนายค่า ผู้ใช้จำเป็นต้องพิจารณาประกอบการตัดสินใจก่อนนำค่าทำนายที่ได้ไปใช้ในกระบวนการผลิตต่อไป

ตารางที่ 12 รายละเอียดตารางระดับความน่าเชื่อถือในการทำนายค่า

ชื่อเขตข้อมูล	คำอธิบาย	ชนิดของข้อมูล	ค่าที่เป็นไปได้ทั้งหมด
ID_level (PK)	หมายเลขระดับความน่าเชื่อถือ	ตัวเลข (Number)	จำนวนเต็ม
Level	ระดับความน่าเชื่อถือ	ข้อความ (Text)	ระดับสูงมาก สูง ปานกลาง และต่ำ
Detail	รายละเอียดของระดับความน่าเชื่อถือ	ข้อความ (Text)	ข้อความใดๆที่มีความยาวไม่เกิน 255 ตัวอักษร

2.3 การออกแบบและสร้างฟอร์มในฐานข้อมูลโดยใช้โปรแกรม MS Access 2007

2.3.1 ฟอร์มหลัก (Main Form) เป็นหน้าหลักที่ปรากฏ เมื่อผู้ใช้ทำการเปิดฐานข้อมูล การทำนายค่าความหวานวัตถุดิบ จะมีส่วนให้ผู้ใช้ใส่รายละเอียดของวัตถุดิบสับปะรดที่ต้องการทำนายค่าความหวาน ได้แก่ เดือนที่รับวัตถุดิบ อำเภอที่ปลูก ผู้รวบรวมวัตถุดิบ และระดับสีสับปะรด ดังแสดงในภาพที่ 24 ในการใส่ข้อมูล ผู้ใช้สามารถกดเลือกข้อมูลที่มี โดยการกดลูกศรลงทางด้านขวาของแต่ละตัวแปรหรือจะพิมพ์ข้อมูลที่ต้องการลงไปในช่วงว่างก็ได้ เมื่อใส่รายละเอียดของชุดข้อมูลวัตถุดิบที่ต้องการทำนายค่าความหวานครบทุกตัวแปรแล้ว จะปรากฏค่าทำนายในฟอร์มย่อยต่างๆ

ภาพที่ 24 หน้าหลักการใช้งานฐานข้อมูลการทำนายค่าความหวาน

2.3.2 ฟอร์มย่อยแสดงจังหวัดที่ปลูกวัตถุดิบ (Subform_Show Province) แสดงจังหวัดที่ปลูกโดยอัตโนมัติ เมื่อผู้ใช้ทำการเลือกอำเภอที่ปลูก เพื่อความสะดวกในการใช้งาน และแสดงข้อมูลให้ผู้ใช้มั่นใจว่าแหล่งปลูกที่เลือกนั้นถูกต้อง ดังแสดงในภาพที่ 25

Brix Prediction Model

กรุณาใส่ข้อมูลสับปะรดรับเข้า

เดือนที่รับวัตถุดิบ มกราคม

อำเภอที่รับวัตถุดิบ ปลวกแดง

จังหวัดที่รับวัตถุดิบ ระยอง

ผู้ส่งมอบ

ระดับสี

Predict
Reset

ภาพที่ 25 ช่องจังหวัดที่ปลูกแสดงโดยอัตโนมัติ เมื่อผู้ใช้เลือกอำเภอที่ปลูก

2.3.3 ฟอรม์ย่อยค่าทำนาย (Subform_Predicted Brix) แสดงค่าทำนายความหวานทั้งหมดที่สามารถทำนายได้ ตามรายละเอียดของชุดข้อมูลที่ถูกเลือกในฟอร์มหลัก ดังแสดงในภาพที่ 26 - 30

ระดับความน่าเชื่อถือในการทำนายค่าบrix

Very High

รายละเอียดของชุดข้อมูลที่ทำนาย

ตัวแปรนำเข้าทุกตัวแปรตรงกับชุดข้อมูลที่สร้างแบบจำลอง

ค่าทำนายที่ดีที่สุด

14.0 brix

จากแบบจำลอง : ANN_AQ

ช่วงค่าทำนายที่ได้จากทุกแบบจำลอง :
ค่าต่ำสุด-สูงสุด 13.5 - 14.6 brix

ค่าบrixเฉลี่ยจากค่าทำนายที่ได้จากทุกแบบจำลอง :
14.0 brix

ค่าความคลาดเคลื่อนแบบจำลองที่ดีที่สุด
Mean Absolute Error : 1.01
Mean Absolute Percentage Error : 8.38 %

รายละเอียดค่าทำนายค่าความหวานของทุกแบบจำลอง

MR_MC :	14.2	ANN_MC :	13.8	CART_MC :	13.6	CHAID_MC :	13.6
MR_A :	14.2	ANN_A :	14.2	CART_A :	13.6	CHAID_A :	13.6
MR_Q :	14.2	ANN_Q :	14.6	CART_Q :	14.2	CHAID_Q :	14.1
MR_AQ :	14.2	ANN_AQ :	14.0	CART_AQ :	13.6	CHAID_AQ :	14.1
MR_S :	14.0	ANN_S :	14.2	CART_S :	14.2	CHAID_S :	13.5

ชุดข้อมูล : เดือน มกราคม จังหวัด ระยอง อำเภอ ปลวกแดง

ผู้ส่งมอบ ขัยยงค์ ระดับสี 2

ภาพที่ 26 ตัวอย่างการแสดงผลค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับสูงมาก

ระดับความน่าเชื่อถือในการทำนายค่าบrix

High

รายละเอียดของชุดข้อมูลที่ทำนาย

ตัวแปรแหล่งรับวัตถุดิบ (Source) ตรงกับชุดข้อมูลที่สร้างแบบจำลอง

ช่วงค่าทำนายที่ได้จากทุกแบบจำลอง :
ค่าต่ำสุด-สูงสุด

14.2 — 16.5 brix

ค่าบrixเฉลี่ยจากค่าทำนายที่ได้จากทุกแบบจำลอง :

15.2 brix

ค่าทำนายที่ดีที่สุด

14.2 brix

จากความคลาดเคลื่อนแบบจำลองที่ดีที่สุด

Mean Absolute Error : 1.01

Mean Absolute Percentage Error : 8.38 %

จากแบบจำลอง : ANN_AQ

รายละเอียดค่าทำนายค่าความหวานของทุกแบบจำลอง

MR_MC :	15.3	ANN_MC :	15.5	CART_MC :	15.5	CHAID_MC :	15.3
MR_A :	15.3	ANN_A :	16.1	CART_A :	15.1	CHAID_A :	16.5
MR_Q :	15.4	ANN_Q :	14.2	CART_Q :	14.8	CHAID_Q :	14.8
MR_AQ :	14.2	ANN_AQ :	14.2	CART_AQ :	14.4	CHAID_AQ :	16.5
MR_S :	15.3	ANN_S :	14.7	CART_S :	15.0	CHAID_S :	15.4

ชุดข้อมูล : เดือน พฤษภาคม จังหวัด กาญจนบุรี อำเภอ ปอพลอย

ผู้ส่งมอบ รุ่งนภา ระดับสี 2

ภาพที่ 27 ตัวอย่างการแสดงผลการทำนายความหวานที่ได้จากชุดข้อมูลในระดับสูง

ระดับความน่าเชื่อถือในการทำนายค่าบrix

Medium

รายละเอียดของชุดข้อมูลที่ทำนาย

ตัวแปรอำเภอ/ผู้ส่งมอบตรงกับชุดข้อมูลที่สร้างแบบจำลอง

ช่วงค่าทำนายที่ได้จากทุกแบบจำลอง :
ค่าต่ำสุด-สูงสุด

15.3 — 16.5 brix

ค่าบrixเฉลี่ยจากค่าทำนายที่ได้จากทุกแบบจำลอง :

15.6 brix

ค่าทำนายที่ดีที่สุด

15.5 brix

จากความคลาดเคลื่อนแบบจำลองที่ดีที่สุด

Mean Absolute Error : 1.03

Mean Absolute Percentage Error : 8.54 %

จากแบบจำลอง : ANN_A

รายละเอียดค่าทำนายค่าความหวานของทุกแบบจำลอง

MR_MC :	15.3	ANN_MC :	15.5	CART_MC :	15.5	CHAID_MC :	15.3
MR_A :	15.3	ANN_A :	15.5	CART_A :	16.0	CHAID_A :	16.5
MR_Q :		ANN_Q :		CART_Q :		CHAID_Q :	
MR_AQ :		ANN_AQ :		CART_AQ :		CHAID_AQ :	
MR_S :		ANN_S :		CART_S :		CHAID_S :	

ชุดข้อมูล : เดือน พฤษภาคม จังหวัด พิษณุโลก อำเภอ นครไทย

ผู้ส่งมอบ ไสว ระดับสี 2

ภาพที่ 28 ตัวอย่างการแสดงผลการทำนายความหวานที่ได้จากชุดข้อมูลในระดับปานกลาง

ระดับความน่าเชื่อถือในการทำนายค่าบrix

Low

รายละเอียดของชุดข้อมูลที่ทำนาย

เป็นผู้ส่งมอบที่รับวัตถุดิบเข้าใหม่

ช่วงค่าทำนายที่ได้จากทุกแบบจำลอง :
ค่าต่ำสุด-สูงสุด

14.3 – 16.0 brix

ค่าบrixเฉลี่ยจากค่าทำนายที่ได้จากทุกแบบจำลอง :

14.8 brix

ค่าทำนายที่ดีที่สุด

14.7 brix

จากแบบจำลอง : ANN_A

ค่าความคลาดเคลื่อนแบบจำลองที่ดีที่สุด

Mean Absolute Error : 1.03

Mean Absolute Percentage Error : 8.54 %

รายละเอียดค่าทำนายค่าความหวานของทุกแบบจำลอง

MR_MC :	14.7	ANN_MC :	14.3	CART_MC :	14.8	CHAID_MC :	14.8
MR_A :	14.7	ANN_A :	14.7	CART_A :	16.0	CHAID_A :	14.8
MR_Q :		ANN_Q :		CART_Q :		CHAID_Q :	
MR_AQ :		ANN_AQ :		CART_AQ :		CHAID_AQ :	
MR_S :		ANN_S :		CART_S :		CHAID_S :	

ชุดข้อมูล : เดือน มีนาคม จังหวัด ระยอง อำเภอ เขาสงเมษา

ผู้ส่งมอบ อื่นๆ ระดับสี 2

ภาพที่ 29 ตัวอย่างการแสดงผลค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับต่ำ

ระดับความน่าเชื่อถือในการทำนายค่าบrix

Very Low

รายละเอียดของชุดข้อมูลที่ทำนาย

เป็นอำเภอและผู้ส่งมอบที่รับวัตถุดิบเข้าใหม่

ช่วงค่าทำนายที่ได้จากทุกแบบจำลอง :
ค่าต่ำสุด-สูงสุด

14.3 – 14.8 brix

ค่าบrixเฉลี่ยจากค่าทำนายที่ได้จากทุกแบบจำลอง :

14.6 brix

ค่าทำนายที่ดีที่สุด

14.8 brix

จากแบบจำลอง : CART_MC

ค่าความคลาดเคลื่อนแบบจำลองที่ดีที่สุด

Mean Absolute Error : 1.12

Mean Absolute Percentage Error : 9.20 %

รายละเอียดค่าทำนายค่าความหวานของทุกแบบจำลอง

MR_MC :	14.7	ANN_MC :	14.3	CART_MC :	14.8	CHAID_MC :	14.8
MR_A :		ANN_A :		CART_A :		CHAID_A :	
MR_Q :		ANN_Q :		CART_Q :		CHAID_Q :	
MR_AQ :		ANN_AQ :		CART_AQ :		CHAID_AQ :	
MR_S :		ANN_S :		CART_S :		CHAID_S :	

ชุดข้อมูล : เดือน มีนาคม จังหวัด อื่นๆ อำเภอ อื่นๆ

ผู้ส่งมอบ อื่นๆ ระดับสี 2

ภาพที่ 30 ตัวอย่างการแสดงผลค่าทำนายความหวานที่ได้จากชุดข้อมูลในระดับต่ำมาก

2.3.4 ฟอรม์ย่อยค่าจริงในอดีต (Subform_Historical Brix) แสดงคุณลักษณะของวัตถุดิบที่เป็นข้อมูลจริงในอดีต ซึ่งใช้เป็นข้อมูลในการสร้างแบบจำลองการทำนายค่าตามรายละเอียดของชุดข้อมูลที่เลือกในฟอรม์หลัก โดยแสดงค่าเฉพาะชุดข้อมูลที่อยู่ในระดับความน่าเชื่อถือสูงเท่านั้น (ภาพที่ 31)

Historical Data (ข้อมูลในช่วงปี 2550-2551)							
Date	Brix	%Acid	pH	Month	Aumphoe	Quota	Color
01/08/50	15.6	0.24	4.65	มกราคม	ปลวกแดง	ชัยยงค์	2
01/08/50	15.0	0.36	3.35	มกราคม	ปลวกแดง	ชัยยงค์	2
01/04/50	13.8	0.48	3.75	มกราคม	ปลวกแดง	ชัยยงค์	2
01/14/51	13.0	0.49	3.50	มกราคม	ปลวกแดง	ชัยยงค์	2
01/14/51	12.0	0.61	3.50	มกราคม	ปลวกแดง	ชัยยงค์	2
01/04/50	13.4	0.32	3.75	มกราคม	ปลวกแดง	ชัยยงค์	2

ภาพที่ 31 ตัวอย่างการแสดงค่าความหวานจริงในอดีตที่ใช้ในการสร้างแบบจำลอง

2.3.5 ฟอรม์ย่อยคำนวณค่าความหวานสารบรรจุ (Subform_Calculate PM) ใช้ในการคำนวณค่าความหวานของสารบรรจุ (Packing Media) ที่ควรเตรียม เมื่อได้ค่าความหวานวัตถุดิบจากการทำนายค่าทางคณิตศาสตร์และสถิติ ดังแสดงในภาพที่ 32 โดยในช่องค่าทำนายความหวาน (Predicted RM Brix) จะเป็นค่าแสดงอัตโนมัติจากแบบจำลองที่ดีที่สุด ทั้งนี้ผู้ใช้สามารถแก้ไขค่าในช่องนี้ใหม่ได้ โดยแก้ไขตัวเลขในช่องดังกล่าวได้ตามต้องการ

Calculate Packing Media Brix	
กรุณาใส่ข้อมูลที่ต้องการคำนวณ	
Predicted RM Brix	RM Weight
14.0 brix	g
Final Brix	Net Weight
brix	g
ค่าปริมาตร น้ำ PM ที่จะต้องเตรียม	
PM Brix	brix

ภาพที่ 32 ส่วนคำนวณค่าความหวานของสารบรรจุที่ควรเตรียมในกระบวนการผลิต

ในส่วนนี้ผู้ใช้จำเป็นต้องทราบข้อมูลน้ำหนักของวัตถุดิบที่จะใช้ (Raw Material Weight) ความหวานสุดท้ายของผลิตภัณฑ์ (Final Brix) และน้ำหนักสุทธิของผลิตภัณฑ์สับปะรดกระป๋อง (Net Weight) จากภาพที่ 32 เมื่อผู้ใช้ใส่ตัวเลขต่างๆในช่องทั้งหมดแล้ว ฐานข้อมูลจะแสดงค่าความหวานของสารบรรจุ (Packing Media Brix) ที่ต้องเตรียมในกระบวนการผลิตให้โดยอัตโนมัติ การคำนวณในส่วนนี้ใช้ทฤษฎีการเท่ากันของสารก่อนและหลังผสม ดังนี้

$$B_1W_1 + B_2W_2 = B_3W_3 \quad (18)$$

ดังนั้นจะได้

$$B_2 = \frac{B_3W_3 - B_1W_1}{W_2} = \frac{B_3W_3 - B_1W_1}{(W_3 - W_1)}$$

โดย B_1, W_1 คือ ค่าความหวาน (°Brix) และน้ำหนักของวัตถุดิบสับปะรด (กรัม) ตามลำดับ
 B_2, W_2 คือ ค่าความหวาน (°Brix) และน้ำหนักของสารบรรจุ (กรัม) ตามลำดับ
 B_3, W_3 คือ ค่าความหวาน (°Brix) และน้ำหนักของผลิตภัณฑ์สุดท้าย (กรัม) ตามลำดับ

สรุปและข้อเสนอแนะ

สรุป

จากการศึกษาเปรียบเทียบความสามารถในการทำนายค่าความหวานวัตถุดิบสับปะรดของโรงงานกรณีศึกษา โดยการคัดเลือกปัจจัยที่เกี่ยวข้องกับค่าความหวานวัตถุดิบ โดยใช้ข้อมูลในด้านฤดูกาล ระดับความสุก และแหล่งที่มาของวัตถุดิบเป็นตัวแปรอิสระ และเปรียบเทียบแบบจำลองการถดถอย แบบจำลองเครือข่ายประสาทเทียม แบบจำลองต้นไม้การตัดสินใจ Classification and Regression Trees (CART) และแบบจำลองต้นไม้การตัดสินใจ Chi-Square Automatic Iteration Detector (CHAID) พบว่าแบบจำลองที่ให้ค่าความคลาดเคลื่อนน้อยที่สุด คือ แบบจำลองเครือข่ายประสาทเทียม โดยปัจจัยที่ใช้เป็นตัวแปรอิสระ ได้แก่ เดือนที่รับวัตถุดิบ (ปัจจัยด้านฤดูกาล) ระดับเกรดสีวัตถุดิบ (ปัจจัยความสุกสับปะรด) อำเภอที่ปลูก และผู้รวบรวมวัตถุดิบ (ปัจจัยด้านแหล่งที่มาของวัตถุดิบ)

อย่างไรก็ตาม แบบจำลองที่สร้างขึ้นในงานวิจัยนี้สามารถนำไปใช้ในการทำนายค่าระดับความหวานได้ทั้งหมด เนื่องจากค่าความคลาดเคลื่อนในการทำนายจากแบบจำลองทุกแบบจำลองมีค่าไม่แตกต่างกันมากนัก ในการศึกษาครั้งนี้ผู้วิจัยได้ทำการสร้างระบบสนับสนุนการตัดสินใจ เพื่อการทำนายค่าระดับความหวานวัตถุดิบและการคำนวณค่าความหวานของสารบรรจุ เพื่อประโยชน์ในการใช้งานในกระบวนการผลิตของโรงงานกรณีศึกษา โดยแสดงทุกค่าทำนายที่ได้จากแบบจำลองทั้งหมด ทั้งนี้เพื่อให้ฐานข้อมูลดังกล่าวมีความสามารถในการทำนายค่าครอบคลุมถึงชุดข้อมูลที่นอกเหนือจากชุดข้อมูลที่เป็นข้อมูลซึ่งใช้ในการสร้างแบบจำลอง โดยผู้วิจัยได้ทำการแบ่งระดับความน่าเชื่อถือออกเป็นระดับต่างๆ และมีค่าทำนายความหวานจากแบบจำลองที่ให้ค่าความคลาดเคลื่อนต่ำที่สุด เพื่อเป็นข้อมูลประกอบการตัดสินใจในการคัดเลือกค่าทำนายที่เหมาะสม เพื่อนำไปใช้ในกระบวนการผลิตต่อไป

การสร้างระบบสนับสนุนการตัดสินใจ เพื่อทำนายค่าความหวานของวัตถุดิบสับปะรดและกำหนดคุณลักษณะของสารบรรจุที่ต้องเตรียมขึ้นล่วงหน้า โดยการสร้างแบบจำลองทางคณิตศาสตร์และสถิติต่างๆนี้สามารถช่วยลดการสูญเสียเวลาและทรัพยากรจากความล่าช้าในกระบวนการผลิต มีผลให้กระบวนการผลิตสับปะรดบรรจุกระป๋องเป็นไปอย่างต่อเนื่องมากขึ้น สามารถลดการรอคอยในระหว่างการผลิต และทำให้ผลผลิตของโรงงานแปรรูปสับปะรดกระป๋องสูงมากขึ้น

ข้อเสนอแนะ

เนื่องจากวัตถุดิบสับปะรดมีแหล่งวัตถุดิบที่หลากหลาย ตลอดจนมีการเพิ่มของแหล่งวัตถุดิบทุกปี จึงควรเก็บรวบรวมข้อมูลที่นำมาใช้ในการสร้างแบบจำลองให้มากขึ้นและควรมีการปรับข้อมูลให้ทันสมัยอยู่เสมอ เพื่อเพิ่มประสิทธิภาพในการทำนายให้มีค่าความถูกต้องสูงขึ้น



เอกสารและสิ่งอ้างอิง

กัลยา วาณิชขัณษา. 2549. การใช้ SPSS for Windows ในการวิเคราะห์ข้อมูล. ครั้งที่ 9.

สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

กรมวิชาการเกษตร กระทรวงเกษตรและสหกรณ์. 2543. เกษตรดีที่เหมาะสมสำหรับสับปะรด.

กระทรวงเกษตรและสหกรณ์, กรุงเทพฯ.

กล้าณรงค์ ศรีรอด. 2546. เทคโนโลยีของน้ำตาล เล่ม 1 คุณสมบัติและเทคโนโลยี. ครั้งที่ 1.

สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

เกตุอร ทองเครือ. 2536. การปลูกสับปะรด คำแนะนำที่ 37 กรมส่งเสริมการเกษตร. ครั้งที่ 7. โรง

พิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย, กรุงเทพฯ.

ทศพร รอดทอง. 2546. แบบจำลองทางเศรษฐมิติอุปสงค์การนำเข้าและอุปทานการส่งออก

สับปะรดกระป๋องของประเทศไทย. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

ทรงศิริ แด่สมบัติ. 2548. การวิเคราะห์การถดถอย. ครั้งที่ 3. สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์,

กรุงเทพฯ.

ธัญนันท์ ฤทธิมณี. 2549. การพยากรณ์คุณภาพของส้มสายน้ำผึ้งหลังการขนส่งทางรถบรรทุกโดย

ใช้โครงข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเชียงใหม่.

บุญเสริม กิจศิริกุล. 2545. อัลกอริทึมสำหรับการทำเหมืองข้อมูล. วิทยานิพนธ์ปริญญาโท,

จุฬาลงกรณ์มหาวิทยาลัย.

นนทนิ แสงโสภณ. 2548. อินไซด์ Access2003. โปรวิชั่น, กรุงเทพฯ.

สกวรัตน์ ใจพัฒนกร. 2550. การวิเคราะห์และออกแบบระบบสารสนเทศ. ครั้งที่ 1. สำนักพิมพ์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

สิริยาภรณ์ ไกรมาก. 2552. แบบจำลองการทำนายเพื่อประเมินคุณภาพของผลิตภัณฑ์กาแฟแก้ว และบด. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

สมเกียรติ เกตุเอี่ยม. 2548. เทคนิคการพยากรณ์. ครั้งที่ 2. การกิจเอกสารและตำรา กลุ่มงานบริการ การศึกษา มหาวิทยาลัยทักษิณ, สงขลา.

สุรัสวดี วงศ์จันทร์สุข และ สัจจะ จรัสรุ่งรวิธร. 2549. คู่มือใช้งาน Access 2003 ฉบับสมบูรณ์. ครั้งที่ 1. ไอดีซี, นนทบุรี.

สำนักส่งเสริมการค้าสินค้าเกษตร กรมการค้าภายใน กระทรวงพาณิชย์. 2552. สับปะรด. แหล่งที่มา:

http://agri.dit.go.th/web_dit_sec3/admin/uploadfiles/upload_files/pineapple1251.pdf, 3 มีนาคม 2552.

สำนักงานเศรษฐกิจอุตสาหกรรม กระทรวงอุตสาหกรรม. 2545. รายงานภาวะอุตสาหกรรมแปรรูปสับปะรด. แหล่งที่มา:

<http://www.oie.go.th/industrystatus2/37.pdf>, 20 ตุลาคม 2551.

อัมรินทร์ ศรีวนิชย์. 2548. การทำนายความชอบของเบียร์โดยใช้แบบจำลองความถดถอยและแบบจำลองเครือข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

Breiman, L., J. H. Friedman, R. A. Olshen and C. G. Stone. 1984. **Classification and Regression Trees**. Wadsworth International Group, California.

Berry, M.J.A. and G. S. Linoff. 2004. **Data Mining Techniques**. 2nd ed. Wiley Publishing Inc., Indiana.

Fausett, L. 1994. **Fundamentals of Neural Networks: Architectures, Algorithm and Applications.** Prentice Hall Inc., New Jersey.

Funahashi, K. 1989. On the approximate realization of continuous mappings by Neural Networks. **Neural Networks** 2 (3): 183-192.

Hornik, K., M. Stinchcombe and H. White. 1989. Multilayer feed forward networks are universal approximators. **Neural Networks** 2 (5): 359-366.

Hassoun, M. H. 1995. Fundamentals of Artificial Neural Networks. Asco Trade Typesetting Ltd., Massachusetts.

Kass, G. 1980. An exploratory technique for investigating large quantities of categorical data. **Applied Statistics** 29 (2): 119-127.

Kaul, M., R. L. Hill and C. Walthall. 2005. Artificial neural networks for corn and soybean yield prediction. **Agricultural Systems** 2005 (85): 1-18.

Kutner, H. M., C. J. Nachtsheim and J. Neter. 2005. **Applied Linear Statistical Methods.** The McGraw-Hill companies, New Jersey.

Mahjoobi, J. and A. Etemad-Shahidi. 2008. An alternative approach for the prediction of significant wave heights based on classification and regression trees. **Applied Ocean Research** 30: 172-177.

Mitchell, T. M. 1997. **Machine learning.** The McGraw-Hill Companies, New York.

Page, G. F., J. B. Gomm, and D. Willam. 1993. **Application of Neural Networks to Modeling and Control.** Chapman and Hall, London.

- Petitti, D. 2000. **Meta-analysis Decision analysis and Cost effectiveness analysis: Methods for Quantitative synthesis in Medicine.** 2nd ed. Oxford University Press, New York.
- Rai, P., G. C. Majumdar, S. DasGupta and S. De. 2005. Prediction of the viscosity of clarified fruit juice using artificial neural network: a combined effect of concentration and temperature. **Journal of Food Engineering** 2005 (68): 527-533.
- Rawlings, J. O., S. G. Pantula and D. A. Dickey. 1998. **Applied Regression Analysis: a Research Tool.** Springer, New York.
- Scheffer, J. 2002. Data mining in the survey setting: why do children go off the rails?. **Res. Lett. Inf. Math. Sci.** 3: 161-189.
- SPSS Inc Illinois: Integral Solutions Limited. 2007. **Clementine 12.0 Algorithms Guide.** Chicago.
- Thuraisingham, B. 1999. **Data mining: technologies, techniques, tools and trends.** CRC press, New York.
- Tso, G. K. F. and K. K. W. Yau. 2007. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. **Energy** 32: 1761-1768.
- Uno, Y., S. O. Prasher, R. Lacroix, P. K. Goel, Y. Karimi, A. Viau and R.M. Patel. 2005. Artificial neural networks to predict corn yield from compact airborne spectrographic imager data. **Computers and Electronics in Agriculture** 2005 (47): 149-161.
- Vayssieres, Marc P., Richard E. Plant and Barbara H. Allen-Diaz. 2000. Classification trees an alternative non-parametric approach for predicting species distributions. **Journal of Vegetation Science** 11 (5): 679-694.

Yang,C.C., S. O. Prasher, P. Enright, C. Madramootoo, M. Burgess, P.K. Goel and I. Callum.

2003. Application of decision tree technology for image classification using remote sensing data. **Agricultural Systems** 76: 1101-1117.







ภาคผนวก ก

รายละเอียดของตัวแปรอิสระที่เป็นข้อมูลที่ใช้สร้างแบบจำลองต่างๆ

ตารางผนวกที่ ก1 รายละเอียดอำเภอที่ปลูกซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการทำนายค่าความหวานวัดดัชนีสับปะรด

รหัสอำเภอ	อำเภอที่ปลูก	จังหวัดที่ปลูก
A01	เขาชะเมา	ระยอง
A02	ปลวกแดง	ระยอง
A03	วังจันทร์	ระยอง
A04	ปากท่อ	ราชบุรี
A05	เมืองชัยภูมิ	ชัยภูมิ
A06	นครไท	พิษณุโลก
A07	ด่านช้าง	สุพรรณ
A08	บ่อพลอย	กาญจนบุรี
A09	หนองปรือ	กาญจนบุรี
A10	หัวหิน	ประจวบคีรีขันธ์
A11	สามร้อยยอด	ประจวบคีรีขันธ์
A12	บ้านไร่	อุทัยธานี
A13	แม่สาย	เชียงราย
A14	แม่จัน	เชียงราย
A15	เขาสมิง	ตราด
A16	ห้วยมุ่น	อุดรดิตถ์
A17*	บ้านคา	ราชบุรี

หมายเหตุ * หมายถึง อำเภอที่ปลูกที่ใช้เป็นตัวแปรอ้างอิงในการกำหนดตัวแปรกลุ่ม

ตารางผนวกที่ ก2 รายละเอียดของผู้รวบรวมวัตถุซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการทำนาย
ค่าความหวานวัตถุดิบสับปะรด

รหัสผู้รวบรวม	ผู้รวบรวมวัตถุ
Q01	ผู้รวบรวมวัตถุที่ 1
Q02	ผู้รวบรวมวัตถุที่ 2
Q03	ผู้รวบรวมวัตถุที่ 3
Q04	ผู้รวบรวมวัตถุที่ 4
Q05	ผู้รวบรวมวัตถุที่ 5
Q06	ผู้รวบรวมวัตถุที่ 6
Q07	ผู้รวบรวมวัตถุที่ 7
Q08	ผู้รวบรวมวัตถุที่ 8
Q09	ผู้รวบรวมวัตถุที่ 9
Q10	ผู้รวบรวมวัตถุที่ 10
Q11	ผู้รวบรวมวัตถุที่ 11
Q12	ผู้รวบรวมวัตถุที่ 12
Q13	ผู้รวบรวมวัตถุที่ 13
Q14	ผู้รวบรวมวัตถุที่ 14
Q15	ผู้รวบรวมวัตถุที่ 15
Q16	ผู้รวบรวมวัตถุที่ 16
Q17	ผู้รวบรวมวัตถุที่ 17
Q18	ผู้รวบรวมวัตถุที่ 18
Q19	ผู้รวบรวมวัตถุที่ 19
Q20	ผู้รวบรวมวัตถุที่ 20
Q21	ผู้รวบรวมวัตถุที่ 21
Q22	ผู้รวบรวมวัตถุที่ 22
Q23	ผู้รวบรวมวัตถุที่ 23
Q24 *	ผู้รวบรวมวัตถุที่ 24

หมายเหตุ * หมายถึง ผู้รวบรวมวัตถุที่ใช้เป็นตัวแปรอ้างอิงในการกำหนดตัวแปรกลุ่ม

ตารางผนวกที่ ก3 รายละเอียดของแหล่งที่มาของวัตถุสืบซึ่งเป็นข้อมูลที่ใช้สร้างแบบจำลองการ
ทำนายค่าความหวาน

รหัสแหล่งที่มาวัตถุสืบ	อำเภอที่ปลูก	ผู้รวบรวมวัตถุสืบ
S01	เขาชะเมา	ผู้รวบรวมวัตถุสืบที่ 24
S02	ปลวกแดง	ผู้รวบรวมวัตถุสืบที่ 1
S03	ปลวกแดง	ผู้รวบรวมวัตถุสืบที่ 3
S04	ปลวกแดง	ผู้รวบรวมวัตถุสืบที่ 9
S05	วังจันทร์	ผู้รวบรวมวัตถุสืบที่ 1
S06	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 2
S07	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 24
S08	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 13
S09	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 14
S10	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 16
S11	บ้านคา	ผู้รวบรวมวัตถุสืบที่ 21
S12	ปากท่อ	ผู้รวบรวมวัตถุสืบที่ 5
S13	ปากท่อ	ผู้รวบรวมวัตถุสืบที่ 15
S14	เมืองชัยภูมิ	ผู้รวบรวมวัตถุสืบที่ 1
S15	เมืองชัยภูมิ	ผู้รวบรวมวัตถุสืบที่ 3
S16	เมืองชัยภูมิ	ผู้รวบรวมวัตถุสืบที่ 6
S17	เมืองชัยภูมิ	ผู้รวบรวมวัตถุสืบที่ 24
S18	เมืองชัยภูมิ	ผู้รวบรวมวัตถุสืบที่ 23
S19	นครไทย	ผู้รวบรวมวัตถุสืบที่ 1
S20	นครไทย	ผู้รวบรวมวัตถุสืบที่ 24
S21	ด่านช้าง	ผู้รวบรวมวัตถุสืบที่ 4
S22	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 1
S23	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 6
S24	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 7
S25	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 8
S26	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 24
S27	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 10
S28	บ่อพลอย	ผู้รวบรวมวัตถุสืบที่ 18

ตารางผนวกที่ ก3 (ต่อ)

รหัสแหล่งที่มาวัตถุ	อำเภอที่ปลูก	ผู้รวบรวมวัตถุ
S29	บ่อพลอย	ผู้รวบรวมวัตถุที่ 20
S30	หนองปรือ	ผู้รวบรวมวัตถุที่ 18
S31	ห้วยหิน	ผู้รวบรวมวัตถุที่ 5
S32	ห้วยหิน	ผู้รวบรวมวัตถุที่ 8
S33	ห้วยหิน	ผู้รวบรวมวัตถุที่ 11
S34	ห้วยหิน	ผู้รวบรวมวัตถุที่ 12
S35	ห้วยหิน	ผู้รวบรวมวัตถุที่ 22
S36	สามร้อยยอด	ผู้รวบรวมวัตถุที่ 24
S37	สามร้อยยอด	ผู้รวบรวมวัตถุที่ 9
S38	สามร้อยยอด	ผู้รวบรวมวัตถุที่ 13
S39	บ้านไร่	ผู้รวบรวมวัตถุที่ 1
S40	บ้านไร่	ผู้รวบรวมวัตถุที่ 24
S41	แม่สาย	ผู้รวบรวมวัตถุที่ 6
S42	แม่สาย	ผู้รวบรวมวัตถุที่ 24
S43	แม่จัน	ผู้รวบรวมวัตถุที่ 6
S44	แม่จัน	ผู้รวบรวมวัตถุที่ 24
S45	เขาสมิง	ผู้รวบรวมวัตถุที่ 9
S46	เขาสมิง	ผู้รวบรวมวัตถุที่ 19
S47	ห้วยมุ่น	ผู้รวบรวมวัตถุที่ 24
S48 *	บ้านคา	ผู้รวบรวมวัตถุที่ 17

หมายเหตุ * หมายถึง แหล่งที่มาของวัตถุที่ใช้เป็นตัวแปรอ้างอิงในการกำหนดตัวแปรกลุ่ม

ตารางผนวกที่ ก4 ตัวแปรหุ่นเดือนที่รับวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม

รหัสเดือนที่รับวัตถุดิบ	m1	m2	m3	m4
Jan	0	1	1	0
Feb	0	1	0	0
Mar	0	0	1	0
Apr	0	1	1	1
May	0	1	0	1
Jun	1	0	1	1
Jul	1	0	0	1
Aug	1	1	1	0
Oct	1	1	0	0
Nov	1	0	1	0
Dec	1	0	0	0

ตารางผนวกที่ ก5 ตัวแปรหุ่นอำเภอที่ปลูกวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม

รหัสอำเภอ	a1	a2	a3	a4	a5
A1	0	0	0	0	1
A2	0	0	0	1	1
A3	0	0	1	0	1
A4	0	1	0	0	1
A5	0	1	0	1	1
A6	0	1	1	0	1
A7	0	1	1	1	1
A8	1	0	0	0	0
A9	1	0	0	1	0
A10	1	0	1	0	0
A11	1	0	1	1	0
A12	1	1	0	0	0
A13	1	1	0	1	0
A14	1	1	1	0	0
A15	1	1	1	1	0
A16	0	0	0	1	0
A17	0	0	1	1	1

ตารางผนวกที่ ก6 ตัวแปรหุ่นผู้รวบรวมวัตถุดิบในแบบจำลองเครือข่ายประสาทเทียม

รหัสผู้รวบรวมวัตถุดิบ	q1	q2	q3	q4	q5
Q1	0	0	1	1	0
Q2	0	0	1	1	1
Q3	0	0	1	0	0
Q4	0	0	0	0	1
Q5	0	1	0	1	0
Q6	0	1	0	1	1
Q7	1	1	1	0	0
Q8	1	1	1	0	1
Q9	1	0	1	1	0
Q10	1	0	0	1	1
Q11	1	0	0	0	0
Q12	1	0	0	0	1
Q13	0	1	1	1	0
Q14	0	1	1	1	1
Q15	0	1	1	0	0
Q16	0	1	0	0	1
Q17	0	0	0	1	0
Q18	0	0	0	1	1
Q19	1	0	1	0	0
Q20	1	0	1	0	1
Q21	1	1	1	1	0
Q22	1	1	0	1	1
Q23	1	1	0	0	0
Q24	1	1	0	0	1

ตารางผนวกที่ ก7 ตัวแปรหุ่นแหล่งที่มาของวัตถุคิบบนแบบจำลองเครือข่ายประสาทเทียม

รหัสแหล่งที่มาวัตถุคิบบ	s1	s2	s3	s4	s5	s6
S1	1	1	0	0	1	1
S2	1	1	0	0	1	0
S3	1	1	0	0	0	1
S4	1	1	0	1	0	0
S5	1	1	0	1	1	1
S6	1	1	0	1	1	0
S7	1	1	1	0	0	1
S8	1	1	1	0	0	0
S9	1	1	1	0	1	1
S10	1	1	1	1	1	0
S11	1	1	1	1	0	1
S12	1	1	1	1	0	0
S13	1	0	0	0	1	1
S14	1	0	0	0	1	0
S15	1	0	0	0	0	1
S16	1	0	0	1	0	0
S17	1	0	0	1	1	1
S18	1	0	0	1	1	0
S19	1	0	1	0	0	1
S20	1	0	1	0	0	0
S21	1	0	1	0	1	1
S22	1	0	1	1	1	0
S23	1	0	1	1	0	1
S24	1	0	1	1	0	0
S25	0	1	0	0	1	1
S26	0	1	0	0	1	0
S27	0	1	0	0	0	1
S28	0	1	0	1	0	0
S29	0	1	0	1	1	1

ตารางผนวกที่ ก7 (ต่อ)

รหัสแหล่งที่มาวัตถุสืบ	s1	s2	s3	s4	s5	s6
S30	0	1	0	1	1	0
S31	0	1	1	0	0	1
S32	0	1	1	0	0	0
S33	0	1	1	0	1	1
S34	0	1	1	1	1	0
S35	0	1	1	1	0	1
S36	0	1	1	1	0	0
S37	0	0	0	0	1	1
S38	0	0	0	0	1	0
S39	0	0	0	0	0	1
S40	0	0	0	1	0	0
S41	0	0	0	1	1	1
S42	0	0	0	1	1	0
S43	0	0	1	0	0	1
S44	0	0	1	0	0	0
S45	0	0	1	0	1	1
S46	0	0	1	1	1	0
S47	0	0	1	1	0	1
S48	0	0	1	1	0	0



ภาคผนวก ข

การแบ่งกลุ่มข้อมูลในการทำนายค่าโดยวิธีต้นไม้การตัดสินใจในแบบจำลองต่างๆ

ตารางผนวกที่ ข1 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง MC (Month/Color) โดยวิธี CART

กลุ่ม	ระดับเกรดสี	เดือนที่รับวัตถุดิบ	ค่าความหวาน ($^{\circ}$ Brix)
1	2	May, Jun, Aug	15.49
2	3	May, Jun, Aug	15.13
3	2	Mar, Apr, Jul	14.76
4	3	Mar, Apr, Jul	14.20
5	2, 3	Feb, Dec	14.05
6	4	May	13.99
7	2, 3	Jan	13.64
8	2	Oct, Nov	13.58
9	4	Jun	13.37
10	4	Mar, Apr, Jul	12.99
11	3	Oct, Nov	12.75
12	4	Feb, Aug, Dec	12.51
13	4	Jan, Oct	12.20
14	4	Nov	11.78
15	5	Apr, May	10.29
16	5	Mar	10.20
17	5	Feb, Dec	9.96
18	5	Jan	9.87
19	5	Jul	9.82
20	5	Jun, Aug	9.71
21	5	Nov	9.21
22	5	Oct	9.10

ตารางผนวกที่ ข2 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง A (Month/Color/Amphoe)
โดยวิธี CART

กลุ่ม	ระดับ เกรดสี่	เดือนที่รับวัตถุดิบ	อำเภอที่ปลูก	ค่าความหวาน (°Brix)
1	2, 3	Mar, Apr, May, Jun, Jul, Aug	A1, A6, A9, A14, A16	16.03
2	2, 3	May, Jun, Aug	A2, A3, A4, A5, A7, A8, A10, A11, A12, A13, A15, A17	15.05
3	2, 3	Mar, Apr, Jul	A2, A3, A4, A5, A7, A8, A10, A11, A12, A13, A15, A17	14.37
4	2, 3	Jan, Feb, Oct, Nov, Dec	-	13.58
5	4	Mar, Apr, May, Jun, Jul	-	13.22
6	4	Jan, Feb, Aug, Oct, Nov, Dec	-	12.18
7	5	Mar, Apr, May, Jul, Aug	A4, A5, A7, A8, A9, A10	10.76
8	5	Jan, Feb, Jun, Dec	A4, A7, A10, A14, A16, A17	10.17
9	5	Mar, Apr, May, Jul, Aug	A2, A6, A13, A14, A17	10.05
10	5	Jan, Feb, Jun, Dec	A2, A5, A6, A8, A12, A13	9.60
11	5	May, Jul, Aug, Dec	A1, A3, A11, A15	9.49
12	5	Oct, Nov	-	9.18
13	5	Jan, Feb, Mar, Apr, Jun	A1, A3, A11, A15	8.11

ตารางผนวกที่ ข3 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง Q (Month/Color/Quota)
โดยวิธี CART

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Jul, Aug	Q2, Q7, Q13, Q18, Q24	15.60
2	2, 3	Apr, May, Jun, Aug	Q1, Q3, Q4, Q5, Q6, Q8, Q9, Q12, Q14, Q15, Q16, Q17, Q19	14.80
3	2, 3	Mar, Apr	Q2, Q7, Q13, Q18, Q24	14.58
4	2, 3	Jan, Feb, Dec	Q1, Q2, Q3, Q4, Q5, Q6, Q9, Q10, Q17, Q18, Q19, Q24	14.17
5	2, 3	Mar, Jul	Q1, Q3, Q4, Q5, Q6, Q8, Q9, Q12, Q14, Q15, Q16, Q17, Q19	14.15
6	2	Oct, Nov	-	13.58
7	2, 3	Jan, Feb, Dec	Q7, Q8, Q11, Q12, Q13, Q14, Q15, Q21, Q22, Q23	13.23
8	4	Mar, Apr, May, Jun, Jul	-	13.22
9	4	Jan, Feb, Aug, Dec	Q1, Q2, Q3, Q4, Q5, Q7, Q8, Q9, Q12, Q17, Q18, Q19, Q20, Q22, Q23, Q24	12.81
10	3	Oct, Nov	-	12.75
11	4	Oct, Nov	Q1, Q2, Q3, Q4, Q5, Q7, Q8, Q9, Q12, Q17, Q18, Q19, Q20, Q22, Q23, Q24	12.03
12	4	Jan, Feb, Aug, Oct, Nov, Dec	Q6	11.86
13	4	Jan, Feb, Aug, Oct, Nov, Dec	Q10, Q11, Q13, Q14, Q15, Q21	11.34
14	5	Jan, Feb, Mar, Apr, May, Jul, Oct	Q4, Q10, Q12, Q15, Q18, Q21	11.00
15	5	Mar, Apr, May, Jul, Oct	Q5, Q6, Q8, Q11, Q13, Q17	10.69
16	5	Feb, Mar, May, Aug	Q1, Q7, Q9, Q14, Q16, Q24	10.28
17	5	Jun, Nov, Dec	Q13, Q17	10.28
18	5	Jan, Feb	Q5, Q6, Q8, Q11, Q13, Q17	9.88

ตารางผนวกที่ ข3 (ต่อ)

กลุ่ม	ระดับ เกรดสี่	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
19	5	Apr, Jun, Sep, Dec	Q1, Q7, Q9, Q14, Q16, Q24	9.62
20	5	Jan, Oct, Nov	Q1, Q3, Q9, Q14, Q19, Q22, Q23	9.31
21	5	Feb, Mar, Apr, May, Jun, Jul, Aug, Dec	Q2, Q3, Q19	9.18
22	5	Jun, Nov, Dec	Q5, Q6, Q15, Q18	8.86
23	5	Jan, Oct, Nov	Q2, Q7, Q20, Q24	8.75

ตารางผนวกที่ ๔4 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง AQ (Month/Color/Amphoe/Quota)
โดยวิธี CART

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	อำเภอที่ปลูก	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	Mar, Apr, May, Jun, Jul, Aug	A1, A6, A9, A14, A16	-	16.03
2	2, 3	Mar, Apr, May, Jun, Jul, Aug	A2, A3, A4, A5, A7, A8, A10, A11, A12, A13, A15, A17	Q2, Q7, Q13, Q18	15.47
3	2, 3	Mar, Apr, May, Jun, Jul, Aug	A2, A3, A4, A5, A7, A8, A10, A11, A12, A13, A15, A17	Q1, Q3, Q4, Q5, Q6, Q8, Q9, Q12, Q14, Q15, Q16 Q17, Q19, Q24	14.42
4	2, 3	Jan, Feb, Oct, Nov, Dec	-	-	13.58
5	4	Mar, Apr, May, Jun, Jul	-	-	13.22
6	4	Jan, Feb, Aug, Oct, Nov, Dec	-	Q1, Q2, Q3, Q4, Q5, Q7, Q8, Q9, Q12, Q17, Q18, Q19, Q20, Q22, Q23, Q24, Q6	12.45
7	4	Jan, Feb, Aug, Oct, Nov, Dec	-	Q10, Q11, Q13, Q14, Q15, Q21	11.86
8	4	Jan, Feb, Aug, Oct, Nov, Dec	-	Q1, Q7	11.34
9	5	Feb, Mar, Apr, May, Jun, Jul, Aug, Dec	A2, A5, A6, A8, A11, A12, A13, A14, A16, A17	Q1, Q7	10.53
10	5	Jan, Feb, Mar, Apr, May, Jul, Oct	-	Q2, Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17 Q18	10.51
11	5	Jun, Nov, Dec	A14, A17	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	10.28

ตารางผนวกที่ ๔ (ต่อ)

กลุ่ม	ระดับ เกรดสี่	เดือนที่รับวัตถุดิบ	อำเภอที่ปลูก	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
12	5	Feb, Mar, Apr, May, Jun, Jul, Aug, Dec	A2, A5, A6, A8, A11, A12, A13, A14, A16, A17	Q2, Q3, Q9, Q14, Q16, Q24	9.78
13	5	Jan, Oct, Nov	-	Q1, Q2, Q3, Q7, Q9, Q14, Q16, Q19, Q20, Q22, Q23, Q24	9.16
14	5	Feb, Mar, Apr, May, Jun, Jul, Aug, Dec	A1, A3, A15	Q1, Q2, Q3, Q7, Q9, Q14, Q16, Q19, Q20, Q22, Q23, Q24	8.97
15	5	Jun, Nov, Dec	A4, A8, A10, A13	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	8.78

ตารางผนวกที่ ข5 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง S (Month/Color/Source)
โดยวิธี CART

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	แหล่งที่มาวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Aug	S3, S6, S20, S44, S47	17.14
2	2, 3	Mar, Apr, Jul	S1, S8, S16, S19, S24, S25, S28, S30, S45	15.59
3	2, 3	May, Jun, Aug	S1, S4, S5, S7, S8, S9, S10, S13, S14, S16, S19, S24, S36, S40, S41, S42, S43, S46, S48	14.96
4	2, 3	Jan, Feb, Dec	S1, S2, S3, S4, S6, S7, S8, S12, S14, S19, S21, S22, S23, S27, S28, S31, S40, S46, S48	14.19
5	2, 3	Mar, Apr, Jul	S2, S3, S4, S7, S9, S12, S13, S14, S15, S21, S23, S32, S34, S36, S46, S48	14.16
6	4	May	-	13.99
7	2, 3	Oct, Nov	-	13.17
8	2, 3	Jan, Feb, Dec	S9, S11, S13, S17, S18, S24, S26, S32, S33, S34, S35, S38	13.12
9	4	Mar, Apr, Jun, Jul,	-	13.09
10	4	Jan, Feb, Aug, Oct, Dec	S2, S3, S4, S6, S7, S18, S19, S21, S24, S26, S28, S31, S32, S34, S36, S40, S46, S48	12.91
11	4	Nov	-	11.78
12	4	Jan, Feb, Aug, Oct, Dec	S1, S8, S9, S11, S12, S13, S14, S17, S23, S27, S33, S35, S37, S38	11.62
13	5	Mar, Apr, May	S4, S8, S12, S14, S16, S24, S33	11.34
14	5	Jan, Feb, Dec	S4, S7, S11, S12, S13, S21, S26, S27, S32, S34, S46, S48	10.62
15	5	Mar, Apr, May	S2, S13, S21, S23, S34	10.54
16	5	Jun, Jul, Aug	S2, S3, S6, S7, S8, S9, S10, S48, S14, S20, S28, S30, S32, S36, S42, S43, S47	10.27

ตารางผนวกที่ ๕ (ต่อ)

กลุ่ม	ระดับ เกรด	เดือนที่รับวัตถุดิบ	แหล่งที่มาวัตถุดิบ	ค่าความหวาน (°Brix)
17	5	Oct, Nov	S3, S5, S19, S48	9.88
18	5	Mar, Apr, May	S1, S6, S7, S9, S15, S19, S25, S36, S42, S44, S46	9.42
19	5	Jan, Feb, Dec	S1, S2, S3, S6, S8, S9, S14, S17, S18, S19, S23, S24, S28, S31, S33, S35, S38, S40	9.31
20	5	Jun, Jul, Aug	S1, S4, S5, S41, S46	9.03
21	5	Oct, Nov	S2, S4, S6, S7, S9, S13, S14, S22, S23, S24, S29, S36, S39, S40, S46	8.88

ตารางผนวกที่ ข6 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง MC (Month/Color) โดยวิธี CHAID

กลุ่ม	ระดับเกรดดี	เดือนที่รับวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Aug	15.32
2	2	Mar, Apr, Jul	14.76
3	2	Feb, Dec	14.25
4	3	Mar, Apr, Jul	14.20
5	4	May	13.99
6	3	Feb, Dec	13.86
7	2, 3	Jan	13.64
8	2	Oct, Nov	13.58
9	4	Mar, Apr, Jun, Jul	13.09
10	3	Oct, Nov	12.75
11	4	Jan, Feb, Aug, Oct, Dec	12.35
12	4	Nov	11.78
13	5	Mar, Apr, May	10.26
14	5	Jan, Feb, Jun, Sep, Aug, Dec	9.86
15	5	Oct, Nov	9.18

ตารางผนวกที่ ข7 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง A (Month/Color/Amphoe)
โดยวิธี CHAID

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	อำเภอที่ปลูก	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Aug	A2, A6, A8, A11, A14, A16	16.53
2	2, 3	May, Jun, Aug	A1, A3, A4, A5, A12, A13, A15, A17	15.00
3	2, 3	Feb, Dec	A2, A6, A12, A15	14.92
4	2	Mar, Apr, Jul	-	14.76
5	2	Dec	A1, A5, A7, A11, A17	14.65
6	3	Mar, Apr, Jul	-	14.20
7	4	May	-	13.99
8	3	Dec	A1 A5 A7 A11 A17	13.90
9	2, 3	Jan	-	13.64
10	2, 3	Feb, Dec	A4, A8, A10	13.62
11	2	Oct, Nov	-	13.58
12	2, 3	Feb	A1, A5, A7, A11, A17	13.54
13	3	Oct, Nov	A3, A4, A5, A15, A17	13.32
14	4	Jan, Feb, Aug, Oct, Dec	A2, A7, A12, A15	13.09
15	4	Mar, Apr, Jun, Jul	-	13.09
16	3	Oct, Nov	A2, A6, A8, A11, A12	12.32
17	4	Jan, Feb, Aug, Oct, Dec	A5, A6, A8, A10, A17	12.29
18	4	Nov	-	11.78
19	4	Jan, Feb, Aug, Oct, Dec	A1, A4, A11	11.39
20	5	Mar, Apr, May	A2, A4, A5, A7, A8, A10	10.75
21	5	Jan, Feb, Jun, Jul, Aug, Dec	A4, A7, A9, A10, A14, A16, A17	10.23
22	5	Jan, Feb, Jun, Jul, Aug, Dec	A2, A5, A6, A8, A12, A13	9.67
23	5	Mar, Apr, May	A1, A6, A11, A13, A14, A15, A17	9.64
24	5	Oct, Nov	-	9.18
25	5	Jan, Feb, Jun, Jul, Aug, Dec	A1, A3, A11, A15	9.06

ตารางผนวกที่ ข8 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง Q (Month/Color/Quota)
โดยวิธี CHAID

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Aug	Q2, Q3, Q7, Q14, Q24	15.83
2	2, 3	May, Jun, Aug	Q1, Q6, Q9, Q13, Q15, Q16, Q17, Q19	14.83
3	2	Mar, Apr, Jul	-	14.76
4	2, 3	Jan	Q2, Q3, Q4, Q6	14.78
5	2, 3	Feb, Dec	Q1, Q3, Q4, Q9, Q18, Q19, Q24	14.62
6	2	Feb, Dec	Q5, Q10, Q17	14.37
7	3	Mar, Apr, Jul	-	14.20
8	2, 3	Jan	Q1, Q5	14.14
9	4	May	-	13.99
10	3	Feb, Dec	Q5, Q10, Q17	13.69
11	2	Oct, Nov	-	13.58
12	2, 3	Jan	Q9, Q13, Q15, Q17	13.54
13	2, 3	Feb, Dec	Q6, Q7, Q11, Q12, Q13, Q14, Q15	13.44
14	4	Mar, Apr, Jun, Jul	-	13.09
15	4	Jan, Feb, Aug, Oct, Dec	Q2, Q3, Q4, Q7, Q8, Q9, Q12, Q17, Q18, Q19, Q23	12.96
16	2, 3	Jan	Q7, Q8, Q11, Q14, Q21, Q22, Q23, Q24	12.92
17	3	Oct, Nov	-	12.75
18	4	Jan, Feb, Aug, Oct, Dec	Q1, Q5, Q6, Q22, Q24	12.27
19	4	Nov	-	11.78
20	4	Jan, Feb, Aug, Oct, Dec	Q10, Q11, Q13, Q14, Q15, Q21	11.36
21	5	Apr	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	11.21
22	5	Jun, Dec	Q17	10.83

ตารางผนวกที่ ๗๘ (ต่อ)

กลุ่ม	ระดับ เกรดสี่	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	ค่าความหวาน (°Brix)
23	5	Jan, Feb, Mar, May, Jul, Oct	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	10.39
24	5	Mar, May, Aug	Q1, Q3, Q7, Q9, Q14, Q16, Q24	10.33
25	5	Feb, Jul, Dec	Q1, Q3, Q7, Q9, Q14, Q16, Q24	9.77
26	5	Jan, Apr, Jun, Nov	Q1, Q3, Q7, Q9, Q14, Q16, Q24	9.40
27	5	Jun, Dec	Q5, Q6, Q13, Q15, Q18	9.32
28	5	Nov	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	9.21
29	5	-	Q2, Q19, Q20, Q22, Q23	8.94
30	5	Oct	Q1, Q3, Q7, Q9, Q14, Q16, Q24	8.82

ตารางผนวกที่ ข9 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง AQ (Month/Color/Amphoe/Quota)
โดยวิธี CHAID

กลุ่ม	ระดับ เกรดสี	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	อำเภอที่ปลูก	ค่าความหวาน (°Brix)
1	2, 3	May, Jun, Aug	-	A2 A6 A8 A11 A14 A16	16.53
2	2, 3	Feb, Dec	Q1, Q3, Q4, Q9, Q18, Q19, Q24	A2 A6 A12 A15 A17	15.02
3	2, 3	May, Jun, Aug	-	A1 A3 A4 A5 A12 A13 A15 A17	15.00
4	2, 3	Jan	Q2, Q3, Q4, Q6	-	14.78
5	2	Mar, Apr, Jul	-	-	14.76
6	2	Feb, Dec	Q5, Q10, Q17	-	14.37
7	3	Mar, Apr, Jul	-	-	14.20
8	2, 3	Jan	Q1, Q5	-	14.14
9	4	May	-	-	13.99
10	2, 3	Feb, Dec	Q1, Q3, Q4, Q9, Q18, Q19, Q24	A1 A5 A7 A8	13.94
11	3	Feb, Dec	Q5, Q10, Q17	-	13.69
12	2	Oct, Nov	-	-	13.58
13	2, 3	Jan	Q9, Q13, Q15, Q17	-	13.54
14	2, 3	Feb, Dec	Q6, Q7, Q11, Q12, Q13, Q14, Q15	-	13.44
15	3	Oct, Nov	-	A3 A4 A5 A15 A17	13.32
16	4	Mar, Apr, Jun, Jul	-	-	13.09
17	4	Jan, Feb, Aug, Oct, Dec	Q2, Q3, Q4, Q7, Q8, Q9, Q12, Q17, Q18, Q19, Q23	-	12.96
18	2, 3	Jan	Q7, Q8, Q11, Q14, Q21, Q22, Q23, Q24	-	12.92
19	3	Oct, Nov	-	A2 A6 A8 A11 A12	12.32
20	4	Jan, Feb, Aug, Oct, Dec	Q1, Q5, Q6, Q22, Q24	-	12.27

ตารางผนวกที่ ๗9 (ต่อ)

กลุ่ม	ระดับ เกรดสี่	เดือนที่รับวัตถุดิบ	ผู้รวบรวมวัตถุดิบ	อำเภอที่ปลูก	ค่าความหวาน (°Brix)
21	4	Nov	-	-	11.78
22	4	Jan, Feb, Aug, Oct, Dec	Q10, Q11, Q13, Q14, Q15, Q21	-	11.36
23	5	Apr	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	-	11.21
24	5	Apr, Jun, Dec	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	A14, A17	10.50
25	5	Jan, Feb, Mar, May, Jul, Oct	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	-	10.39
26	5	Mar, May, Aug	Q1, Q3, Q7, Q9, Q14, Q16, Q24	-	10.33
27	5	Feb, Jul, Dec	Q1, Q3, Q7, Q9, Q14, Q16, Q24	-	9.77
28	5	Jan, Apr, Jun, Nov	Q1, Q3, Q7, Q9, Q14, Q16, Q24	-	9.40
29	5	Nov	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	-	9.21
30	5	-	Q2, Q19, Q20, Q22, Q23	-	8.94
31	5	Apr, Jun, Dec	Q4, Q5, Q6, Q8, Q10, Q11, Q12, Q13, Q15, Q17, Q18, Q21	A4, A8, A10, A13	8.90
32	5	Oct	Q1, Q3, Q7, Q9, Q14, Q16, Q24	-	8.82

ตารางผนวกที่ ข10 การแบ่งกลุ่มข้อมูลการทำนายค่าแบบจำลอง S (Month/Color/Source)
โดยวิธี CHAID

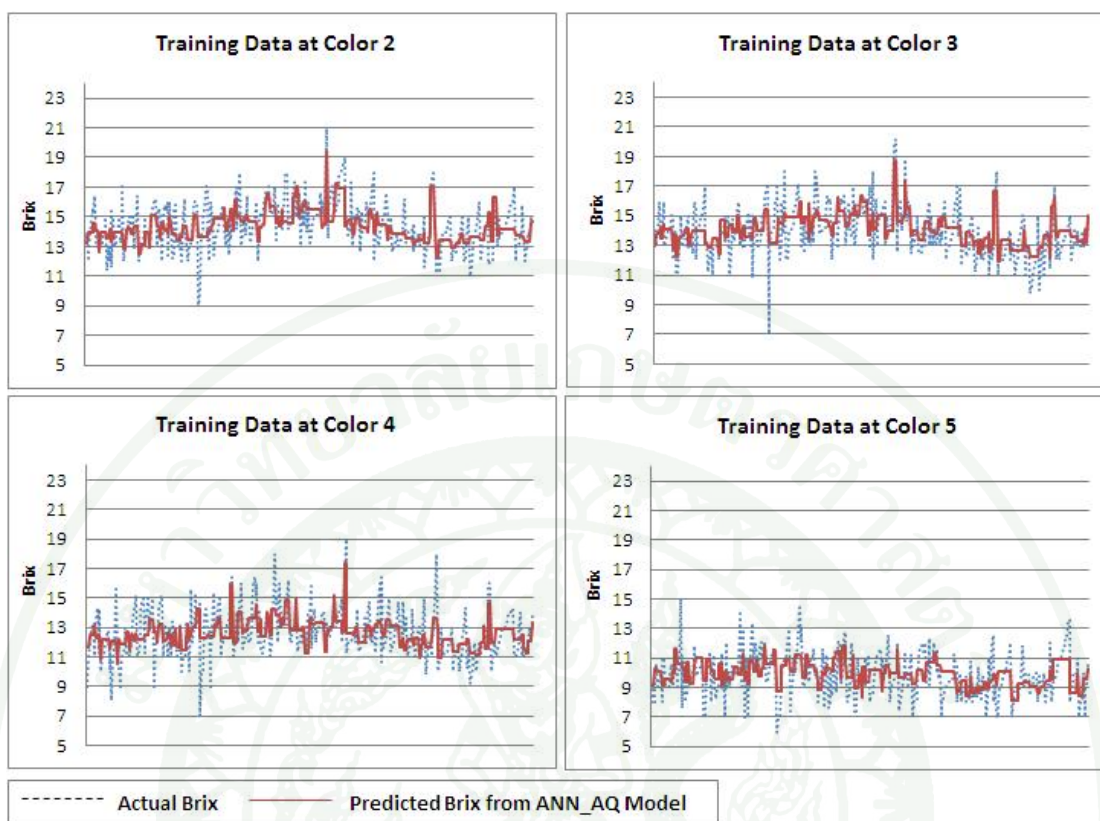
กลุ่ม	ระดับ เกรดสี	แหล่งที่มาวัตถุดิบ	เดือนที่รับวัตถุดิบ	ค่าความหวาน (°Brix)
1	2, 3	S5, S6, S20, S25, S30, S43, S44, S47		16.21
2	2, 3	S7, S9, S12, S14, S23, S24, S27, S31, S34, S36, S40, S48	May, Aug	15.42
3	2	S1, S3, S8, S10, S16, S21, S28, S41, S42, S45, S46		15.14
4	2, 3	S7, S9, S12, S14, S23, S24, S27, S31, S34, S36, S40, S48	Apr, Jun	14.72
5	2, 3	S2, S4, S13, S15, S18, S19, S33	Feb, May, Aug	14.62
6	3	S1, S3, S8, S10, S16, S21, S28, S41, S42, S45, S46		14.49
7	4	S5, S6, S16, S20, S21, S28, S34, S41, S43, S44, S47		14.23
8	2, 3	S7, S9, S12, S14, S23, S24, S27, S31, S34, S36, S40, S48	Jan, Feb, Mar, Jul, Dec	14.00
9	2, 3	S2, S4, S13, S15, S18, S19, S33	Mar, Apr, Jul	13.99
10	2	S7, S9, S12, S14, S23, S24, S27, S31, S34, S36, S40, S48	Oct, Nov	13.57
11	2, 3	S2, S4, S13, S15, S18, S19, S33	Jan, Dec	13.49
12	4	S1, S2, S4, S9, S12, S13, S14, S19, S23, S29, S35	Apr, May	13.41
13	2	S2, S4, S13, S15, S18, S19, S33	Oct, Nov	13.22
14	4	S3, S7, S8, S10, S15, S18, S24, S25, S26, S30, S31, S32, S36, S39, S40, S42, S46, S48	Feb, Mar, May, Jun, Jul, Dec	13.22
15	2	S11, S17, S22, S26, S29, S32, S35, S38, S39		13.05

ตารางผนวกที่ ข10 (ต่อ)

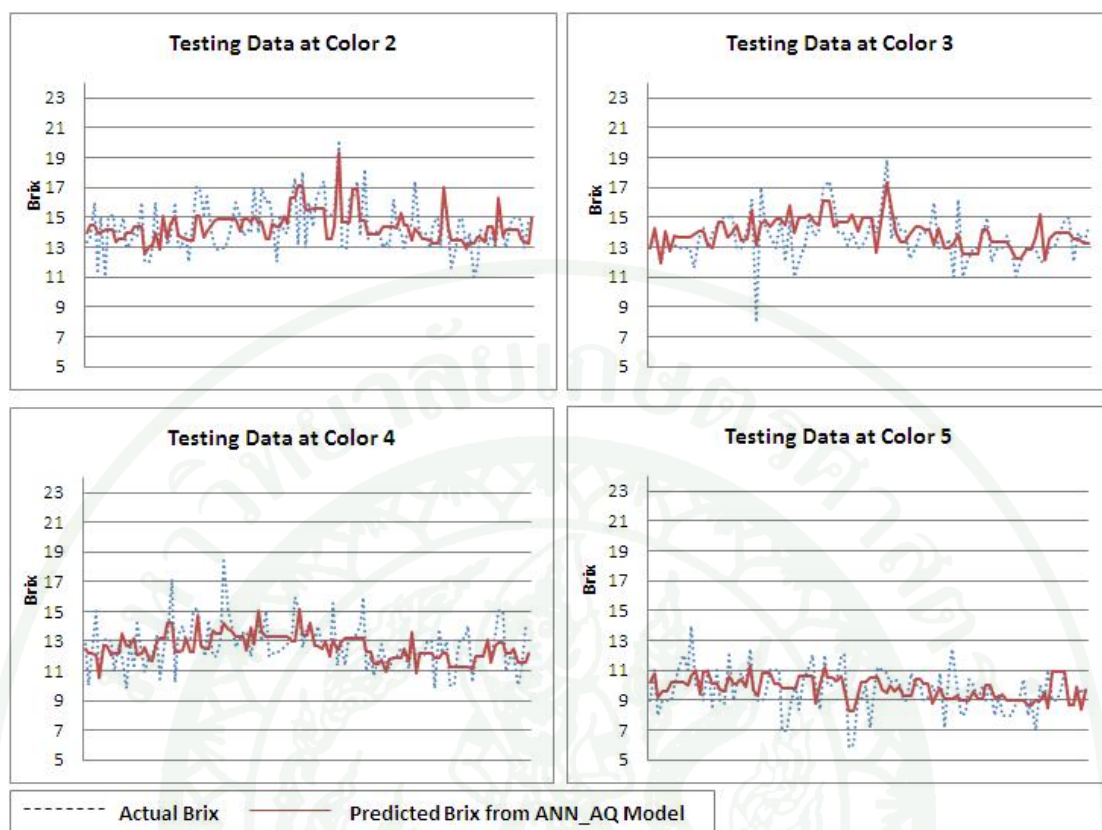
กลุ่ม	ระดับ เกรดสี่	แหล่งที่มาวัตถุดิบ	เดือนที่รับวัตถุดิบ	ค่าความหวาน (°Brix)
16	3	S7, S9, S12, S14, S23, S24, S27, S31, S34, S36, S40, S48	Oct, Nov	12.85
17	3	S2, S4, S13, S15, S18, S19, S33	Oct, Nov	12.46
18	4	S1, S2, S4, S9, S12, S13, S14, S19, S23, S29, S35	Jan, Feb, Mar, Jun, Jul, Aug, Oct	12.26
19	4	S3, S7, S8, S10, S15, S18, S24, S25, S26, S30, S31, S32, S36, S39, S40, S42, S46, S48	Jan, Aug, Oct, Nov	12.25
20	3	S11, S17, S22, S26, S29, S32, S35, S38, S39		12.12
21	4	S1, S2, S4, S9, S12, S13, S14, S19, S23, S29, S35	Nov, Dec	11.54
22	5	S8, S12, S13, S14, S20, S23, S25, S26, S30, S32, S33, S42, S43, S44, S48	Apr	11.41
23	5	S11, S16, S21, S27, S34, S47		11.34
24	4	S11, S17, S22, S27, S33, S37, S38		11.27
25	5	S8, S12, S13, S14, S20, S23, S25, S26, 30, S32, S33, S42, S43, S44, S48	Jan, Feb, Mar, May, Jun, Jul, Dec	10.18
26	5	S2, S3, S4, S7, S9, S10, S15, S19, S24, S28, S36, S39, S41	Mar, May, Jul, Aug	10.12
27	5	S2, S3, S4, S7, S9, S10, S15, S19, S24, S28, S36, S39, S41	Jan, Feb, Apr, Jun	9.61
28	5	S8, S12, S13 S14, S20, S23, S25, S26, S30, S32, S33, S42, S43, S44, S48	Oct, Nov	9.29
29	5	S1, S5, S6, S17, S18, S22, S29, S31, S35, S38, S40, S46		9.04
30	5	S2, S3, S4, S7, S9, S10, S15, S19, S24, S28, S36, S39, S41	Oct, Nov, Dec	9.04



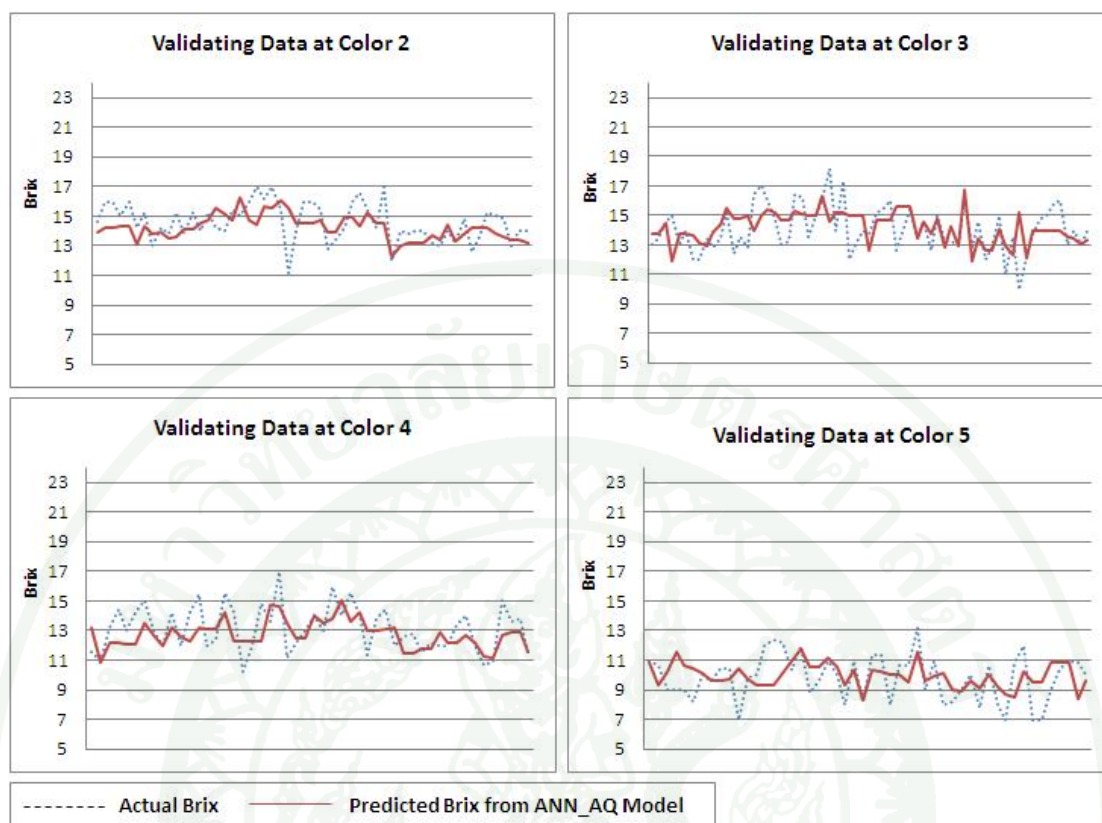
ภาคผนวก ค
การตรวจสอบค่าความคลาดเคลื่อนของแบบจำลองต่างๆ



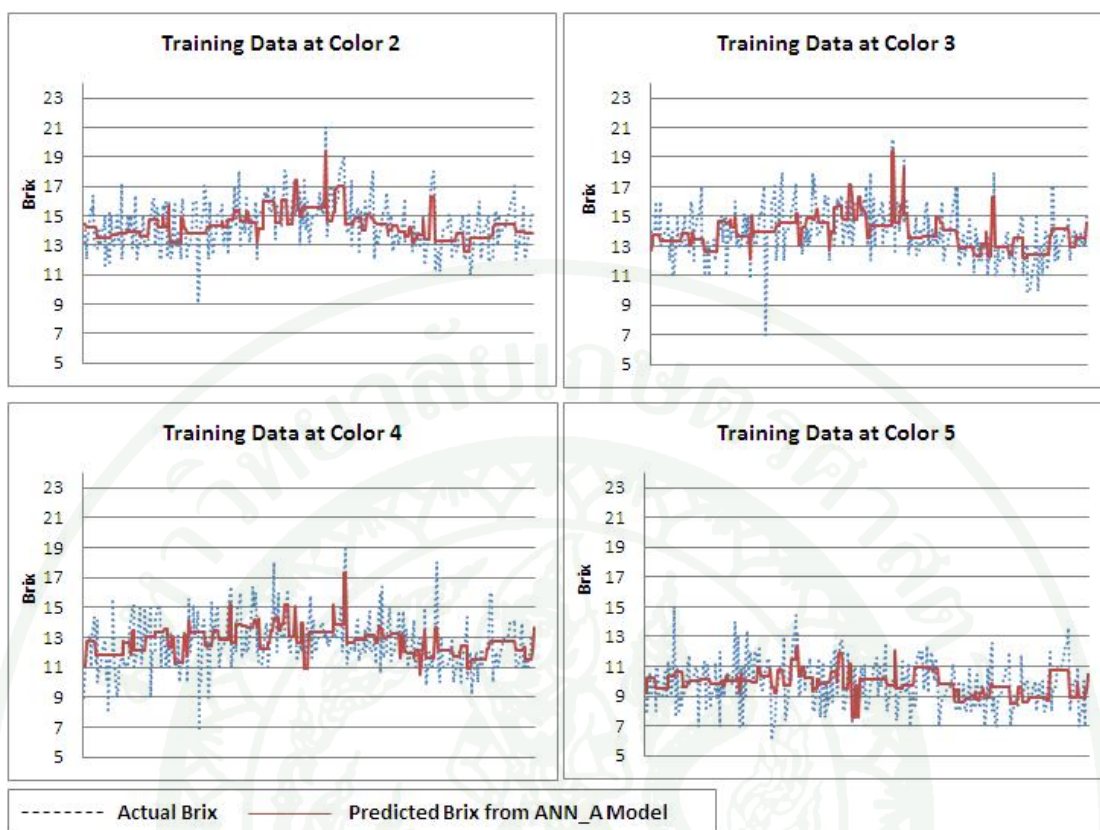
ภาพผนวกที่ ค1 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูง
และสูงมาก โดยแบบจำลอง AQ (Month/Color/Amphoe/Quota) วิธี
เครือข่ายประสาทเทียม



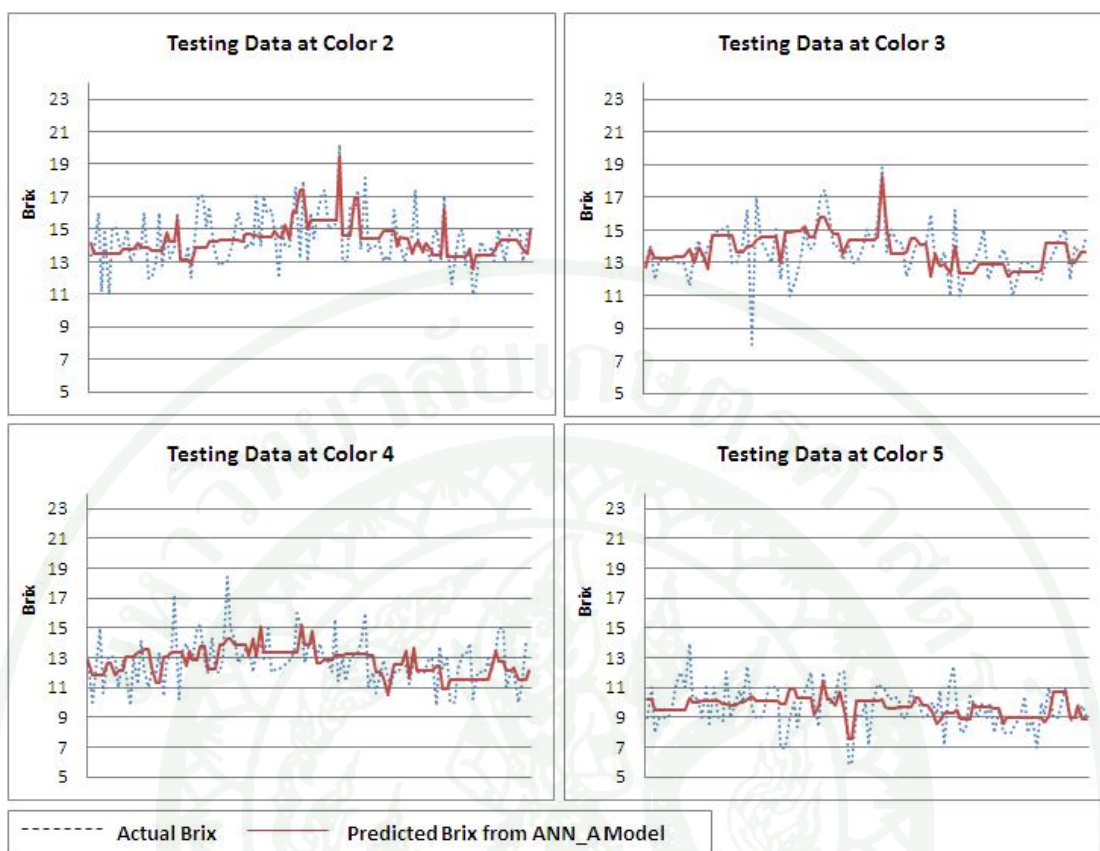
ภาพผนวกที่ ค2 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูง
และสูงมาก โดยแบบจำลอง AQ (Month/Color/Amphoe/Quota) วิธี
เครือข่ายประสาทเทียม



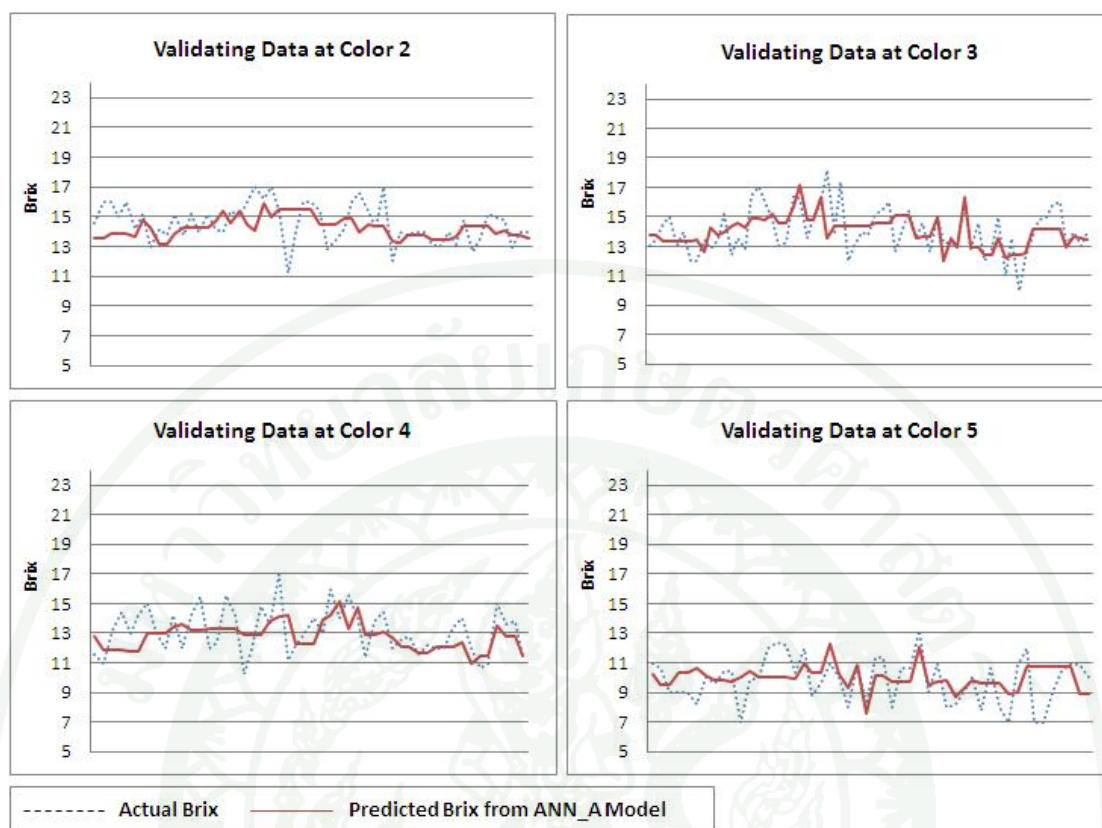
ภาพผนวกที่ ค3 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือสูง
และสูงมาก โดยแบบจำลอง AQ (Month/Color/Amphoe/Quota) วิธี
เครือข่ายประสาทเทียม



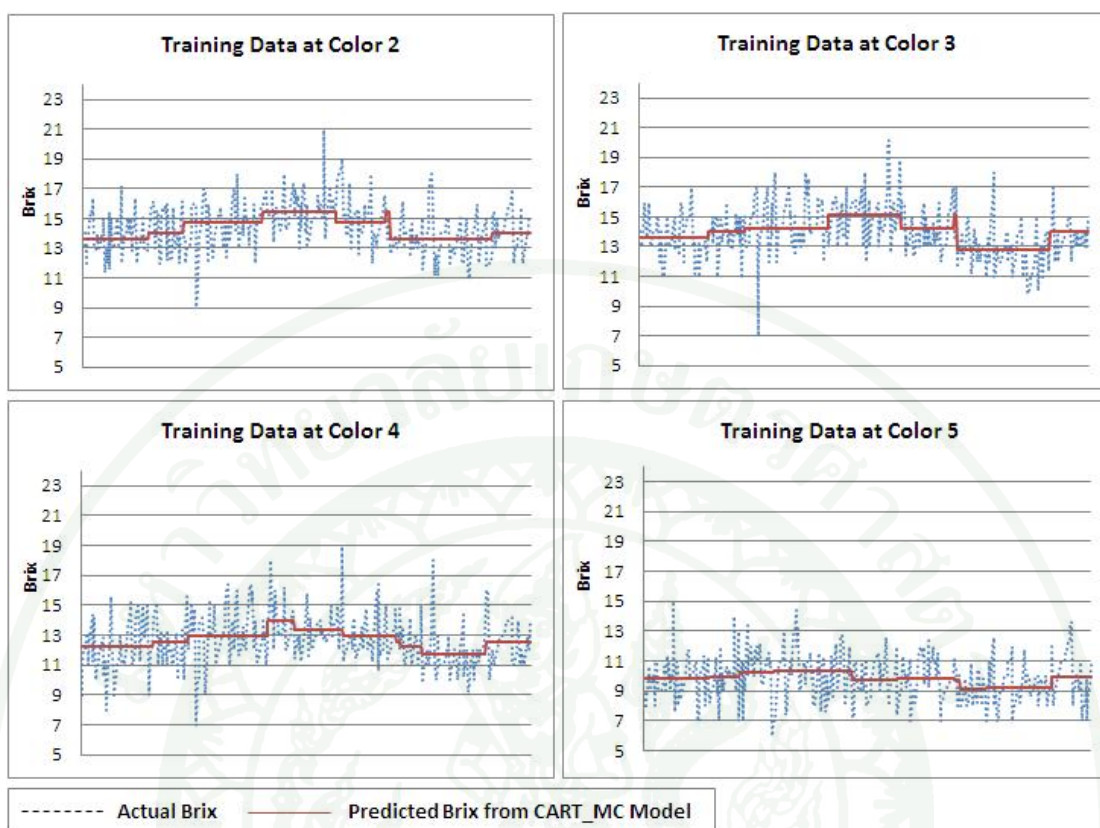
ภาพผนวกที่ ค4 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุด ในระดับความน่าเชื่อถือปาน
กลางและต่ำ โดยแบบจำลอง A (Month/Color/Amphoe) วิธีเครือข่าย
ประสาทเทียม



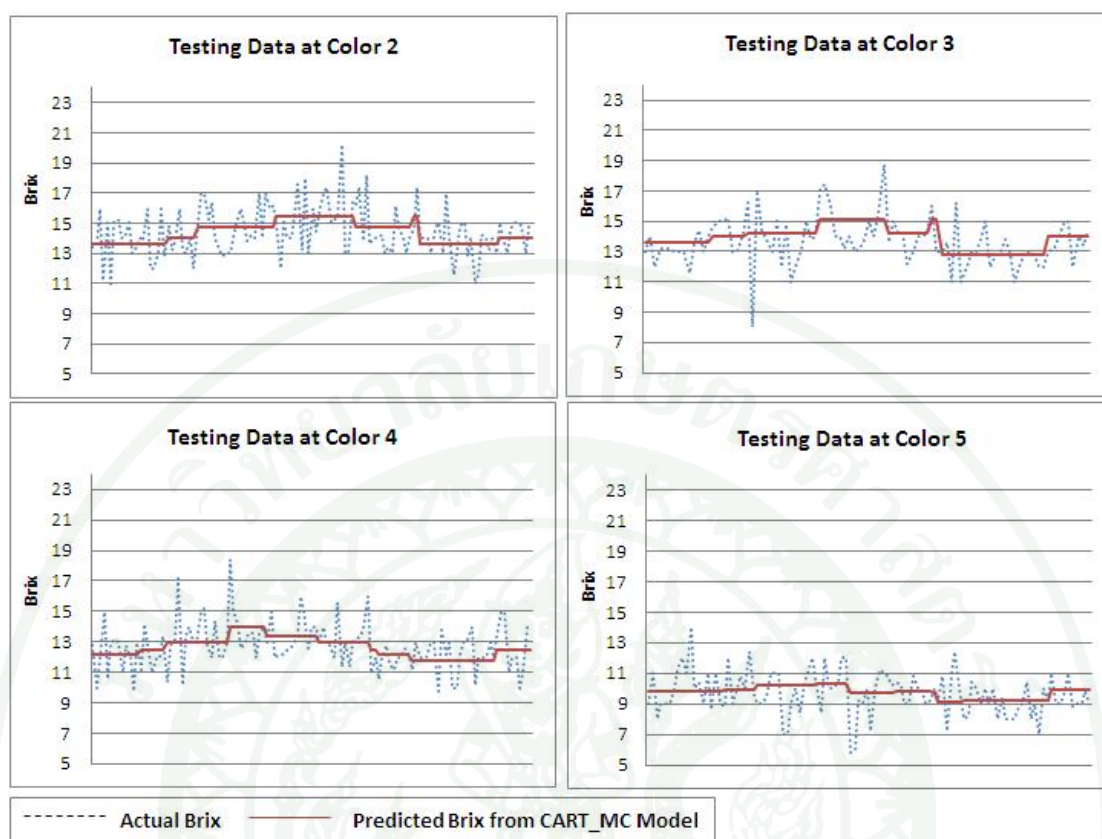
ภาพผนวกที่ ค5 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือปาน
กลางและต่ำ โดยแบบจำลอง A (Month/Color/Amphoe) วิเคราะห์ช่วย
ประสาทเทียม



ภาพผนวกที่ ๑๖ แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนายในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือปานกลางและต่ำ โดยแบบจำลอง A (Month/Color/Amphoe) วิธีเครือข่ายประสาทเทียม



ภาพผนวกที่ ค7 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดเรียนรู้ของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำ
มาก โดยแบบจำลอง MC (Month/Color) วิธี CART



ภาพผนวกที่ ค8 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำ
มาก โดยแบบจำลอง MC (Month/Color) วิธี CART



ภาพผนวกที่ ค9 แผนภาพเปรียบเทียบค่าความหวานจริงและค่าทำนายที่ได้จากการทำนาย
ในข้อมูลชุดทดสอบของแบบจำลองที่ดีที่สุดในระดับความน่าเชื่อถือต่ำ
มาก โดยแบบจำลอง MC (Month/Color) วิธี CART

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นางสาวณฤมล ชมบุญมี
วัน เดือน ปี ที่เกิด	วันที่ 31 สิงหาคม 2524
สถานที่เกิด	สมุทรสาคร
ประวัติการศึกษา	วท.บ. (เทคโนโลยีชีวภาพ) มหาวิทยาลัยเกษตรศาสตร์ (พ.ศ. 2546)
ตำแหน่งหน้าที่การงานปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและรางวัลทางวิชาการ	เข้าร่วมประชุมเสนอผลงานวิจัยในประชุมวิชาการสถิติ ประยุกต์ระดับชาติ คณะสถิติประยุกต์ สถาบันบัณฑิต พัฒนบริหารศาสตร์ ประจำปี 2552 ณ โรงแรมเอ-วัน เดอะรอยัลครุส พัทยา จังหวัดชลบุรี เมื่อวันที่ 21-22 พฤษภาคม พ.ศ. 2552 เรื่องการพัฒนาแบบจำลองเพื่อ ทำนายคุณลักษณะของวัตถุดิบในสายการผลิตสับปะรด บรรจุกระป๋อง
ทุนการศึกษาที่ได้รับ	ทุนวิจัยมหาบัณฑิต สกว. สาขาวิทยาศาสตร์และ เทคโนโลยี ภายใต้โครงการเชื่อมโยงภาคการผลิตกับ งานวิจัย ทุน สกว. – อุตสาหกรรม ปี 2551 (MRG- WI515S025)