สหกิจศึกษา

แบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ

THE ANALYSIS MODEL FOR THE TREND OF RESTAURANT'S DISSOLUTION WITH DECISION TREE METHOD

สิริกร ศิโรรัตน์ธัญโชค

สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2561

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

THE ANALYSIS MODEL FOR THE TREND OF RESTAURANT'S DISSOLUTION WITH DECISION TREE METHOD

SIRIKORN SIRORATTANYACHOK

CO-OPERATIVE EDUCATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR OF SCIENCE
PROGRAM IN APPLIED MATHEMATICS
DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2018

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อสหกิจศึกษา	แบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ The Analysis Model for The Trend of Restaurant's Dissolution with Decision Tree Method
ชื่อนักศึกษา	นางสาวสิริกร ศิโรรัตน์ธัญโชค
ปริญญา	วิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)
ภาควิชา	คณิตศาสตร์
ปีการศึกษา	2561
อาจารย์ที่ปรึกษา	ดร.บุษยมาส พิมพ์พรรณชาติ

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์) ประจำปีการศึกษา 2561

คณะกรรมการสอบ	ลายมือชื่อ
รศ.ไพรบูลย์ พันธรักษ์พงษ์ ประธานกรรมการ	
ดร.วรรณพร สรรประเสริฐ กรรมการ	
ดร.บุษยมาส พิมพ์พรรณชาติ กรรมการและอาจารย์ที่ปรึกษา	

ลิขสิทธิ์ของคณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อสหกิจศึกษา	แบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ The Analysis Model for The Trend of Restaurant's Dissolution with Decision Tree Method
ชื่อนักศึกษา	นางสาวสิริกร ศิโรรัตน์ธัญโชค รหัสนักศึกษา 58050168
ปริญญา	วิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)
ภาควิชา	คณิตศาสตร์
คณะ	วิทยาศาสตร์
มหาวิทยาลัย	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.)
ปีการศึกษา	2561
อาจารย์ที่ปรึกษา	ดร.บุษยมาส พิมพ์พรรณชาติ

บทคัดย่อ

โครงงานสหกิจศึกษาของภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เกิดขึ้นโดยมีวัตถุประสงค์เพื่อให้นักศึกษาได้เรียนรู้การปฏิบัติงานในสถานประกอบการจริง โดยนำความรู้ที่ได้ศึกษาจากสถาบันมาประยุกต์ใช้กับงานที่ได้รับมอบหมาย เพื่อพัฒนาทักษะทางด้านวิชาการ การสื่อสาร เรียนรู้ที่จะพัฒนาตนเอง ความรับผิดชอบ และการทำงานร่วมกับผู้อื่น

การเข้าร่วมโครงการสหกิจครั้งนี้ ผู้จัดทำได้มีโอกาสฝึกงานกับ บริษัท วงใน มีเดีย จำกัด ซึ่งเป็นเว็บไซต์และแอปพลิเคชันที่ได้รับการยอมรับว่า เป็นแอปฯ ค้นหาร้านอาหารอันดับ 1 ของไทยที่มีข้อมูลมากที่สุด ครอบคลุมทั้งร้านอาหาร ร้านเสริมสวยและสปา สูตรอาหาร โรงแรม ที่พัก ที่เที่ยว จากฐานข้อมูลมากกว่า 200,000 ร้านที่อัพเดทตลอดเวลา รวมทั้งยังได้รับข้อมูลที่ถูกต้องและรีวิวที่มาจากผู้ที่ไปใช้บริการมาจริงเพื่อช่วยประกอบการตัดสินใจ จากสมาชิกที่มีมากกว่า 3 ล้านคนทั่วประเทศ

ผู้จัดทำได้รับมอบหมายงานในตำแหน่ง Data Operations Team ให้สร้างแบบจำลองแนวโน้มการเลิกกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ เพื่อใช้คัดแยกประเภทของร้านอาหารที่มีการเปิดหรือปิดในระบบและถ้าหากพบข้อมูลของร้านที่ไม่เป็นประโยชน์ต่อผู้ใช้งานจะได้สามารถแก้ไขปรับปรุงข้อมูลให้เป็นประโยชน์ต่อผู้ใช้งานเว็บไซต์สูงสุด

เอกสารนี้เป็นลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ห้ามมิให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Title	The Analysis Model for The Trend of Restaurant's Dissolution With Decision Tree Method
Students	Miss Sirikorn Sirorattanyachok Student ID 58050168
Degree	Bachelor of Science (Applied Mathematics)
Department	Mathematics
Faculty	Science
University	King Mongkut's Institute of Technology Ladkrabang (KMITL)
Academic Year	2018
Advisor	Dr. Busayamas Pimpunchart

Abstract

Mathematics department's Cooperative Education Program of science faculty of King Mongkut's Institute of Technology Ladkrabang was created. The objective is let student learn about working in the real working place. Using knowledge from studying to applied with the work which is assigned. Improve skill about technical communication, self-learning, responsibility and working with others.

In this time, I have a chance to join with Wongnai Media Co.,Ltd. Wongnai is one of the leading lifestyle platforms in Thailand, that will help you discover great restaurants, hotels, tourist attractions, spas, and salons which cover over 200,000 locations from all over Thailand. I work as Data Operations team. I have to making The Analysis Model for The Trend of Restaurant's Dissolution with Decision Tree Method for Classification open or close restaurant on website. If finding things that do not benefit the users will have to update the information to be useful to most users of the Website.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

จากการที่ข้าพเจ้าได้มาปฏิบัติงานสหกิจศึกษา ณ บริษัท วงใน มีเดีย จำกัด ตั้งแต่ วันที่ 7 สิงหาคม พ.ศ. 2561 ถึง วันที่ 14 ธันวาคม พ.ศ. 2561 ส่งผลให้ข้าพเจ้าได้รับความรู้และประสบการณ์ที่มีประโยชน์มากมาย ในการจัดทำรายงานสหกิจครั้งนี้สำเร็จลงได้ด้วยดีเนื่องจากความร่วมมือและสนับสนุนจากหลายฝ่าย ดังนี้

1. คุณวันจักร ปาณะเจริญ ซึ่งเป็น อดีตManager Data Operations team
2. คุณวรวิทย์ ชัยวารีวงศ์ ซึ่งเป็น Acting Lead Data Operations team
3. Data Operations team ทุกท่าน
4. ดร.บุษยมาส พิมพ์พรรณชาติ ซึ่งเป็น อาจารย์ที่ปรึกษาสหกิจศึกษา
5. ประธานและคณะกรรมการสอบสหกิจศึกษา

ขอขอบคุณ คุณพ่อ คุณแม่ เพื่อนๆที่เป็นกำลังใจและบุคลากรท่านอื่นๆที่ไม่ได้กล่าวนามทุกท่านที่ได้ให้คำปรึกษาแนะนำช่วยเหลือในการจัดทำและซักถามความคืบหน้าของรายงาน ซึ่งเป็นแรงผลักดันให้ข้าพเจ้าสามารถทำรายงานเล่มนี้จนสำเร็จลุล่วงลงได้ด้วยดี

สุดท้ายนี้ ขอกราบขอบพระคุณสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ซึ่งเป็นสถานศึกษาที่ให้ความรู้แก่นักศึกษาทุกคน เพื่อให้นักศึกษาได้นำความรู้ไปใช้ประโยชน์ต่อตนเองและผู้อื่น รวมทั้งมีความรู้เพื่อที่จะนำไปประยุกต์ในการประกอบอาชีพสืบไป

สิริกร ศิโรรัตน์ธัญโชค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อ ..... ก

ABSTRACT ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฉ

สารบัญรูป ..... ช

บทที่ 1 บทนำ 9

1.1 ความเป็นมาและความสำคัญของการทำสหกิจศึกษา 9

1.2 วัตถุประสงค์ของสหกิจศึกษา 10

1.3 ขอบเขตของการทำสหกิจศึกษา 10

1.4 วิธีการดำเนินการสหกิจศึกษา 10

1.5 ประโยชน์ที่คาดว่าจะได้รับ 10

1.6 สถานที่ทำสหกิจศึกษา [1] 11

1.7 ระยะเวลาในการดำเนินงานสหกิจศึกษา 11

บทที่ 2 ความรู้พื้นฐาน 13

2.1 การสร้างแบบจำลองทางคณิตศาสตร์ 13

2.2 วิธีการต้นไม้ตัดสินใจ (DECISION TREE) 15

2.3 ตัววัดประสิทธิภาพของโมเดล CLASSIFICATION 20

2.4 ความรู้พื้นฐาน SQL 21

2.6 R STUDIO 29

บทที่ 3 วิธีการดำเนินงาน 36

3.1 แผนผังแสดงภาพรวมกระบวนการดำเนินงาน 36

3.2 ตัวแปรที่นำมาศึกษาในการสร้างแบบจำลอง 38

3.3 วิธีการดำเนินงาน 39

3.3.1 ตัวอย่างตารางข้อมูลที่เรียกมาใช้งาน (Raw Data) 39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 1 ระยะเวลาดำเนินงาน 11
ตารางที่ 2 วิธีการที่จะนำมาสร้างต้นไม้ตัดสินใจ 17
ตารางที่ 3 ตารางConfusion Matrix 20
ตารางที่ 4 ข้อมูลที่เรียกมาใช้งานจากระบบฐานข้อมูล 39
ตารางที่ 5 ข้อมูลที่จะนำไปใช้งาน 41
ตารางที่ 6 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 80% และ testing data 20 %... 55
ตารางที่ 7 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 70% และ testing data 30 %... 58
ตารางที่ 8 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 60% และ testing data 40 %... 60
ตารางที่ 9 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 50% และ testing data 50 %... 62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

ภาพที่ **หน้า**

ภาพที่ 1 โลโก้ บริษัท วงใน มีเดีย จำกัด 11
ภาพที่ 2 แผนผังขั้นตอนการสร้างแบบจำลองทางคณิตศาสตร์ 13
ภาพที่ 3 การกรองข้อมูลลำดับที่1 26
ภาพที่ 4 การกรองข้อมูลลำดับที่2 26
ภาพที่ 5 การกรองข้อมูลลำดับดับที่3 27
ภาพที่ 6 การกรองข้อมูลลำดับดับที่4 27
ภาพที่ 7 การกรองข้อมูลลำดับดับที่5 27
ภาพที่ 8 การกรองข้อมูลลำดับดับที่6 28
ภาพที่ 9 การกรองข้อมูลลำดับดับที่7 29
ภาพที่ 10 โปรแกรม R Studio 29
ภาพที่ 11 โครงสร้างองค์กร 31
ภาพที่ 12 แผนผังแสดงภาพรวมกระบวนการดำเนินงาน 36
ภาพที่ 13 ขั้นตอนการทำความสะอาดข้อมูล (Cleaning data) 40
ภาพที่ 14 Decision Tree C5.0 Algorithm (training 80 % and testing 20 %) 47
ภาพที่ 15 Decision Tree CART Algorithm (training 80 % and testing 20 %) 48
ภาพที่ 16 Decision Tree C5.0 Algorithm (training 70 % and testing 30 %) 49
ภาพที่ 17 Decision Tree CART Algorithm (training 70 % and testing 30 %) 50
ภาพที่ 18 Decision Tree C5.0 Algorithm (training 60 % and testing 40 %) 51
ภาพที่ 19 Decision Tree CART Algorithm (training 60 % and testing 40 %) 52
ภาพที่ 20 Decision Tree C5.0 Algorithm (training 50 % and testing 50 %) 53
ภาพที่ 21 Decision Tree CART Algorithm (training 50 % and testing 50 %) 54
ภาพที่ 22 กราฟแสดงความถี่ของจำนวนร้านอาหารที่ปิดไปแล้วของแต่ละเขตในจังหวัดกรุงเทพมหานคร 65
ภาพที่ 23 ตัวอย่างของร้านอาหารที่ปิดในระบบในเขตวัฒนาที่จะนำมาใช้ในการสร้างแบบจำลอง .. 66
ภาพที่ 24 ตัวอย่างของร้านอาหารที่ยังเปิดอยู่ในระบบในเขตวัฒนาที่จะนำมาใช้ในการสร้าง 66
ภาพที่ 25 Decision Tree ของแบบจำลองที่พัฒนาใหม่ 71
ภาพที่ 26 แสดงขั้นตอนการเลือก Attribute สำหรับตัวเลข 72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพที่ 27 Decision Tree บางส่วนจากการทำแบบจำลองครั้งที่1 73
ภาพที่ 28 แสดงผลจากการใช้แบบจำลองจากการทำครั้งที่ 1 76
ภาพที่ 29 แสดงผลจากการใช้แบบจำลองที่พัฒนาใหม่ 76
ภาพที่ 30 ผลลัพธ์จากการใช้แบบจำลองจากการทำครั้งที่ 1 77
ภาพที่ 31 ผลลัพธ์จากการแบบจำลองที่พัฒนาใหม่ 78
ภาพที่ 32 ตาราง correlation 80
ภาพที่ 33 แสดงผลจากการใช้แบบจำลองที่ตัดตัวแปรจากการตรวจสอบcorrelation 81
ภาพที่ 34 แสดงผลจากการใช้แบบจำลองที่พัฒนาใหม่ 81
ภาพที่ 35 หน้าต่าง application ก่อนยังไม่มี e-menu lineman 82
ภาพที่ 36 หน้าต่าง application เมื่อมี e-menu lineman 82
ภาพที่ 37 การคัดแยกประเภทของอาหารตามสั่งและอาหารจานเดียว 83
ภาพที่ 38 หน้าต่าง log in ระบบฐานข้อมูล 84
ภาพที่ 39 ตัวอย่างโค้ดในการ query 84

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

ในบทนี้จะกล่าวถึง ความเป็นมาของการทำสหกิจศึกษา วัตถุประสงค์ของการทำสหกิจศึกษา ขอบเขตของการทำสหกิจศึกษา วิธีการดำเนินงาน ประโยชน์ที่คาดว่าจะได้รับและสถานที่ทำสหกิจศึกษา ซึ่งมีรายละเอียดดังต่อไปนี้

1.1 ความเป็นมาและความสำคัญของการทำสหกิจศึกษา

เนื่องจากหลักสูตรวิทยาศาสตรบัณฑิต สาขาคณิตศาสตร์ประยุกต์ มีแผนการศึกษาทางเลือกแบบหกิจศึกษา (Cooperative Education) เป็นการศึกษาที่เน้นการปฏิบัติงานจริงในสถานประกอบการอย่างมีระบบ โดยจัดให้มีการเรียนในมหาวิทยาลัยร่วมกับการจัดให้นักศึกษาไปปฏิบัติงานจริง ณ สถานประกอบการที่ให้ความร่วมมืออย่างมีหลักการและเป็นระบบ ทำให้นักศึกษาได้มีโอกาสสร้างความเข้าใจและเรียนรู้จากการปฏิบัติงาน สามารถพัฒนาทักษะทางวิชาชีพเพื่อให้มีคุณภาพตรงตามที่สถานประกอบการต้องการมากที่สุด โดยมีสหกิจศึกษาเป็นกลไกความร่วมมืออย่างเป็นรูปธรรมระหว่างสถานประกอบการและมหาวิทยาลัย และยังก่อให้เกิดประโยชน์ร่วมกัน นักศึกษาสหกิจศึกษาจะเข้าปฏิบัติงานแบบเต็มเวลาอย่างต่อเนื่องเสมือนหนึ่งเป็นพนักงานชั่วคราวในสถานประกอบการเป็นระยะเวลา 1 ภาคการศึกษา และได้รับประสบการณ์จากการปฏิบัติงานในสถานประกอบการจริง ซึ่งนักศึกษาไม่สามารถเรียนรู้ได้ในห้องเรียนปกติ นักศึกษาจะได้รับการพัฒนาทักษะของงานอาชีพที่ผสมผสานระหว่างภาคทฤษฎีและภาคปฏิบัติได้อย่างลงตัว อันจะส่งผลให้นักศึกษาเป็นบัณฑิตที่มีคุณภาพสูงเป็นที่ต้องการของตลาดแรงงานมากขึ้น และกระบวนการสหกิจศึกษาจะก่อให้เกิดการประสานงานอย่างใกล้ชิดกับสถานประกอบการ ทำให้มหาวิทยาลัยสามารถพัฒนาหลักสูตรได้ตลอดเวลา ส่วนสถานประกอบการจะได้บัณฑิตที่มีคุณภาพตรงตามความต้องการของสถานประกอบการต่างๆ

ผู้จัดทำได้เลือกทำสหกิจศึกษากับทางบริษัท วงใน มีเดีย จำกัด จำกัด (Wongnai Media Co,.Ltd.) ซึ่งเป็นเว็บไซต์และแอปพลิเค-ชันที่ได้รับการยอมรับว่า เป็นแอปฯ ค้นหาร้านอาหารอันดับ 1 ของไทยที่มีข้อมูลมากที่สุด ครอบคลุมทั้งร้านอาหาร ร้านเสริมสวยและสปา สูตรอาหาร โรงแรม ที่พัก ที่เที่ยว จากฐานข้อมูลมากกว่า 200,000 ร้านที่อัพเดทตลอดเวลา รวมทั้งยังได้รับข้อมูลที่ถูกต้องและรีวิวที่มาจากผู้ที่ไปใช้บริการมาจริงเพื่อช่วยประกอบการตัดสินใจ จากสมาชิกที่มีมากกว่า 3 ล้านคนทั่วประเทศ ซึ่งผู้จัดทำได้ทำสหกิจศึกษาเป็นระยะเวลารวมทั้งสิ้น 18 สัปดาห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของสหกิจศึกษา

1. เพื่อเรียนรู้การปฏิบัติงานในสถานประกอบการจริง โดยนำความรู้ที่ได้ศึกษาจากสถาบันมาประยุกต์ใช้กับงานที่ได้รับมอบหมาย
2. เรียนรู้โฟลว์กระบวนการทำงาน และศึกษาการวิเคราะห์ข้อมูลในฐานะ Data Admin
3. เรียนรู้กระบวนการเพื่อปรับปรุงและพัฒนาแบบจำลองให้ดียิ่งขึ้น

1.3 ขอบเขตของการทำสหกิจศึกษา

1. สร้างแบบจำลองการวิเคราะห์แนวโน้มการปิดกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ

2. ทดสอบค่าความถูกต้องและค่าความแม่นยำของแบบจำลองการวิเคราะห์แนวโน้มการปิดกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ

1.4 วิธีการดำเนินการสหกิจศึกษา

1. ยื่นเรื่องขอทำสหกิจศึกษาในตำแหน่ง Data Operations team กับทางบริษัท วงในมีเดีย จำกัด

2. ภาควิชาอนุญาตให้ทำสหกิจศึกษาในตำแหน่งตามข้อ 1 และขอให้คณะจัดทำหนังสือส่งตัว

3. รายงานตัวเข้าทำงานที่บริษัทและรับมอบหมายงานจากพนักงานพี่เลี้ยงในระยะเวลาที่กำหนดคือ 18 สัปดาห์ (วันที่ 7 สิงหาคม พ.ศ. 2561 ถึง วันที่ 14 ธันวาคม พ.ศ. 2561)

4. ระหว่างปฏิบัติงานเสนอหัวข้อเรื่องที่จะทำรายงานสหกิจศึกษาต่ออาจารย์ที่ปรึกษา โดยเลือกจากปัญหาที่ทางทีมพบเจอและกำลังดำเนินการหาวิธีการแก้ไขอยู่ในปัจจุบัน

5. จัดทำรายงานสหกิจศึกษา

5. แก้ไขรายงานสหกิจศึกษาให้มีความถูกต้อง

6. นำเสนอสหกิจศึกษาต่อคณะกรรมการสอบ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้เรียนรู้กระบวนการสร้างแบบจำลองการวิเคราะห์แนวโน้มการปิดกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ

2. ได้เรียนรู้เกี่ยวกับเครื่องมือที่ใช้ในการแยกประเภทของข้อมูล (Classification)

3. ได้ประสบการณ์ในการทำงานภายในองค์กรด้านข้อมูล ส่งต่อข้อมูลไปยังผู้ใช้งานให้ได้ประสิทธิภาพมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 สถานที่ทำสหกิจศึกษา [1]

ภาพที่ 1 โลโก้ บริษัท วงใน มีเดีย จำกัด

บริษัท วงใน มีเดีย จำกัด

117/8,12,13,16,17 อาคารปานจิตต์ ทาวเวอร์ ชั้น 5, 9, 10, 14, 15

ซอย สุขุมวิท 55 ถนน สุขุมวิท แขวงคลองตันเหนือ เขตวัฒนา กรุงเทพมหานคร 10110

1.7 ระยะเวลาในการดำเนินงานสหกิจศึกษา

ระยะเวลาที่ปฏิบัติงาน ผู้จัดทำได้เริ่มปฏิบัติสหกิจศึกษาตั้งแต่วันที่ 7 สิงหาคม พ.ศ. 2561 ถึง วันที่ 14 ธันวาคม พ.ศ. 2561 (รวมเป็นระยะเวลา 18 สัปดาห์)

ตารางที่ 1 ระยะเวลาดำเนินงาน

กิจกรรมดำเนินงาน	ระยะเวลา					
	ปี 2560					
	ก.ค.	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.
1. ยื่นเรื่องขอทำสหกิจศึกษากับบริษัท วงใน มีเดีย จำกัด						
2. ได้รับอนุมัติให้ทำสหกิจศึกษา และได้รับมอบหมายงานในData Operations team						
3 เรียนรู้เกี่ยวกับขั้นตอนการทำงานของData Operations team						
4.1 ได้รับมอบหมายให้สร้างแบบจำลองการวิเคราะห์						

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 ขั้นตอนการทำความสะอาดข้อมูล (Cleaning Data) 40
3.3.3 ตัวอย่างตารางข้อมูลที่นำมาใช้งาน (Data Set) 41
3.4 ตรวจสอบค่าPERFORMANCEของแบบจำลองครั้งที่1 55
3.5 การพัฒนาแบบจำลอง 65
3.6 ตรวจสอบค่าPERFORMANCEของแบบจำลองที่พัฒนาใหม่ 70
บทที่ 4 ผลการทดสอบและอภิปรายผล 75
4.1 ผลการทดสอบแบบจำลอง 75
4.2 การอภิปรายผล 79
4.3 การอภิปรายผลเพิ่มเติม 80
4.4 งานที่ได้รับมอบหมายเพิ่มเติม 82
บทที่ 5 สรุปผลการทำสหกิจศึกษาและข้อเสนอแนะ 84
5.1 สรุปผลการทำสหกิจศึกษา 85
5.2 ปัญหาอุปสรรค 86
5.3 แนวทางแก้ปัญหา 86
5.4 ข้อเสนอแนะ 86
เอกสารอ้างอิง 87
ภาคผนวก 93
ภาคผนวก ก 94
ภาคผนวก ข 99
ภาคผนวก ค 101

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แนวโน้มการปิดกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ 4.2 ศึกษาและค้นคว้า เรื่อง Decision Tree 4.3 ศึกษาเรียนรู้การ Classification Using R Studio 4.4 พิจารณาข้อมูลเพื่อเลือกใช้คุณสมบัติที่จะมาใช้ในการสร้างแบบจำลอง และทำความสะอาดข้อมูลเพื่อนำมาใช้งาน 4.5 ดำเนินการ Classification ด้วยโปรแกรม R Studio						
5.1 บันทึกผลการดำเนินเนินงาน 5.2 ดำเนินการแก้ไขหากมีข้อผิดพลาดที่เกิดขึ้น ปรึกษาพี่เลี้ยงและอาจารย์ที่ปรึกษาสหกิจ						
6. ทดสอบค่าความแม่นยำและความถูกต้องของแบบจำลอง						
7. ปฏิบัติงานอื่นๆตามที่ได้รับมอบหมาย						
8. จัดทำรูปเล่มสหกิจศึกษาพร้อมทั้งจัดทำการนำเสนอ						

ในบทที่ 2 อธิบายถึงความรู้พื้นฐานที่ควรทราบในการทำงานที่ได้รับมอบหมายนี้ บทที่ 3 อธิบายถึงวิธีการดำเนินการ หรือกระบวนการทำงานว่าเป็นอย่างไร ในบทที่ 4 กล่าวถึงผลของการดำเนินงานสหกิจศึกษาที่ได้รับหมายและสรุปสุดท้ายอยู่ในบทที่ 5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

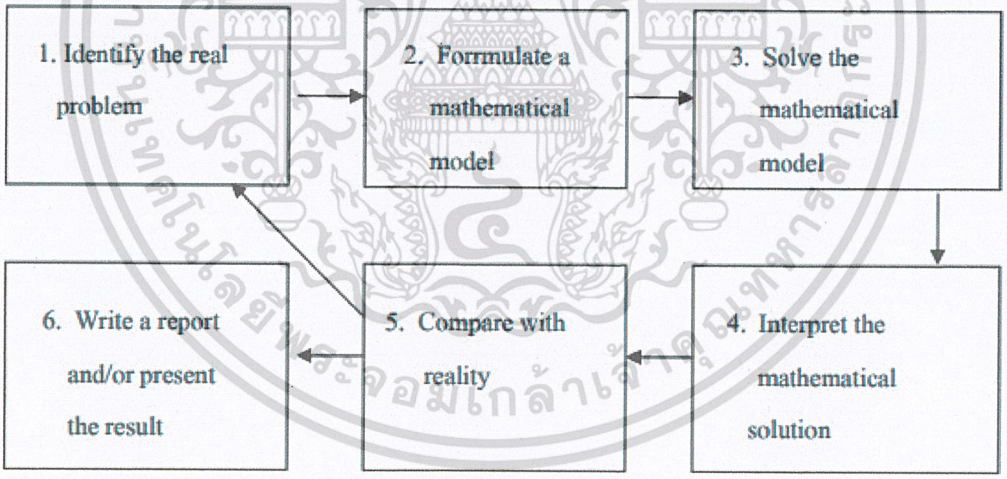
บทที่ 2
ความรู้พื้นฐาน

เนื่องจากงานที่ผู้จัดทำได้รับมอบหมายคือการสร้างแบบจำลองการวิเคราะห์แนวโน้มการปิดกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ ผู้จัดทำจึงจำเป็นจะต้องมีความรู้ในเรื่องต่อไปนี้

2.1 การสร้างแบบจำลองทางคณิตศาสตร์

โลกของคณิตศาสตร์ (Mathematical World) เป็นโลกของนามธรรมที่มีการกำหนดสัญลักษณ์แทนนามธรรมเหล่านั้น มีการนิยามข้อตกลงต่างๆ เพื่อใช้ในการศึกษาระบบ และ พิสูจน์ผลที่ได้เป็นกฎหรือทฤษฎีเพื่อการนำ ไปใช้ ดังนั้นการสร้างตัวแบบเชิงคณิตศาสตร์จึงจัดว่าเป็นการจำลองสิ่งที่อยู่ในโลกแห่งความเป็นจริง ให้เป็นตัวแบบเชิงคณิตศาสตร์ที่อยู่ในรูป นามธรรม

และใช้กฎเกณฑ์ทางคณิตศาสตร์หาคำตอบของตัวแบบเพื่อนำผลกลับไปประยุกต์ใช้กับโลกแห่งความเป็นจริง ซึ่งสามารถสรุปเป็นขั้นตอนได้ 6 ขั้นตอน ดังต่อไปนี้



ภาพที่ 2 แผนผังขั้นตอนการสร้างแบบจำลองทางคณิตศาสตร์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 1 ทำความเข้าใจปัญหา (Identify the real problem)

การทำความเข้าใจปัญหา ขั้นตอนแรกในการสร้างตัวแบบ ต้องวิเคราะห์ให้ทราบว่า ปัญหาคืออะไร มีอะไรบ้างที่เกี่ยวข้องกับปัญหา มีคำถามมากมายที่ต้องทำความเข้าใจโจทย์ ปัญหาในสถานการณจริง เช่น ปัญหานี้ต้องการทราบอะไร มีวัตถุประสงค์และเป้าหมายอะไร จะตัดสินผลที่ออกมาอย่างไร แหล่งข้อมูลมาจากไหนเชื่อถือได้หรือไม่ มีคำตอบเป็นแบบ เดียวหรือไม่ จำแนกปัญหาว่าเป็น แบบมีคำตอบแน่นอน(deterministic) หรือ แบบมีคำตอบ ไม่แน่นอน (stochastic) ต้องใช้การสร้างสถานการณ์จำลอง (simulation) หรือไม่ คำถามหรือคำตอบ ดังกล่าวมาแล้วต้องนิยาม กำหนดขอบเขต ให้ตรงประเด็นและชัดเจน

ขั้นที่ 2 สร้างตัวแบบเชิงคณิตศาสตร์ (Formulate a mathematical model)

ขั้นตอนนี้เป็นขั้นตอนการสร้างตัวแบบเชิงคณิตศาสตร์ จากปัญหาที่ได้วิเคราะห์หรือ ทำให้ชัดเจนแล้วในขั้นที่ 1 ทดลองสร้างตัวแบบที่ซับซ้อนน้อยที่สุดก่อน เขียนแผนภาพตาม ความเหมาะสม เขียนรายการปัจจัยที่เกี่ยวข้อง รวบรวมข้อมูลและทดสอบเนื้อหารายละเอียด อธิบายพฤติกรรมของตัวแปร รวบรวมข้อมูลเพิ่มเติมถ้าจำเป็น แสดงตัวแปรแต่ละตัวด้วย สัญลักษณ์ที่เหมาะสมพร้อมทั้งกำหนดหน่วย กำหนดข้อสมมติที่ต้องการสร้าง เขียน ความสัมพันธ์และสมการของตัวแปรในโจทย์โดยใช้ความรู้ทางคณิตศาสตร์ เช่น การเป็นสัดส่วน ความสัมพันธ์เชิงเส้นและไม่เชิงเส้น ความสัมพันธ์จากการทดลอง หลักการ inputoutput กฎการเคลื่อนที่ของนิวตัน สมการเชิงผลต่างและสมการเชิงอนุพันธ์ เมทริกซ์ความน่าจะเป็น การกระจายเชิงสถิติ เป็นต้น

ขั้นที่ 3 หาผลลัพธ์ของตัวแบบ (Solve the mathematical model)

การหาคำตอบทางคณิตศาสตร์ของตัวแบบ อาจจะใช้วิธีเกี่ยวกับพีชคณิต หรือใช้วิธีเชิงตัวเลข ใช้แคลคูลัสและกราฟ เขียนโปรแกรมคอมพิวเตอร์หรือใช้โปรแกรมสำเร็จรูปที่ เหมาะสม หาค่าของตัวแปรที่ต้องการ อาจจะเป็นรูปแบบตารางหรือรูปภาพ

ขั้นที่ 4 แปลความหมายของผลลัพธ์ (Interpret the mathematical solution)

ขั้นตอนนี้เป็นการแปลความหมาย และตรวจสอบผลลัพธ์ที่หาได้จากวิธีการเชิงคณิตศาสตร์ เช่น พิจารณาค่าของตัวแปรที่หาค่าได้ ว่ามีครื่องหมาย และขนาดถูกต้องหรือไม่ มีค่าเพิ่มหรือลดตามที่ควรจะเป็นหรือไม่พิจารณาค่ามากและค่าน้อยของตัวแปรเพื่อตรวจสอบ พฤติกรรมความไวต่อสิ่งกระตุ้น ได้คำตอบที่ดีที่สุดตามที่คาดไว้หรือไม่หรือต้องเปลี่ยน เงื่อนไขเริ่มต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่ 5 ตรวจสอบผลลัพธ์กับข้อมูลจริง (Compare with reality)

ผลลัพธ์ที่ได้สามารถตรวจสอบกับข้อมูลจริงได้หรือไม่คำตอบเชิงคณิตศาสตร์มี ความหมายหรือไม่ การทำนายสอดคล้องกับข้อมูลจริงหรือไม่ ประเมินตัวแบบที่สร้างขึ้นว่าได้ ครบตามวัตถุประสงค์หรือไม่ ตัวแบบสามารถปรับปรุงให้ดีขึ้นได้อีกหรือไม่ ผลลัพธ์ที่ได้ ก่อนหน้านี้ ชี้ให้เห็นว่าต้องคำนวณหาค่าตัวแปรจากตัวแบบที่ปรับปรุงใหม่เพื่อความแม่นยำที่ดีกว่าหรือไม่ ถ้าต้องการทำใหม่ก็ต้องกลับไปเริ่มที่ขั้นที่ 1 หรือถ้าไม่ต้องการให้ไปที่ขั้นที่ 6 ขั้นตอนนี้สำคัญมาก เพราะมีบ่อยครั้งที่ต้องสร้างตัวแบบหลายรอบก่อนที่จะได้ผลเป็นที่น่าพอใจ

ขั้นที่ 6 เขียนรายงาน (Write a report)

การเขียนรายงานต้องทราบว่า เขียนเพื่อใคร ผู้อ่านต้องการทราบอะไร ต้องการรายละเอียดในรายงานมากน้อยเพียงใด จะสร้างรายงานอย่างไร จึงจะทำให้ลักษณะที่สำคัญ ชัดเจน และผลที่ต้องการทราบปรากฏอยู่ ขั้นตอนนี้อาจจะไม่ต้องทำ ถ้าไม่ทราบว่าจะเขียนให้ใครอ่าน

2.2 วิธีการต้นไม้ตัดสินใจ (Decision Tree)

การเรียนรู้ของต้นไม้ตัดสินใจ (Decision Tree) เป็นการเรียนรู้โดยการจำแนกประเภท (Classification) ข้อมูลออกเป็นกลุ่ม (class) ต่างๆ โดยใช้คุณลักษณะ (attribute) ข้อมูลในการจำแนกประเภท ต้นไม้ตัดสินใจที่ได้จากการเรียนรู้ทำให้ทราบว่า คุณลักษณะใดเป็นตัวกำหนดการจำแนกประเภท และคุณลักษณะแต่ละตัวมีความสำคัญมากน้อยต่างกันอย่างไร

ส่วนประกอบของผลลัพธ์วิธีการต้นไม้ตัดสินใจ

- โหนดภายใน (internal node) คือ คุณลักษณะต่างๆ ของข้อมูล ซึ่งเมื่อข้อมูลใดๆ ตกลงมาที่โหนด จะใช้คุณลักษณะนี้เป็นตัวตัดสินใจว่าข้อมูลจะไปในทิศทางใด โดยโหนดภายในที่เป็นจุดเริ่มต้นของต้นไม้ เรียกว่า โหนดราก
- กิ่ง (branch, link) เป็นค่าของคุณลักษณะในโหนดภายในที่แตกกิ่งนี้ออกมา ซึ่งโหนดภายในจะแตกกิ่งเป็นจำนวนเท่ากับจำนวนค่าของคุณลักษณะในโหนดภายในนั้น
- โหนดใบ (leaf node) คือกลุ่มต่างๆ ซึ่งเป็นผลลัพธ์ในการจำแนกประเภทข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนวิธีการสร้างDecision Tree

- ต้นไม้ตัดสินใจสร้างโดยวิธีแบบ top-down recursive
- เริ่มต้นด้วยนำตัวอย่างการสอน มาสร้างเป็นราก
- Attribute ควรอยู่ในรูปของ Categorical คือ ข้อมูลชนิดกลุ่ม หากเป็นข้อมูลที่อยู่ในรูป Continuous หรือ Numeric เป็นข้อมูลมีความต่อเนื่องกัน ควรทำแบ่งข้อมูลให้เป็นกลุ่มก่อน
- การสร้างต้นไม้ตัดสินใจมีพื้นฐานมาจากวิธีการเลือก Attribute

เมื่อไหร่ถึงจะหยุดการสร้างต้นไม้

- เมื่อทุกข้อมูลใน node นั้นเป็น Class เดียวกัน
- เมื่อทุกข้อมูลใน node นั้นมีค่าของ Attribute เหมือนกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2 วิธีการที่จะนำมาสร้างต้นไม้ตัดสินใจ

	Splitting criterion	Pruning criterion	Other features
CART	➢ Gini index ➢ Twoing	Cross-validation post-pruning	➢ Regression/Classif. ➢ Nominal/Numeric attributes ➢ Missing values ➢ Oblique splits ➢ Nominal splits grouping
ID3	➢ Information Gain	Pre-pruning	➢ Classification ➢ Nominal attributes
C4.5 C5.0	➢ Information Gain ➢ Information Gain Ratio	Statistical based post-pruning	➢ Classification ➢ Nominal/numeric attributes ➢ Missing values ➢ Rule generator ➢ Multiple nodes split

- เกณฑ์ที่ใช้ในการเลือก Attribute

1. Gini index

$$\text{Gini}(t_i) = 1 - \sum_{i=1}^{N} [p(t_i)]^2$$

$$\text{Gini}_{\text{split}}(T) = \frac{N1}{N}\text{Gini}(t_1) + \frac{N2}{N}\text{Gini}(t_2) + \cdots + \frac{Nn}{N}\text{Gini}(t_n)$$

ตัวอย่าง หน้าถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

age	income	student	Credit_rating	Buys_computer
<= 30	High	No	Fair	No
<= 30	High	No	Excellent	No
31...40	High	No	Fair	Yes
>40	Medium	no	Fair	Yes
>40	low	Yes	Fair	Yes
>40	low	Yes	Excellent	No
31...40	low	Yes	Excellent	Yes
<= 30	Medium	No	Fair	No
<= 30	low	Yes	Fair	Yes
>40	Medium	Yes	Fair	Yes
<= 30	Medium	Yes	Excellent	Yes
31...40	Medium	No	Excellent	Yes
31...40	High	Yes	Fair	Yes
>40	Medium	No	Excellent	No

	N1	N2	N3
C0	2	4	3
C1	3	0	2

C0 = yes

C1 = no

$$\text{Gini}(t_1) = 1 - [\left(\frac{2}{5}\right)^2 + \left(\frac{3}{5}\right)^2] = 0.48$$

$$\text{Gini}(t_2) = 1 - [\left(\frac{4}{4}\right)^2 + \left(\frac{0}{4}\right)^2] = 0$$

$$\text{Gini}(t_3) = 1 - [\left(\frac{3}{5}\right)^2 + \left(\frac{2}{5}\right)^2] = 0.48$$

$$\text{Gini}_{\text{split}}(\text{Age}) = \frac{5}{14}(0.48) + \frac{4}{14}(0) + \frac{5}{14}(0.48) = 0.343$$

ดังนั้นจะได้ว่า

$\text{Gini}_{\text{split}}(\text{Age}) = 0.343$, $\text{Gini}_{\text{split}}(\text{Income}) = 0.491$, $\text{Gini}_{\text{split}}(\text{Student}) = 0.365$,

$\text{Gini}_{\text{split}}(\text{credit_rating}) = 0.426$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากการคำนวณการเลือก Attribute ที่มีค่า $\mathbf{Gini_{split}(T)}$ น้อยที่สุด มาใช้ในการพิจารณาเป็นราก (root) ของต้นไม้ตัดสินใจ จากนั้นทำการเลือก Attribute ต่อโดยตัดข้อมูลของข้อมูลที่พิจารณาก่อนหน้าออกไป ทำเช่นนี้ไปจนกระทั่งจนครบfeatureของตัวอย่างที่นำมาพิจารณา

1. Information Gain / Gain Ratio

โดยที่ Information Gain(D,A) = $\mathbf{Entropy(D) - \sum_{i=1}^{v} \frac{|D_i|}{|D|} x Entropy(D_i)}$

และ $\mathbf{Entropy(D) = \sum_{i=1}^{v} - \frac{|D_i|}{|D|} x \log_2 \frac{|D_i|}{|D|}}$

$$\mathbf{Gain\ Ratio(D,A) = \frac{Information\ Gain(D,A)}{Entropy(D,A)}}$$

ตัวอย่าง

age	income	student	Credit_rating	Buys_computer
<= 30	High	No	Fair	No
<= 30	High	No	Excellent	No
31...40	High	No	Fair	Yes
>40	Medium	no	Fair	Yes
>40	low	Yes	Fair	Yes
>40	low	Yes	Excellent	No
31...40	low	Yes	Excellent	Yes
<= 30	Medium	No	Fair	No
<= 30	low	Yes	Fair	Yes
>40	Medium	Yes	Fair	Yes
<= 30	Medium	Yes	Excellent	Yes
31...40	Medium	No	Excellent	Yes
31...40	High	Yes	Fair	Yes
>40	Medium	No	Excellent	No

หา entropy ตัวอย่าง $\mathbf{Entropy(D) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.940}$

$$\mathbf{Gain(D,income) = 0.940 - \frac{4}{14}\left(-\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4}\right)}$$
$$\mathbf{- \frac{6}{14}\left(-\frac{4}{6}\log_2\frac{4}{6} - \frac{2}{6}\log_2\frac{2}{6}\right) - \frac{4}{14}\left(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}\right) = 0.029}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เลือกพิจารณา income ก่อน โดยที่ income แบ่งออกเป็น 3 ระดับ คือ

low มีจำนวน 4 ตัวอย่าง จากทั้งหมด 14 ตัวอย่าง (yes : 3 จาก 4 , no : 1 จาก 4)

medium มีจำนวน 6 ตัวอย่าง จากทั้งหมด 14 ตัวอย่าง (yes : 4 จาก 6 , no : 2 จาก 6)

high มีจำนวน 4 ตัวอย่าง จากทั้งหมด 14 ตัวอย่าง (yes : 2 จาก 4 , no : 2 จาก 4)

หา SplitInformation (D,income) โดยที่ income มี 3 ระดับ คือ low , medium , high จำนวน 4 , 6 , 4 ตามลำดับ

SplitInformation (D,income) = $-\frac{4}{14}\log_2\frac{4}{14}-\frac{6}{14}\log_2\frac{6}{14}-\frac{4}{14}\log_2\frac{4}{14}=1.557$

$\text{Gain Ratio(D, income)} = \frac{0.029}{1.557} = 0.019$

ทำเช่นนี้ไปจนกระทั่งครบทุก feature และเลือก gain ratio มากที่สุดมาเป็นattribute ตามลำดับ

2.3 ตัววัดประสิทธิภาพของโมเดล Classification

1. Confusion Matrix

Confusion Matrix คือ การประเมินผลลัพธ์การทำนาย (หรือผลลัพธ์จากโปรแกรม) เปรียบเทียบกับผลลัพธ์จริงๆ ที่หาโดยค่า

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าจริง และคนบอกว่ามันจริง

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง และคนบอกว่ามันไม่จริง

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าจริง แต่คนบอกว่าไม่จริง

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง แต่คนบอกว่าจริง

ตารางที่ 3 ตารางConfusion Matrix

		Prediction	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. Precision

Precision คือ ค่าที่บอกว่าโปรแกรมทำนายว่าจริง ถูกต้องเท่าไร หาได้จาก TP/(TP+FP)

3. Recall

Recall (True Positive Rate) คือ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าจริง เป็นอัตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก TP/(TP+FN)

True Negative Rate (TNR) คือ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าไม่จริง เป็นอัตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก TN/(TN+FP)

False Positive Rate (TPR) คือ ค่าที่บอกว่าโปรแกรมทำนายว่าจริง เป็นอันตราส่วนเท่าไรของไม่จริงทั้งหมด หาได้จาก FP/(TN+FP)

False Negative Rate (FNR) คือ ค่าที่บอกว่าโปรแกรมทำนายว่าไม่จริง เป็นอันตราส่วนเท่าไรของจริงทั้งหมด หาได้จาก FN/(TP+FN)

4. Accuracy

Accuracy คือ ค่าที่บอกว่าโปรแกรมสามารถทำนายได้แม่นยำขนาดไหน

หาได้จาก (TP+TN)/(TP+TN+FP+FN)

2.4 ความรู้พื้นฐาน SQL

SQL มาจากคำว่า Structured Query Language เป็นภาษามาตรฐานในการเข้าถึงฐานข้อมูล เราสามารถใช้งานภาษา SQL ได้จากโปรแกรมต่างๆ ที่ต้องทำการกับระบบฐานข้อมูล เช่น ใช้ SQL ในการทำการดึงข้อมูล (Retrieve Data) จากฐานข้อมูล และมันเป็นมาตรฐานกลางที่ใช้ในระบบฐานข้อมูลต่างๆ โดยเป็นมาตรฐานของ ANSI (American National Standard Institute)

เราจะใช้ SQL ทำอะไรได้บ้าง

- ใช้ในการสืบค้นข้อมูลในฐานข้อมูล
- ใช้ใส่ข้อมูลเพิ่มเข้าไปยังฐานข้อมูล
- ใช้ปรับปรุงข้อมูลในฐานข้อมูล
- ใช้ลบรายการที่เราไม่ต้องการออกจากฐานข้อมูล
- ใช้สร้างฐานข้อมูลขึ้นมาใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ใช้สร้างตาราง(Table) ในฐานข้อมูล
- ใช้สร้าง Stored Procedure ในฐานข้อมูล
- ใช้สร้าง Views ในฐานข้อมูล
- ใช้กำหนดสิทธิ์ให้กับตาราง (Table), Procedure และ Views

โดยแม้ว่า SQL เป็นมาตรฐานแต่ก็ยังมีหลาย เวอร์ชั่น ซึ่งก็จะมีความแตกต่างกันออกไป ในแต่ละผลิตภัณฑ์อีกด้วย แต่โครงสร้างหลักในการ SELECT, INSERT INTO, UPDATE หรือ DELETE ก็จะมีโครงสร้างเดียวกัน

เราสามารถนำ SQL ไปใช้งานในระบบใดได้บ้าง

- ใช้กับเว็บไซต์ เพื่อแสดงผลข้อมูลจากฐานข้อมูล DBMS ไม่ว่าจะเป็น Microsoft Access, SQL Server, MySQL, Oracle
- ใช้ร่วมกับระบบฐานข้อมูล RDBMS ไม่ว่าจะเป็น MS SQL Server, IBM DB2, Oracle, MySQL และ Microsoft Access
- ใช้ในการกำหนดในระบบวิเคราะห์ข้อมูล (Analysis Tools) ที่เปิดช่องให้เราสามารถทำการใส่ หรือ ปรับปรุง SQL ได้ด้วยตัวเอง

ดังนั้นหากเราสามารถใช้งาน SQL ได้ก็ย่อมจะเป็นประโยชน์ในการทำความเข้าใจในการทำงานกับระบบฐานข้อมูลมากยิ่งขึ้น อ่านความหมายของสิ่งที่ใช้งานได้อย่างลึกซึ่งขึ้นอีกด้วย

การเรียกข้อมูลจากตารางสามารถทำได้หลายรูปแบบทั้งที่มีเงื่อนไขหรือแม้แต่การเรียกดูข้อมูลจากหลายตาราง

- รูปแบบของคำสั่ง SELECT เพื่อใช้เรียกดูข้อมูลทุกคอลัมน์จากตารางใดตารางหนึ่ง

รูปแบบ

```
Select *
FROM <table name>;
```

ในการเรียกดูข้อมูลทั้งตาราง จะใช้เครื่องหมาย * เพื่อแทนการเรียกดูข้อมูลจากทุกคอลัมน์ของตาราง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การเรียกดูข้อมูลเพียงบางคอลัมน์

รูปแบบคำสั่ง SELECT เพื่อใช้เรียกดูข้อมูลเพียงบางคอลัมน์จากตารางใดตารางหนึ่ง

รูปแบบ

```
Select   <column 1 , column 2,...>
FROM <table name>;
```

ในการเรียกดูข้อมูลเพียงบางคอลัมน์ จำเป็นต้องระบุชื่อคอลัมน์ให้ถูกต้อง ตามที่ระบุไว้ในโครงสร้าง และหากเรียกดูมากกว่าหนึ่งคอลัมน์ จะต้องมีเครื่องหมาย , คั่นระหว่างชื่อคอลัมน์

- การเรียกดูข้อมูลโดยไม่ให้แสดงข้อมูลซ้ำกัน

เมื่อต้องการดูข้อมูลจากตารางโดยไม่ให้แสดงข้อมูลที่ซ้ำซ้อนออกมา จะใช้คำว่า DISTINCT กำกับ หลังคำว่า SELECT

- การเรียกดูข้อมูลโดยแสดงผลข้อมูลของคอลัมน์เป็นนิพจน์ทางคณิตศาสตร์

ในบางครั้งการเรียกดูข้อมูลของบางคอลัมน์ (Column list) อาจจะต้องทำการคำนวณประกอบก่อน เพื่อแสดงผลข้อมูลออกมา

- การเรียกดูข้อมูลโดยการจัดเรียงข้อมูล

ข้อมูลที่ถูกเรียกสามารถถูกจัดเรียงตามคอลัมน์ที่ต้องการ เช่นให้เรียงชื่อตามตัวอักษร หรือให้เรียงข้อมูลตามเงินเดือน เป็นต้น ในการจัดเรียงข้อมูลอาจจะเรียงจากน้อยไปหามาก (Ascending) หรือจากมากไปหาน้อย (Descending)

รูปแบบคำสั่งที่ใช้แสดงผลลัพธ์โดยการจัดเรียงข้อมูล คือ

รูปแบบ

```
SELECT [*| DISTINCT] <column 1, ...>
FROM <table name>
[ORDER BY <column 1,...>][ASC][DESC]];
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- การเรียกดูข้อมูลแบบมีเงื่อนไข

การเรียกดูข้อมูลแบบมีเงื่อนไข เป็นการระบุค่าเฉพาะของข้อมูลที่ต้องการเรียกดู อาจจะใช้เงื่อนไขเพื่อดึงข้อมูลบางแถวจากตาราง เช่น

การเรียกดูข้อมูลแบบมีเงื่อนไข จะใช้วลีหรืออนุประโยค WHERE ต่อท้ายวลีหรืออนุประโยค FROM รูปแบบของคำสั่งเป็นดังนี้

รูปแบบ

```
SELECT [ * | DISTINCT] <column 1, column 2, ...>
FROM <table name>
[WHERE <CONDITION>];
```

ในอนุประโยค WHERE ประกอบด้วยส่วนสำคัญ 3 ส่วน คือ

1. ชื่อคอลัมน์
2. โอเปอเรเตอร์การเปรียบเทียบ ซึ่งแบ่งออกเป็นโอเปอเรเตอร์ทางตรรกะ (Logical Operator) โอเปอเรเตอร์ SQL รวมถึงการใช้โอเปอเรเตอร์บูลีน
3. ข้อมูลเฉพาะที่ต้องการเป็นเงื่อนไขของชื่อคอลัมน์ที่ระบุในข้อ 1 อาจเป็นค่าคงที่ (Constant) กลุ่มของข้อมูลหรือนิพจน์ (Expression) หรือชื่อคอลัมน์อื่นที่ต้องการนำมาเปรียบเทียบข้อมูลเฉพาะที่เป็นประเภทตัวอักษรหรือวัน เดือน ปี (DATE) เมื่อนำมาเป็นเงื่อนไขเฉพาะจะต้องมีเครื่องหมาย ‘ ‘ กำกับ

- การเรียกดูข้อมูลแบบมีเงื่อนไข โดยใช้โอเปอเรเตอร์ของ SQL

โอเปอเรเตอร์ของ SQL ที่ใช้ในการกำหนดเงื่อนไขของข้อมูลในอนุประโยค WHERE มีดังต่อไปนี้

4.1 BETWEEN ... AND ... เป็นโอเปอเรเตอร์ที่กำหนดเงื่อนไขของคอลัมน์เป็นค่าระหว่างค่าสองค่า การใช้ BETWEEN ... AND ...นี้ จะแสดงต่อท้ายชื่อคอลัมน์ที่ถูกระบุให้เป็นเงื่อนไข

4.2 IN เป็นโอเปอเรเตอร์ที่ใช้กับเงื่อนไขของคอลัมน์ที่ต้องการระบุเงื่อนไขเป็นกลุ่มของข้อมูล โดย IN จะแสดงต่อท้ายชื่อคอลัมน์ที่ถูกระบุเป็นเงื่อนไข และกลุ่มของข้อมูลที่เป็นข้อมูลเฉพาะของคอลัมน์ที่เป็นเงื่อนไขนี้ จะระบุอยู่ในวงเล็บ และมีเครื่องหมาย , คั่น

4.3 LIKE เป็นโอเปอเรเตอร์ที่ใช้ในการค้นหาข้อมูลของคอลัมน์ที่เก็บข้อมูลประเภทตัวอักษรเท่านั้น โดยยังไม่ทราบค่าที่แน่นอนทั้งหมดของข้อมูลที่จะค้นหา หรือรู้เพียงบางตัวอักษรเท่านั้น โอเปอเรเตอร์ LIKE จะระบุต่อท้ายชื่อคอลัมน์ที่เป็นเงื่อนไขโดยจะใช้สัญญลักษณ์ที่ช่วยในการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค้นหาข้อมูลคือ * และ ? จะต้องมีเครื่องหมาย '' กำกับเสมอ ความหมายของสัญลักษณ์ทั้งสองคือ

*	ใช้แทนจำนวนตัวอักษรได้หลายตัว
?	ใช้แทนจำนวนตัวอักษรที่ไม่ทราบค่า 1 ตัว

4.4 IS NULL เป็นโอเปอเรเตอร์ที่ใช้ในการแสดงค่าของคอลัมน์ที่มีค่าเป็นค่าว่าง หรือไม่มีค่า นอกจากโอเปอเรเตอร์ของ SQL ทั้ง 4 ที่กล่าวมาแล้วข้างต้นโอเปอเรเตอร์เหล่านี้ ยังสามารถใช้เป็นเงื่อนไขในเชิงปฏิเสธโดยใช้คำว่า NOT นำหน้า

- การเรียกดูข้อมูลแบบมีหลายเงื่อนไข โดยใช้โอเปอเรเตอร์boolean

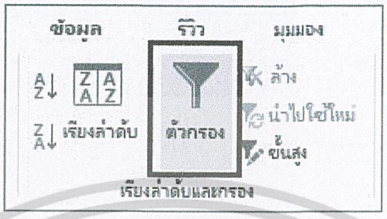
การเรียกดูข้อมูลที่มีเงื่อนไขมากกว่าหนึ่งเงื่อนไข สามารถใช้โอเปอเรเตอร์บูลีนเป็นตัวเชื่อมโยงเงื่อนไขดังกล่าว โอเปอเรเตอร์บูลีน ประกอบด้วย

AND	ใช้เชื่อมเงื่อนไข โดยข้อมูลที่จะแสดงออกมาจะต้องเป็นจริงตามเงื่อนไขทั้งสอง
OR	ใช้เชื่อมเงื่อนไขสองเงื่อนไข โดยข้อมูลที่จะแสดงออกมา จะเป็นจริงตามเงื่อนไขใดเงื่อนไขหนึ่ง
NOT	ใช้แสดงหน้าเงื่อนไขใดเงื่อนไขหนึ่ง เพื่อให้ได้ข้อมูลที่ไม่เป็นตามเงื่อนไขที่ระบุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

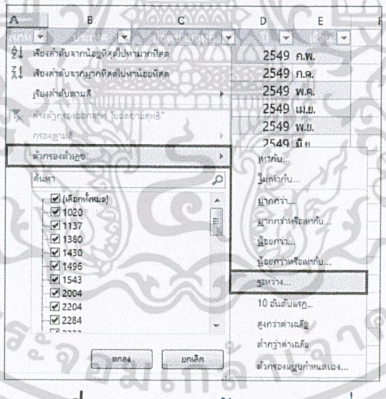
2.5 ความรู้พื้นฐาน Microsoft Excel

- การกรองข้อมูล(filter)
  - เลือกข้อมูลที่คุณต้องการกรอง เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด คอลัมน์ควรจะมีส่วนหัวด้วย



ภาพที่ 3 การกรองข้อมูลลำดับที่1

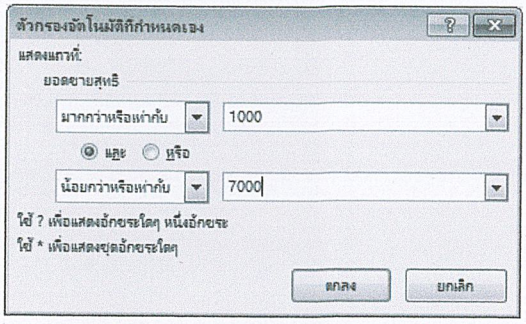
  - คลิกลูกศร ในส่วนหัวของคอลัมน์ แล้วคลิก ตัวกรองข้อความ หรือ ตัวกรองตัวเลข
  - คลิกตัวดำเนินการเปรียบเทียบตัวใดตัวหนึ่ง ตัวอย่างเช่น ถ้าต้องการแสดงตัวเลขที่อยู่ภายในขีดจำกัดล่างและบน ให้เลือก ระหว่าง



ภาพที่ 4 การกรองข้อมูลลำดับที่2

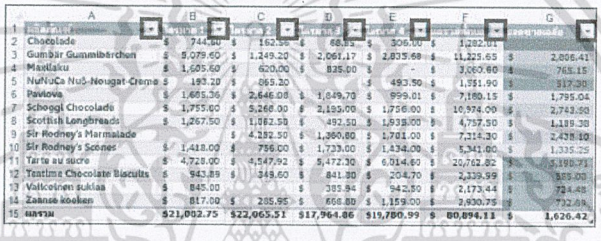
  - ในกล่องตัวกรองอัตโนมัติแบบกำหนดเอง พิมพ์ หรือเลือกเกณฑ์การกรองข้อมูลของคุณ ตัวอย่างเช่น เพื่อแสดงตัวเลขทั้งหมดระหว่าง 1000 และ 7000 ในมีค่ามากกว่าหรือเท่ากับ กล่อง ชนิด1000 และมีค่าน้อยกว่า หรือเท่ากับกล่อง พิมพ์7000

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 5 การกรองข้อมูลลำดับดับที่3

- คลิก ตกลง เพื่อใช้ตัวกรอง
- กรองข้อมูลในตาราง เมื่อคุณใส่ข้อมูลในตาราง ตัวควบคุมการกรองจะถูกเพิ่มที่ส่วนหัวของตารางโดยอัตโนมัติ

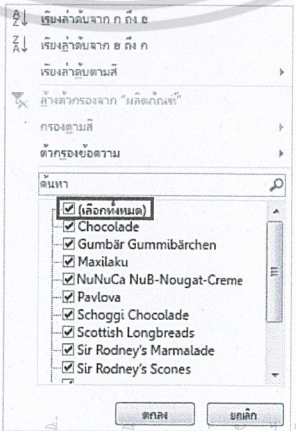


	A	B	C	D	E	F	G
1	ผลิตภัณฑ์	ยอดขาย	ต้นทุน	กำไร	ส่วนลด	ยอดขายสุทธิ	ค่าขนส่ง
2	Chocolade	$ 744.60	$ 162.56	$ 68.12	$ 306.00	$ 1,282.01	
3	Gumbär Gummibärchen	$ 5,079.60	$ 1,249.20	$ 2,061.17	$ 2,835.68	$ 11,225.65	$ 2,806.41
4	Maxilaku	$ 1,695.60	$ 620.00	$ 825.00	$ -	$ 3,060.60	$ 765.15
5	NuNuCa Nuß-Nougat-Creme	$ 187.20	$ 965.20	$ -	$ 493.50	$ 1,551.90	$ 317.30
6	Pavlova	$ 1,665.36	$ 2,646.06	$ 1,849.70	$ 999.01	$ 7,180.15	$ 1,795.04
7	Schoggi Chocolade	$ 1,755.00	$ 5,268.00	$ 2,193.00	$ 1,756.00	$ 10,974.00	$ 2,743.50
8	Scottish Longbreads	$ 1,267.50	$ 1,062.50	$ 492.50	$ 1,938.00	$ 4,757.50	$ 1,189.38
9	Sir Rodney's Marmalade		$ 4,282.50	$ 1,360.80	$ 1,701.00	$ 7,314.30	$ 2,438.10
10	Sir Rodney's Scones	$ 1,418.00	$ 756.00	$ 1,733.00	$ 1,434.00	$ 5,341.00	$ 1,335.25
11	Tarte au sucre	$ 4,728.00	$ 4,547.92	$ 5,472.30	$ 6,014.60	$ 20,762.82	$ 5,190.71
12	Teatime Chocolate Biscuits	$ 943.89	$ 349.60	$ 841.80	$ 204.70	$ 2,339.99	$ 585.00
13	Valkoinen suklaa	$ 845.00		$ 385.94	$ 942.50	$ 2,173.44	$ 734.48
14	Zaanse koeken	$ 817.00	$ 285.95	$ 666.80	$ 1,159.00	$ 2,930.75	$ 732.69
15	ผลรวม	$21,002.75	$22,065.51	$17,964.86	$19,780.99	$ 80,894.11	$ 1,626.42

ภาพที่ 6 การกรองข้อมูลลำดับดับที่4

เพื่อการกรองอย่างรวดเร็ว ทำดังต่อไปนี้

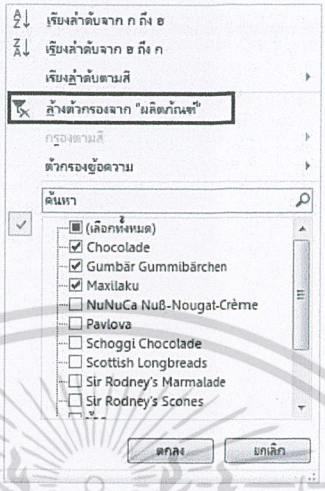
- คลิกที่ลูกศร ที่ส่วนหัวของตารางของคอลัมน์ที่คุณต้องการกรอง
- ในรายการข้อความหรือตัวเลข ให้ยกเลิกการเลือกกล่อง (เลือกทั้งหมด) ตรงด้านบนสุดของรายการ จากนั้นเลือกกล่องรายการที่คุณต้องการให้แสดงในตาราง



ภาพที่ 7 การกรองข้อมูลลำดับดับที่5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

➢ คลิก ตกลง ลูกศรตัวกรองที่อยู่ในส่วนหัวของตารางจะเปลี่ยนเป็นไอคอนนี้ เพื่อแสดงว่ามีการกรอง คลิกเพื่อเปลี่ยนหรือล้างการกรอง



ภาพที่ 8 การกรองข้อมูลลำดับดับที่6

- การPivot Table

PivotTable คืออีกหนึ่งเครื่องมือระดับสูงที่มีอยู่ในโปรแกรม Microsoft Excel ซึ่งมีมาตั้งนานแล้ว ใช้สำหรับการสรุปข้อมูลจำนวนมาก รองรับการแสดงลัพธ์ในรูปแบบของกราฟ หรือข้อมูลตารางที่มีการสรุปผลลัพธ์ สามารถเลือกรูปแบบของรายงานได้หลากหลาย ทำให้ลดเวลาในการจัดการข้อมูลได้มาก การใช้งานไม่ยุ่งยาก เพียงแค่เริ่มต้นด้วยการบันทึกข้อมูล และเลือกคำสั่ง PivotTable จากนั้น เราก็เลือกรูปแบบของข้อมูล และรายงานตามที่เราต้องการเท่านั้น

วิธีการใช้งาน PivotTable สำหรับการเริ่มต้น

1. เปิดเอกสารหรือ sheet ที่มีข้อมูล

2. คลิกคลุมข้อมูลทั้งหมด

3. คลิกเมนู Insert

4. คลิกเลือกคำสั่ง PivotTable

5. จะพบหน้าต่างใหม่ Create PivotTable (จะแสดงว่าเราต้องการสร้างจากข้อมูลไหนบ้าง)

6. ที่คำสั่ง Choose where you want the PivotTable report to be placed

- เลือก New Worksheet ถ้าต้องการให้สร้าง Sheet ใหม่ (แนะนำให้เลือกข้อนี้)

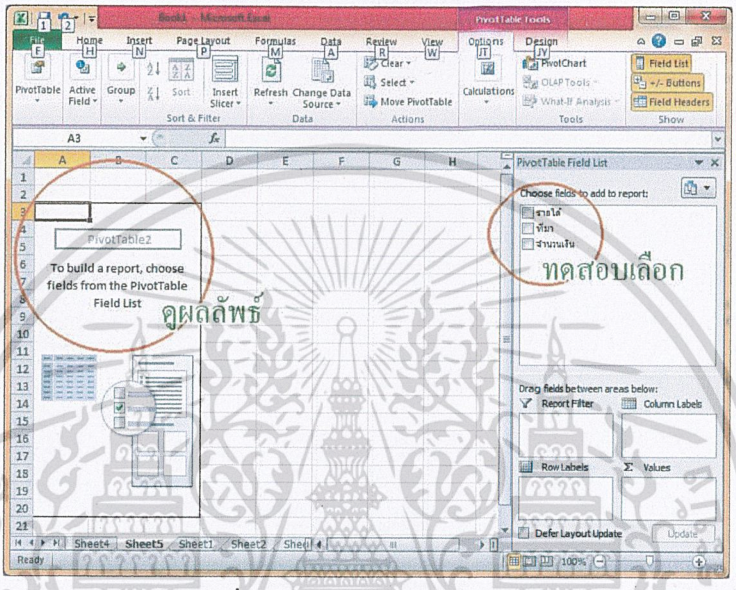
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เลือก Existing Worksheet ถ้าต้องการให้แสดงผลลัพธ์ใน Sheet ที่ต้องการ

7. คลิก OK ได้เลย เพราะเราได้ทำการเลือกข้อมูลไว้แล้ว ตั้งแต่ตอนต้น

8. ดูผลลัพธ์

9. จะพบว่า ที่คอลัมภ์ขวามือ ให้คลิกทุก Fields ดูผลลัพธ์ทางซ้ายมือที่ได้อีกครั้ง



ภาพที่ 9 การกรองข้อมูลลำดับดับที่7

2.6 R Studio

ภาพที่ 10 โปรแกรม R Studio

R เป็นภาษาคอมพิวเตอร์ภาษาหนึ่งที่เป็นที่ใช้คำนวณทางสถิติ เพราะเป็นโปรแกรมโอเพนซอร์ท (open source) ที่ใช้งานได้ฟรี รองรับระบบปฏิบัติการหลายระบบ ปฏิบัติการทั้ง Windows Mac OS หรือ Linux โดยการทำงานของ R ที่นิยมเนื่องจากมี built-in function ทางด้านสถิตที่มีปริมาณมาก รวมถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่ และ มีความสามารถแสดงผลการทำงานในรูปแบบกราฟ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

R เป็นภาษาคอมพิวเตอร์ภาษาหนึ่งที่เป็นที่ใช้คำนวณทางสถิติ เพราะเป็นโปรแกรมโอเพนซอร์ท (open source) ที่ใช้งานได้ฟรี รองรับระบบปฏิบัติการหลายระบบ ปฏิบัติการทั้ง Windows Mac OS หรือ Linux โดยการทำงานของ R ที่นิยมเนื่องจากมี built-in function ทางด้านสถิตที่มีปริมาณมาก รวมถึง สถิติที่ใช้ในการวิเคราะห์ข้อมูลขนาดใหญ่ และ มีความสามารถแสดงผลการทำงานในรูปแบบกราฟ แต่จะพบว่าการพัฒนาโปรแกรมด้วยภาษา R อาจจะไม่ปรากฏเครื่องมือในการ (tools) ช่วยเหลือการทำงานมากนัก แต่ R สามารถประมวลผลได้รวดเร็ว จึงได้รับความนิยมโดยเฉพาะการใช้งานเพื่อวิเคราะห์ data sciences และ big data (ภาควิชาภาษาศาสตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย)

การติดตั้งโปรแกรม

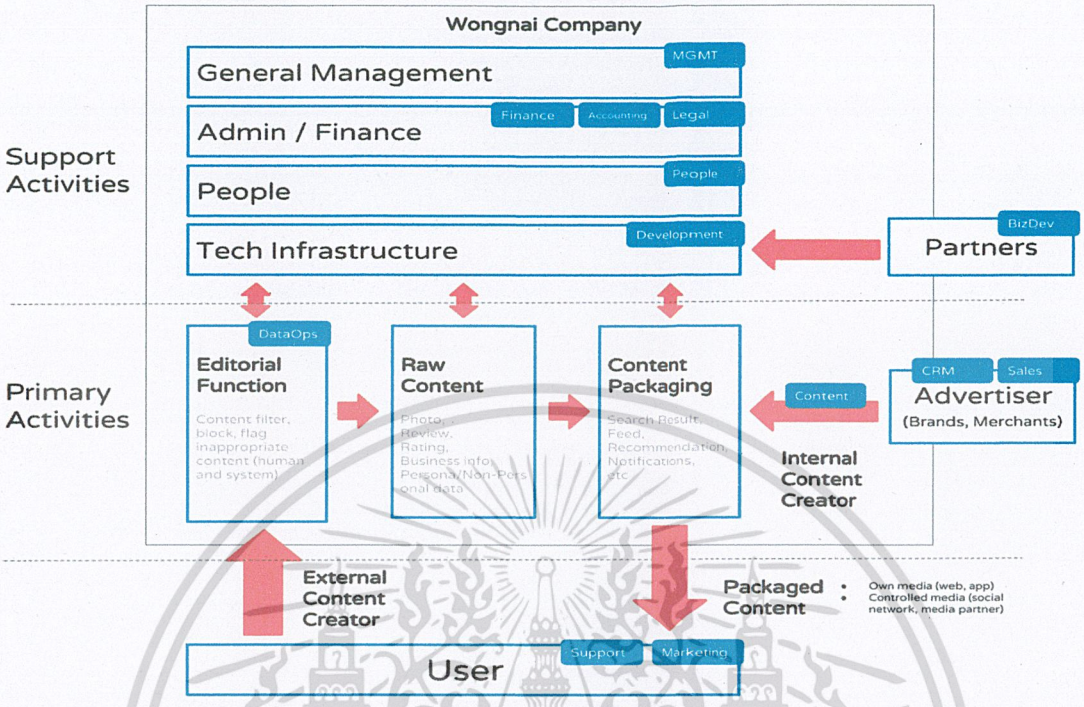
- ในขั้นตอนการติดตั้งโปรแกรม เริ่มต้นจากการดาวน์โหลดโปรแกรมจากเว็บไซต์ https://cran.r-project.org/ และเลือกระบบปฏิบัติการที่ต้องการติดตั้ง
- เมื่อดาวน์โหลดแล้วคลิกเลือกไฟล์ที่ดาวน์โหลด แล้วเลือกภาษาที่ต้องการติดตั้ง
- ระบบจะแสดงหน้าจอการติดตั้งโปรแกรม R กด next และอ่านเงื่อนไข เพื่อเลือกไฟล์ที่ต้องการติดตั้ง และกำหนดพาธ ที่ต้องการติดตั้งโปรแกรมตามต้องการ
- รอการติดตั้ง และกดปุ่ม Finish เมื่อการติดตั้งเสร็จสิ้น

ชนิดของข้อมูล

ในการพัฒนาโปรแกรมสิ่งแรกที่จำเป็นต้องเรียนรู้คือชนิดของตัวแปรที่ใช้ในการเก็บข้อมูลต่าง ๆ โดยตัวแปรในการเก็บข้อมูลนั้นโปรแกรมสงวนที่ในหน่วยความจำเพื่อเก็บค่าของข้อมูล ดังนั้นเมื่อสร้างตัวแปรจะเป็นจองพื้นที่บางส่วนในหน่วยความจำไว้บางส่วน โดยชนิดของตัวแปรหลัก ๆ ในภาษา R ได้แก่ character, wide character, integer, floating point, double floating point, Boolean โดยการประกาศตัวแปรในภาษา R จะแตกต่างจากภาษา c และภาษาจาวา นอกจากนั้นในการพัฒนาโปรแกรมด้วยภาษา R นั้นยัง มีตัวแปรชนิดวัตถุ ที่แตกต่างจากภาษาข้างต้นอยู่หลายชนิด ซึ่งตัวแปรชนิดวัตถุของภาษา R ประกอบด้วย Vectors, Lists, Matrices, Arrays,Factors,Data Frames

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 โครงสร้างองค์กร



ภาพที่ 11 โครงสร้างองค์กร

Support Activities (ไล่จากล่างขึ้นบน)

- Development team (ดูแล Tech Infrastructure): เรียกได้ว่าเป็น “กระดูกสันหลัง” ของ Wongnai ซึ่งเป็น tech company เลยทีเดียว หน้าที่ของแผนกนี้ คือ พัฒนา product และ service ให้ตอบโจทย์ของทุกคนที่มีส่วนร่วม (สมาชิก, ลูกค้า, พาร์ทเนอร์ และตัวบริษัท Wongnai เอง) ซึ่ง development team นี้ เราแบ่งย่อยตาม “ความชำนาญ” ได้เป็น Backend Developer, Frontend Developer, Mobile App Developer, Quality Assurance, UX Engineer, UI Designer,Architect และ Data Scientist

- People team (ดูแล People): ที่ Wongnai เราไม่มีแผนก HR (human-resources) เพราะเราไม่ได้มองพนักงานเป็นแค่ทรัพยากร (resource) แต่เรามองพนักงานของเราเป็น “คน” ซึ่งเป็นส่วนสำคัญที่สุดขององค์กร ดังนั้นเราจึงให้ชื่อทีมที่ดูแล “คนของเรา” ว่า “People team” ซึ่งมีหน้าที่หลัก 3 ประการ คือ สรรหาคนเก่งๆ มาร่วมงาน (Recruit), ดูแลคนของเราให้เดินหน้าไปด้วยกันอย่างมีความสุข (Engage), และพัฒนาคนของเราให้เก่งขึ้นเรื่อยๆ (Grow) เพื่อให้บริษัทมีความสามารถในการแข่งขันได้ในระยะยาว ทีมนี้งานแน่นตลอดเวลา เพราะว่า Wongnai รับคนเพิ่มอยู่ตลอดเพื่อรองรับการเติบโตของบริษัททำให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ต้องสรรหาคนจากหลายๆ ช่องทาง ใช้เครื่องมือ กลยุทธ์สารพัด เพื่อช่วงชิงตัวคนเก่งๆ มาร่วมงานกับเรา และในส่วนการดูแลและพัฒนาคนของเราก็มีทั้งบริหารการจัดการเรื่องสวัสดิการต่างๆ (ขนม-เครื่องดื่มไม่อั้น, ข้าวกลางวันเสิร์ฟที่ออฟฟิศไม่ต้องออกไปเดินหากินเองร้อนๆ, กิจกรรม WeLearn/WeShare ที่สนับสนุนให้พนักงานเรียนรู้ และแบ่งปันอย่างต่อเนื่อง ผ่านงบซื้อหนังสือที่ได้ทุกเดือน โดยมีข้อแม้เพียงแค่ว่าต้องอ่านแล้วเอามาสรุปให้เพื่อนๆ ฟัง รวมถึงการเชิญแขกคนเก่งๆ จากภายนอกหลายวงการมาให้ความรู้ทุกวันศุกร์ และอื่นๆ อีกมาก)

- Finance / Account / Legal team (ดูแล 3 เรื่องตามชื่อทีม): ขอรวม 3 ส่วนนี้เข้าด้วยกัน เนื่องจากตอนนี้ทีมที่ดูแลส่วนนี้เป็นกลุ่มเดียวกัน ซึ่งทีมนี้มีหน้าที่ตามชื่อเลย คือ ดูแลเรื่องการเงิน บัญชี และกฎหมาย ทั้งเงินเข้า เงินออก การลงบัญชีต้องผ่านการตรวจสอบที่เข้มงวดของทีมนี้ให้ถูกต้อง ซึ่งเราก็พยายามพัฒนาให้ได้มาตราฐานดีขึ้นเรื่อยๆ เนื่องจาก Wongnai วางแผนไว้ว่าจะเข้าตลาดหุ้น (IPO) ในปี 2019 นี้

- Management (ดูแลการบริหาร): มีหน้าที่ในการบริหารองค์กรในแง่มุมต่างๆ เช่น ทำให้องค์กรเติบโต และแข่งขันได้อย่างยั่งยืน (CEO), ทำให้องค์กรใช้ technology ให้ตอบโจทย์ของธุรกิจได้สูงที่สุด (CTO), ทำให้องค์กรดำเนินงานได้อย่างราบรื่น (COO)

Primary Activities (ไล่จากซ้ายไปขวา)

- Data Operations team (ดูแลข้อมูล): หน้าที่หลักแบบเต็มๆ คือ “ทำให้ข้อมูลของ Wongnai มีประโยชน์ต่อผู้ใช้งานมากที่สุด” หมายถึง ข้อมูลต้องมีคุณภาพที่ดี (ถูกต้อง ครบถ้วน และทันสมัย) และปริมาณที่มากพอ (ครอบคลุม) ทีมนี้เป็นทีม “หลังบ้าน” ที่คอยกลั่นกรองข้อมูลเพื่อสร้างประสบการณ์การใช้งานที่ดีที่สุด (นอกจาก product ดีแล้ว ข้อมูลก็ต้องดีด้วย) การทำงานของทีมนี้แบ่งเป็นสองรูปแบบ คือ

  - เชิงรับ: เป็นการจัดการกับข้อมูลที่ส่งเข้ามาโดย user ของ Wongnai (ซึ่งเป็นข้อมูลส่วนใหญ่ของ Wongnai กว่า 99%) โดยทีมนี้จะต้องกรองข้อมูลที่ไม่เหมาะสม ผิด หรือซ้ำออกไปจากระบบ

  - เชิงรุก: เป็นการเดินหน้าหาข้อมูลที่ “น่าสนใจ แต่ยังขาดอยู่” มาเพิ่มลงในฐานข้อมูล ซึ่งก็เป็นการทำงานเสริมกับข้อมูลส่วนที่ user ของ Wongnai ส่งเข้ามา เพื่อให้ฐานข้อมูลครบถ้วน และทันสมัยที่สุดตามเป้าที่วางไว้

  คุณสมบัติของพนักงานแผนกนี้: ความละเอียดรอบคอบ รวดเร็ว ไหวพริบ และ ความเป็น perfectionist

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Content Editor team (ดูแลเนื้อหา): มีหน้าที่ผลิตเนื้อหาในรูปแบบต่างๆ ทั้งบทความ รูปภาพ และวิดีโอ ต้องคิดหัวข้อให้น่าสนใจ เนื้อหาแน่น แต่สนุกและย่อยง่าย มีประโยชน์ใช้งานได้จริง เช่น 10 ร้านกาแฟที่ต้องเช็กอิน, สอนทำอาหารแบบง่ายๆ, สอนแต่งหน้าไปทำงานในวันเร่งด่วน, ออกกำลังกายเป็นคู่สนุกกว่า และอื่นๆ ที่น่าจะเคยเห็นผ่านตาใน social network ซึ่งก็ทำขึ้นเพื่อจุดประสงค์ต่างๆ เช่น ทำให้คนรู้จัก Wongnai คุ้นเคย และรู้ว่า Wongnai มีข้อมูลอะไรให้ใช้บ้าง (ร้านอาหาร, สอนทำอาหาร, สั่งอาหาร และร้านบิวตี้) รวมถึงอีกมุมหนึ่งก็เป็นการผลิตเนื้อหาโฆษณา (advertorial) เพื่อตอบโจทย์ของลูกค้าของ Wongnai ในอีกทาง
  คุณสมบัติของพนักงานแผนกนี้: ทักษะในการผลิตเนื้อหา เช่น ถ่ายรูป วีดีโอ การเขียนบทความ กราฟฟิก ความคิดสร้างสรรค์ และต้องเป็นคนที่ติดตามข่าวสาร และเทรนด์ต่างๆ ตลอดเวลา

- Sales team (ดูแลลูกค้า ในเชิงการขาย): เปรียบได้ดังท่อน้ำเลี้ยงขององค์กร แน่นอนว่าตามชื่อเลยว่าหน้าที่หลัก คือ การขาย ซึ่งก็เริ่มตั้งแต่การหาลูกค้า และนัดเข้าไปหาเพื่อนำเสนอและปิดการขายเพื่อหารายได้ให้กับ Wongnai มาเลี้ยงพนักงานทุกคนในบริษัท
  คุณสมบัติของพนักงานแผนกนี้: มีทักษะการสื่อสาร และเจรจาต่อรองที่ดี ทักษะในการขาย และชอบพบปะพูดคุยกับผู้คน

- Customer Relationship Management team (ดูแลลูกค้า ในเชิงการสร้างความสัมพันธ์): ทำหน้าที่ให้บริการลูกค้า ทำงานควบคู่กับ Sales team อย่างใกล้ชิด เพื่อดูแลงานของลูกค้าให้เรียบร้อย และราบรื่น โดยมีเป้าหมายหลัก คือ ช่วยทำให้ธุรกิจของลูกค้าประสบความสำเร็จเมื่อใช้บริการของ Wongnai รวมถึงต้องสร้างความพึงพอใจ และความเชื่อมั่นให้แก่ลูกค้าอีกด้วย
  คุณสมบัติของพนักงานแผนกนี้: มีจิตใจรักในการให้บริการ เข้าใจลูกค้า ละเอียด รวดเร็ว และมีทักษะการสื่อสารที่ดี

- Support team (ดูแล users ในเชิงการให้ความช่วยเหลือ): เป็นแผนกที่คอยรับฟัง feedback หรือคำถามจากผู้ที่ใช้บริการ Wongnai ซึ่งก็ได้แก่ สมาชิก และร้านค้าต่างๆ จากนั้นจึงให้ข้อมูล คำตอบ ประสานงาน หรือช่วยแก้ไขปัญหาให้ได้อย่างรวดเร็วและถูกต้องที่สุด โดยมุ่งให้ผู้ใช้งานได้รับประสบการณ์ที่ดีในใช้บริการของ Wongnai
  คุณสมบัติของพนักงานแผนกนี้: มีจิตใจรักในการให้บริการ มีทักษะการสื่อสาร การแก้ไขปัญหา ไหวพริบ และมีความไวในการทำงานที่ดี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Marketing team (ดูแล user ในเชิงการสร้างการเติบโต): "ทำให้ Wongnai โต" เป็นเป้าหมายหลักของทีม marketing ซึ่งก็จะมีหลายตัววัดการเจริญเติบโตที่ Wongnai ใช้อยู่ ได้แก่ จำนวนผู้ใช้งาน (users), อัตราการใช้งานซ้ำ (monthy active users), ปริมาณการใช้งาน feature ต่างๆ เช่น รีวิว อัพโหลดรูป และอื่นๆ นอกจากนี้ยังจะต้องช่วยหาข้อมูลเชิงลึก (insight) ใหม่ๆ จากการทดสอบ หรือสัมภาษณ์ users เพื่อเป็นข้อมูลให้ development team ใช้ในการพัฒนาผลิตภัณฑ์ โดยสำหรับ marketing team ที่ Wongnai เราได้แบ่งทีมย่อยตามความถนัดออกมากเป็นดังนี้

  - Online marketing team: ดูแลสื่อภายในของ Wongnai ได้แก่ เว็บและแอพที่ใช้ในการสื่อสารกับ user, รวมถึงงานเชิง performance marketing ที่ต้องใช้การคิดวิเคราะห์ทดสอบ ปรับปรุงผล เช่น Web/App Analytics, Search Engine Optimization, Search Engine Marketing (AdWords)
  - Social Media team: ดูแลสื่อข้างนอกที่ Wongnai ใช้งานอยู่ เช่น Facebook, Instagram, Twitter และ LINE@ ทั้งในแง่การคิด จัดการเนื้อหาโพส การสร้าง engagement และการทำโฆษณาใน Social Media
  - Offline marketing & PR team: ดูแลกิจกรรมสมาชิก กิจกรรมการตลาด offline / on ground ต่างๆ รวมถึงช่วยดูแลภาพลักษณ์ของ Wongnai ที่ออกไปทางสื่อต่างๆ (กิจกรรมล่าสุด คือ งาน event อาหาร "Wongnai City of Food 2017 presented by Mali" ซึ่งจัดเป็นครั้งแรก และได้รับการตอบรับเป็นดีมาก)

  คุณสมบัติของพนักงานแผนกนี้: มีความเข้าใจในผู้ใช้งาน สภาพตลาด ตัวผลิตภัณฑ์ มีทักษะในการสื่อสารที่ดี ช่างสังเกต ปรับตัวเร็ว ชอบติดตามเทคโนโลยี ข่าวสาร เทคนิคและชอบทดลองเครื่องมือการตลาดใหม่ๆ

- Business Development team (ดูแลพันธมิตร พัฒนาธุรกิจ): มีหน้าที่ในการสร้างรูปแบบธุรกิจใหม่ๆ รวมถึงสรรหา และสร้างความร่วมมือกับ partner เพื่อสร้างการเติบโตให้กับบริษัทอย่างต่อเนื่อง ผลงานล่าสุดของแผนกนี้ ได้แก่ การทำงานร่วมกับ partner อย่าง LINE MAN ในการให้บริการสั่งอาหารออนไลน์ที่ใหญ่ที่สุดในไทย ด้วยฐานข้อมูลร้านอาหารกว่า 20,000 แห่งที่คัดมาแล้วของ Wongnai หรือการร่วมมือกับ Alipay ในการดูแลนักท่องเที่ยวชาวจีนที่เข้ามาเที่ยวในประเทศไทยปีละกว่า 10 ล้านคนให้เข้าถึงข้อมูลร้านอาหารของไทยในฐานข้อมูลของ Wongnai รวมไปถึงการชำระเงินด้วยระบบ Alipay ได้อีกด้วย

  คุณสมบัติของพนักงานแผนกนี้: มีทักษะในคิดวิเคราะห์เชิงธุรกิจ การวิจัยตลาด ความเข้าในผลิตภัณฑ์อย่างลึกซึ้ง มีทักษะการเจรจาต่อรอง ไหวพริบ และการสื่อสารที่ดีเยี่ยม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

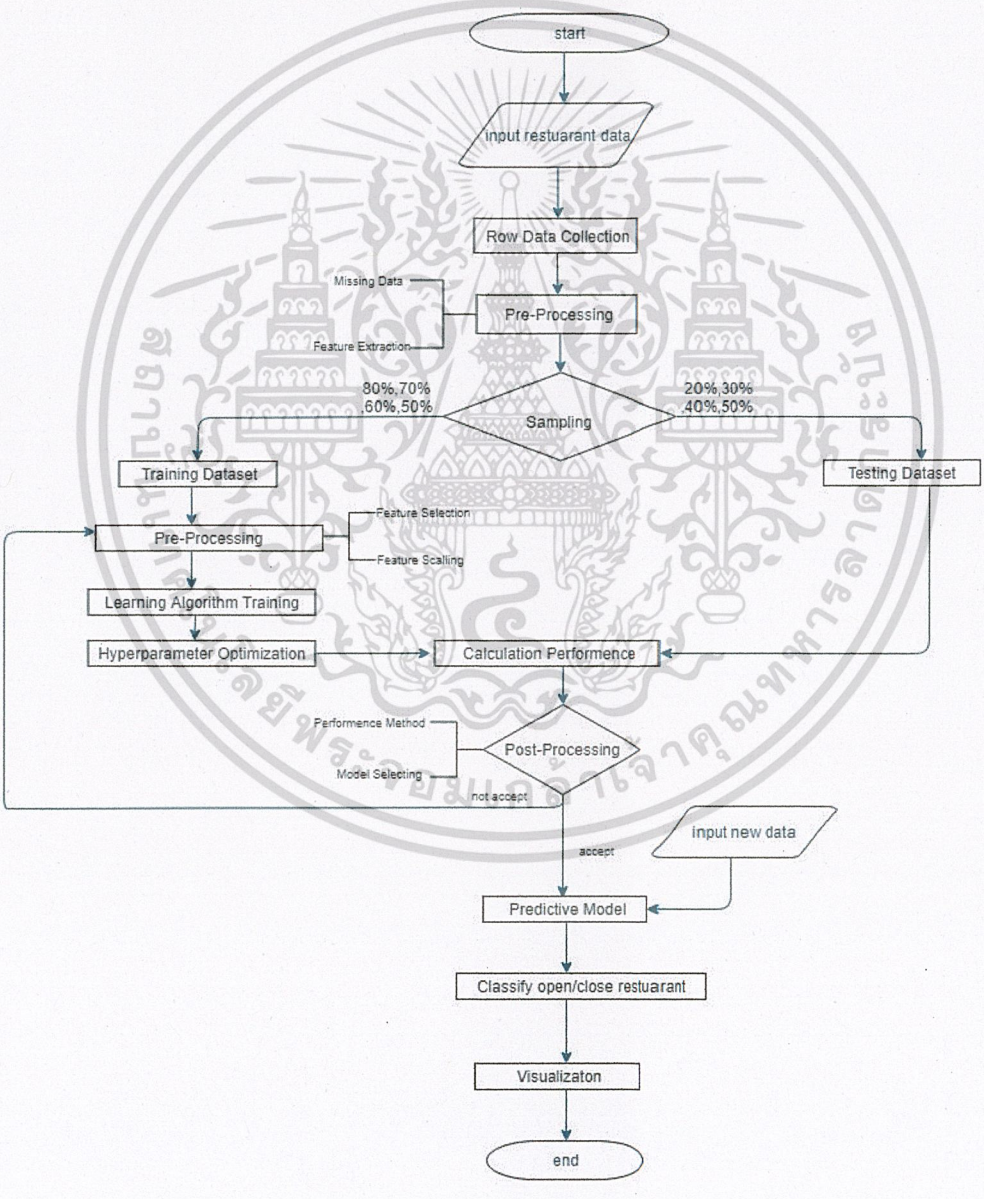
นอกเหนือจากการแบ่งหน้าที่ความรับผิดชอบเป็นรูปแบบแผนกดังข้างต้นแล้ว Wongnai เองยังมีการรวมกลุ่มทีมทำงานในลักษณะ cross functional team ที่เรียกว่า “squad” ซึ่งเป็นทีมย่อยที่ประกอบขึ้นจากสมาชิกจากหลายๆ แผนกมารวมกันเป็นเหมือนบริษัทเล็กๆ ที่มีเป้าหมายในการดูแลส่วนธุรกิจย่อย หรือบริการที่ได้รับมอบหมายให้เติบต่อไป โดย squad มีอำนาจในการตัดสินใจ และจัดการตัวเองได้อย่างคล่องตัว รวมถึงมีอิสระสูงมากในการทำงาน (โดยยึดหลัก speed & impact)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
วิธีการดำเนินงาน

ในบทนี้จะกล่าวถึง วัสดุ อุปกรณ์ เครื่องมือที่จำเป็นต่อการทำงาน และอธิบายถึงขั้นตอนการทดสอบ ที่เป็นลำดับขั้น โดยมีการแสดงผังแสดงกระบวนการทำงานของแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วย วิธีการต้นไม้ตัดสินใจ เพื่อเพิ่มความเข้าใจในแบบจำลองมากยิ่งขึ้น

3.1 แผนผังแสดงภาพรวมกระบวนการดำเนินงาน



ภาพที่ 12 แผนผังแสดงภาพรวมกระบวนการดำเนินงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดการดำเนินงาน

1. เตรียมข้อมูลในการทำแบบจำลอง โดยพิจารณาถึงความเหมาะสม และ มีความสอดคล้อง กับเรื่องของการปิดกิจการของร้านอาหาร

2. ขัดเกลาข้อมูล ทำความสะอาดข้อมูล(Cleaning Data) เพื่อที่จะนำมาใช้ในการสร้างแบบจำลอง

3. ทำการตรวจสอบความถูกต้องข้อมูล ภายหลังการทำความสะอาดข้อมูล(Cleaning Data) ตัวอย่างเช่น ตรวจสอบข้อมูลหาข้อมูลไม่ครบ(Missing Data) อาจจะทำการแทนที่ข้อมูลลงไปเพื่อที่จะให้สามารถนำไปใช้งานต่อได้

4. ทำการแบ่งข้อมูล

   4.1 แบ่งข้อมูลในส่วนของข้อมูลสอน (Training Data) โดยจะแบ่งออกเป็น 80% , 70% , 60% และ 50% ตามลำดับ

   4.2 แบ่งข้อมูลในส่วนของข้อมูลชุดทดสอบ (Testing Data) โดยจะแบ่งออกเป็น 20% , 30% , 40% และ 50% ตามลำดับ

   นั่นคือจะมีการแบ่งข้อมูลของทั้งชุดข้อมูลสอน(Training Data) และ ข้อมูลชุดทดสอบ (Testing Data) คือ ชุดข้อมูลสอน 80% และชุดทดสอบ 20% , ชุดข้อมูลสอน 70% และชุดทดสอบ 30% , ชุดข้อมูลสอน 60% และชุดทดสอบ 40% , ชุดข้อมูลสอน 50% และชุดทดสอบ 50% โดยที่ข้อมูลทั้ง 4 ชุด จะถูกแบ่งแบบสุ่ม (Sampling)

5. เมื่อทำการแบ่งข้อมูลออกเป็นข้อมูลสอน (Training Data) และ ข้อมูลชุดทดสอบ (Testing Data) แล้ว โปรแกรมจะเริ่มทำงานกับชุดข้อมูลสอน ซึ่งจะมีการเลือกรูปแบบของข้อมูลที่เหมาะสมมาใช้งาน เลือกขนาดของต้นไม้ รวมถึงการแบ่งขอบเขตของข้อมูลต้นไม้ตัดสินใจ ตามฟังก์ชั่นของต้นไม้ตัดสินใจ(Decision Tree) ที่เลือกใช้งาน

6. เมื่อโปรแกรมทำการเรียนรู้กับชุดข้อมูลสอน(Training Data) แล้ว จึงทำการทดสอบกับข้อมูลชุดทดสอบ กับข้อมูลชุดทดสอบ(Testing Data) และคำนวณค่าความแม่ยำรวมถึงประสิทธิภาพ (Performance) ของแบบจำลองที่ได้

7. เมื่อทำการทดสอบครบทุกชุดข้อมูลสอน(Training Data)กับข้อมูลชุดทดสอบ(Testing Data) จึงทำการประเมินค่าประสิทธิภาพ (Performance) ของแต่ละแบบจำลอง และทำการเลือกแบบจำลองที่ให้ค่าประสิทธิสูงที่สุดมาใช้งานจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. นำแบบจำลองที่เลือกแล้วนั้นมาใช้งานจริง และ แสดงผล (Visualization) ในรูปแบบที่เข้าใจได้ง่าย

9. เมื่อดำเนินการทุกอย่างเสร็จสิ้นแล้ว ถือเป็นอันจบกระบวนการ

3.2 ตัวแปรที่นำมาศึกษาในการสร้างแบบจำลอง

จากการที่ได้ทำสหกิจมาสักระยะเวลาหนึ่งทำให้พบว่า ปัจจัยที่แสดงถึงการเปิดอยู่ของร้านอาหารมีเพียงไม่กี่ตัวแปรที่จะส่งผลเห็นได้ชัด ได้แก่ การที่ร้านมีรีวิวเข้ามาอยู่ตลอด หรือ มีการรีวิวในช่วงระยะเวลาใกล้ๆกับวันที่ทำการวิเคราะห์แนวโน้มการปิดกิจการของร้านอาหาร อีกปัจจัยหนึ่ง คือ การเช็คอินของผู้ใช้บริการ เพราะเนื่องจากว่า การเช็คอินของร้านที่ผู้ใช้ไปใช้บริการของทางเว็บไซต์วงในมีเดียนั้น จะสามารถเช็คอินได้ก็ต่อเมื่อไปใช้บริการจริงๆ ในระยะบริเวณไม่เกินรัศมีร้าน 50 เมตร

ตัวแปรที่ใช้

- added_date_m วันที่ร้านอาหารเข้ามาสู่ในระบบของเว็บไซต์ ให้หน่วยเป็น เดือน

  โดยในขั้นตอนการทำความสะอาดข้อมูล (Cleaning Data) ได้มีการกำหนดระยะเวลาห่างจากวันที่ 30 กันยายน 2561 เนื่องจากว่าเราทำการวิเคราะห์ในช่วงไตรมาสที่ 4 ของบริษัท จึงเลือกเกณฑ์ว่าข้อมูลควรอยู่ในระยะช่วงของก่อนไตรมาสที่ 4

- lastes_rw_m วันที่มีการรีวิวล่าสุดของร้านอาหาร ให้หน่วยเป็น เดือน

  โดยในขั้นตอนการทำความสะอาดข้อมูล (Cleaning Data) ได้มีการกำหนดระยะเวลาห่างจากวันที่ 30 กันยายน 2561 เนื่องจากว่าเราทำการวิเคราะห์ในช่วงไตรมาสที่ 4 ของบริษัท จึงเลือกเกณฑ์ว่าข้อมูลควรอยู่ในระยะช่วงของก่อนไตรมาสที่ 4

- num_rw จำนวนรีวิวทั้งหมดของร้านอาหาร

- number_chk จำนวนเช็คอินทั้งหมดของร้านอาหาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 วิธีการดำเนินงาน

ตารางที่ 4 ข้อมูลที่เรียกมาใช้งานจากระบบฐานข้อมูล

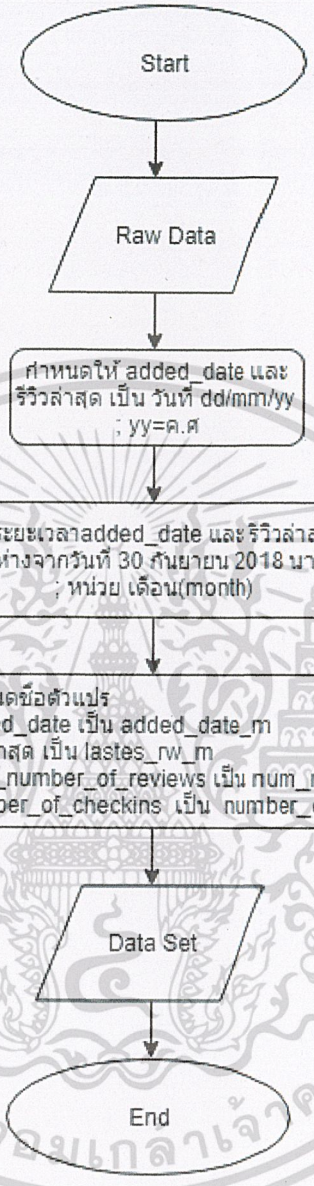
	A	B	C	D	E	F	G	H
1	added_date	id	name	total_number_of_review	number_of_checkins	number_of_been_her	รีวิวล่าสุด	result
2	2015-11-12 17:04:24	215202	Pepper Lunch	36	20	5	1/6/61	open
3	2010-10-05 18:39:03	2730	PIZZERIA LIMONCELLO	84	52	25	13/8/61	open
4	2010-08-01 0:00:00	4887	The Roof Gastro	38	25	20	1/4/61	open
5	2011-08-29 0:16:04	10226	Pepper Lunch	113	52	11	26/8/61	open
6	2010-08-04 0:00:00	8040	เฮ้ยชุนเส็ง	117	156	35	1/9/61	open
7	2014-09-15 14:24:30	171838	Shinsei	52	44	14	2/9/61	open
8	2012-10-25 20:44:43	27238	Pepper Lunch	47	31	8	27/4/61	open
9	2010-11-17 15:33:09	3814	Honmono Sushi	64	29	11	2/9/61	open
10	2010-08-01 0:00:00	4309	Maru Kaiseki	26	11	5	27/7/61	open
11	2012-05-28 19:34:23	18959	OGU OGU	25	18	4	20/4/61	open
12	2012-04-11 17:11:45	15707	iberry	27	15	1	25/7/60	open
13	2012-04-29 22:27:25	17023	Pepper Lunch	31	17	3	23/4/60	open
14	2014-11-24 22:26:37	180326	กาแฟลอยลังกา	31	40	4	6/7/61	open
15	2015-12-27 14:58:13	219279	Ansel & Elliott • Analog Cafe	20	41	7	3/9/61	open
16	2013-02-25 23:23:26	123433	ตำกะเดียว	64	47	4	16/8/60	open
17	2012-07-31 18:57:44	23151	Pepper Lunch	32	20	3	18/8/61	open
18	2015-04-20 12:31:21	194085	Seafood Express	39	4	14	2/9/61	open
19	2016-10-27 18:10:04	252385	เอี๋ยวไก่ สุกี้โบราณ	14	19	5	21/7/61	open
20	2016-03-10 10:05:09	227256	Cold Stone Creamery	19	17	4	24/5/61	open
21	2014-01-19 17:11:28	146807	S&P BAKERY SHOP	19	11	1	19/12/60	open
22	2013-05-27 18:27:32	128260	หนมไท	3	1	0	4/4/60	close
23	2012-11-23 17:43:03	67188	ลาบแจ่วซ์	1	0	0	12/5/56	close
24	2013-11-27 22:24:39	142006	หมูหม่อ	2	1	1	25/5/58	close
25	2012-11-23 17:47:58	72930	แสนคำเล็กแฮนด์โรล	1	0	0	18/10/57	close
26	2012-05-03 15:37:52	17292	Iberry - Paradise Park	25	12	1	24/6/59	close
27	2012-12-29 13:38:45	119794	ก๋วยเตี๋ยวเป็ดวังหลัง	2	4	1	8/4/59	close
28	2010-12-02 11:07:44	7674	Shodai Keisuke	9	0	1	30/11/56	close
29	2014-07-01 17:06:47	159212	Hotdogs & Buns	5	2	0	5/3/58	close
30	2014-05-06 0:49:56	154719	Dream Cones	3	1	0	31/12/57	close
31	2013-06-25 11:16:30	130072	ร้านแม่ศุ๊ค อาหารตามสั่ง	2	3	0	2/7/57	close
32	2015-05-09 10:02:10	195932	Grill Yard	9	14	0	19/12/59	close
33	2014-01-16 12:35:05	146531	Coco Ichibanya	5	2	0	21/8/59	close

3.3.1 ตัวอย่างตารางข้อมูลที่เรียกมาใช้งาน (Raw Data)

จากภาพที่ 13 เป็นภาพที่แสดงตัวอย่างของตารางข้อมูลที่เรียกนำมาใช้งาน จากภาพจะเห็นได้ว่าชื่อคอลัมน์ ที่เรียกมาจะเป็นชื่อคอลัมน์ตามที่มีอยู่ในระบบฐานข้อมูล ซึ่งข้อมูลที่เรียกนำออกมาใช้งานจะยังไม่มีการคำนวณจัดการตามค่าสถิติหรือคำนวณค่าทางคณิตศาสตร์ โดยจำนวนข้อมูลที่นำมาใช้งานจริงมีทั้งสิ้น 27005 ข้อมูล เป็นข้อมูลของร้านที่เปิดการอยู่ 17491 ข้อมูล และ ร้านที่ปิดกิจการไปแล้ว 9514 ข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 ขั้นตอนการทำความสะอาดข้อมูล (Cleaning Data)



ภาพที่ 13 ขั้นตอนการทำความสะอาดข้อมูล (Cleaning data)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.3 ตัวอย่างตารางข้อมูลที่นำมาใช้งาน (Data Set)

ตารางที่ 5 ข้อมูลที่จะนำไปใช้งาน

	A	B	C	D	E	F
1	id	added_date_m	num_rw	number_chk	lastes_rw_m	result
2	141634	58	26	6	3	close
3	217604	33	2	2	28	close
4	249969	23	5	0	9	close
5	199543	39	2	1	39	close
6	219126	33	1	0	33	close
7	337807	3	4	1	0	close
8	184802	44	3	0	39	close
9	328555	5	1	0	5	close
10	152164	53	3	11	9	close
11	151419	54	1	2	40	close
12	245319	25	6	0	14	open
13	140319	58	11	8	2	open
14	23535	73	8	5	13	open
15	2183	95	12	6	4	open
16	213931	35	12	9	3	open
17	59447	70	5	3	8	open
18	161437	50	12	8	1	open
19	12313	81	17	16	2	open
20	10087	85	37	19	2	open
21	162258	50	6	7	8	open

ตารางที่ 5 เป็นภาพที่แสดงตัวอย่างของตารางข้อมูลที่นำมาใช้งาน โดยผ่านการทำความสะอาดข้อมูล (Cleaning Data) ข้อมูลชุดนี้จะนำไปใช้วิเคราะห์ในการสร้างแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วย วิธีต้นไม้ตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Decision Tree : C5.0 and CART โดยใช้ R Studio

```
#preparing data
library(caTools) ; library(C50) ; library(dplyr) ; library(readxl) ; .....(1)
#import data
#filename <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")
filename <- read_excel("pathfile/filename.xlsx",   sheet = "sheetname")..... (2)
DF <- as.data.frame(filename) .....(3)
######################
set.seed(1)
#80%_20%
train80 <- sample_frac(DF , size = 0.8)    #run for keep train80 first then '#' > train80
as same as all the time .....(4)
test20 <- anti_join(DF , train80 , by = "id")
##############################
#C50
names(DF)
tc5_8020 <- C5.0(x = train80[,c(2:5)], y = as.factor(train80$result),control = .....(5)
C5.0Control(noGlobalPruning = TRUE,minCases = seq(1,80,1)))
tc5_8020
summary(tc5_8020)
plot(tc5_8020) .....(6)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#performancec58020
predict.unseenc58020  <- predict(tc5_8020 , test20[,-6] , type = 'class') .....(7)
confusion.matc58020 <- table(test20$result,predict.unseenc58020)
confusion.matc58020 .....(8)
accuract.c5test20 <- sum(diag(confusion.matc58020))/sum(confusion.matc58020)
accuract.c5test20 .....(9)
################################
#rpart
#library(rpart) .....(10)
#library(rpart.plot)
#names(DF)
trpart_8020 <- rpart(formula = result ~ added_date_m+ lastes_rw_m + num_rw +
number_chk , data = train80 , .....(11)
minsplit = seq(10,50,1) , cp = seq(0.001,1,0.001)  , maxsurrogate = seq(1,7,1) , xval =
seq(5,15,1) , maxdepth = seq(3,20,1))
rpart.plot(trpart_8020) .....(12)
################################
#performancerpart8020
predict.unseenrpart8020 <- predict(trpart_8020 , test20[,-6] , type = 'class')
predict.unseenrpart8020 .....(13)
confusion.matrpart8020 <- table(test20$result,predict.unseenrpart8020)
confusion.matrpart8020 .....(14)
accuract.rparttest20 <-
sum(diag(confusion.matrpart8020))/sum(confusion.matrpart8020)
accuract.rparttest20 .....(15)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
######tester#####
library(readxl) .....(16)
test20 <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")
predict.unseenc58020 <- predict(tc5_8020 , test20[,-6] , type = 'class') .....(17)
predict.unseenc58020
#when type is class
test20$class <- predict.unseenc58020
#when type is prob
#test20$close <-predict.unseenc58020[,1]
#test20$open <- predict.unseenc58020[,2]
View(test20) .....(18)
write.csv(test20 , "tester20.csv")
```

อธิบายความหมายของแต่ละขั้นตอนตามหมายเลข ดังต่อไปนี้

1) library(caTools) ; library(C50) ; library(dplyr) ; library(readxl) ;

   การเรียกใช้ library ที่เกี่ยวข้องในการทำต้นไม้ตัดสินใจ

2) filename <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")

   การเรียกใช้ ข้อมูลที่จะนำมาสร้างต้นไม้ตัดสินใจ

3) DF <- as.data.frame(filename)

   การทำให้ข้อมูลอยู่ในรูปแบบของ data frame ใน R

4) train80 <- sample_frac(DF , size = 0.8) และ test20 <- anti_join(DF , train80 , by = "id")

   การแบ่งข้อมูลสอน (Training Data) โดยจะแบ่งเป็น 80 , 70 , 60 และ 50 ตามลำดับ

   การแบ่งข้อมูลชุดทดสอบ (Testing Data) โดยจะแบ่งเป็น 20 , 30 , 40 และ 50 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) tc5_8020 <- C5.0(...)

การเรียกใช้ Decision Tree C5.0 Algorithm กับชุดข้อมูลสอน(Training Data)

6) plot(tc5_8020)

แสดง C5.0 Decision Tree ที่ได้ออกมา

7) predict.unseenc58020 <- predict(tc5_8020 , test20[,-6] , type = 'class')

ทำนายผลข้อมูลชุดทดสอบ (Testing Data) จาก ข้อมูลชุดสอน (Training Data)

8) confusion.matc58020

แสดงผลของ Confusion matrix ว่า ทำนายถูกต้อง หรือ ผิด เป็นจำนวนเท่าไร

9) accuract.c5test20

การคำนวณค่าความถูกต้องของการทำนายจากผลของชุดข้อมูลสอนและข้อมูลชุดทดสอบ

10) #library(rpart) และ #library(rpart.plot)

การเรียกใช้ libraryที่เกี่ยวข้องกับ CART Algorithm

11) trpart_8020 <- rpart(...)

การใช้ Decision tree rpat จะมีคุณสมบัติต่างๆ สามารถสืบค้นได้จาก help ของโปรแกรม

12) rpart.plot(trpart_8020)

แสดง rpart Decision Tree ที่ได้ออกมา

13) predict.unseenrpart8020 <- predict(trpart_8020 , test20[,-6] , type = 'class')

ทำนายผลโดยใช้ข้อมูลชุดทดสอบ (Testing Data) จาก ข้อมูลชุดสอน (Training Data)

14) confusion.matrpart8020

แสดงผลของ Confusion matrix ว่า ทำนายถูกต้อง หรือ ผิด เป็นจำนวนเท่าไร

15) accuract.rparttest20

การคำนวณค่าความถูกต้องของการทำนายจากผลของชุดข้อมูลสอนและข้อมูลชุดทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

16) library(readxl) และ test20 <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")

ในขั้นตอนนี้จะเป็นการที่นำข้อมูลจริงๆมาทดสอบแบบจำลองของเราที่ได้มาว่า การนำมาใช้จริงนั้น ผลที่ได้ออกมาจะดีขึ้นเท่าไร ก็จะมีการอ่านข้อมูลที่เราต้องการจะให้มัน classification ออกมา

17) predict.unseenc58020

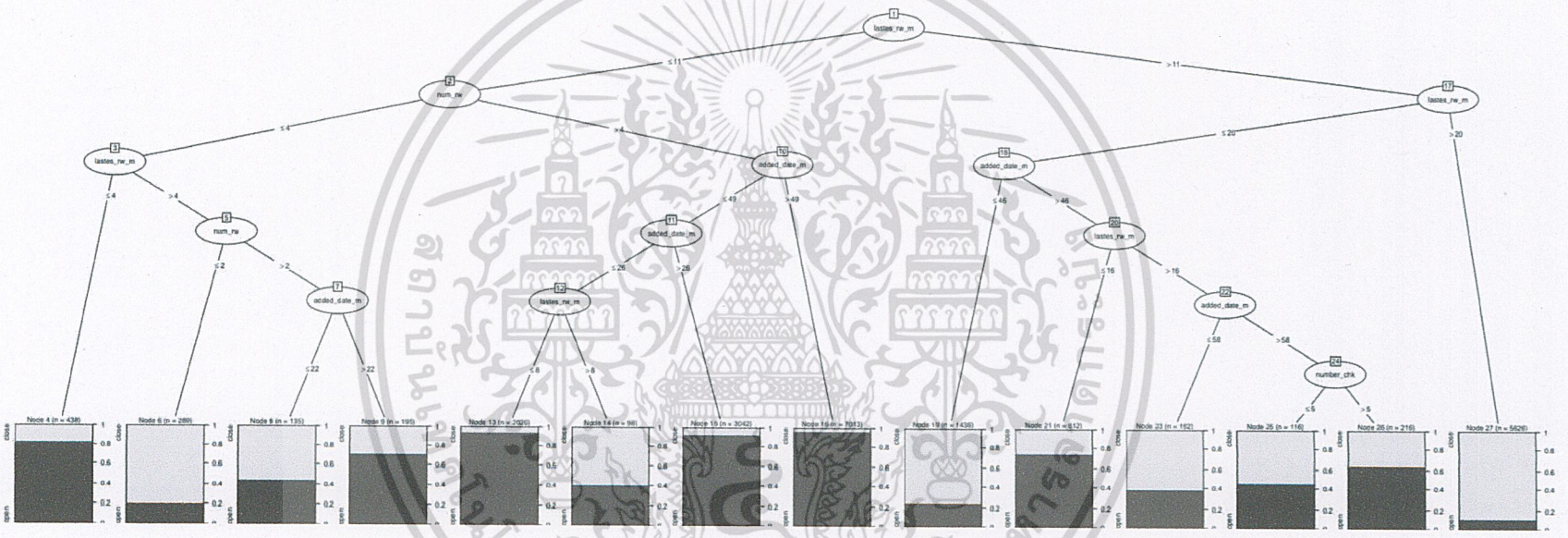
classification จากข้อมูลจริงว่ามีร้านที่ปิดกิจการ และ ร้านที่ยังเปิดบริการ เป็นเท่าไร

18) View(test20)

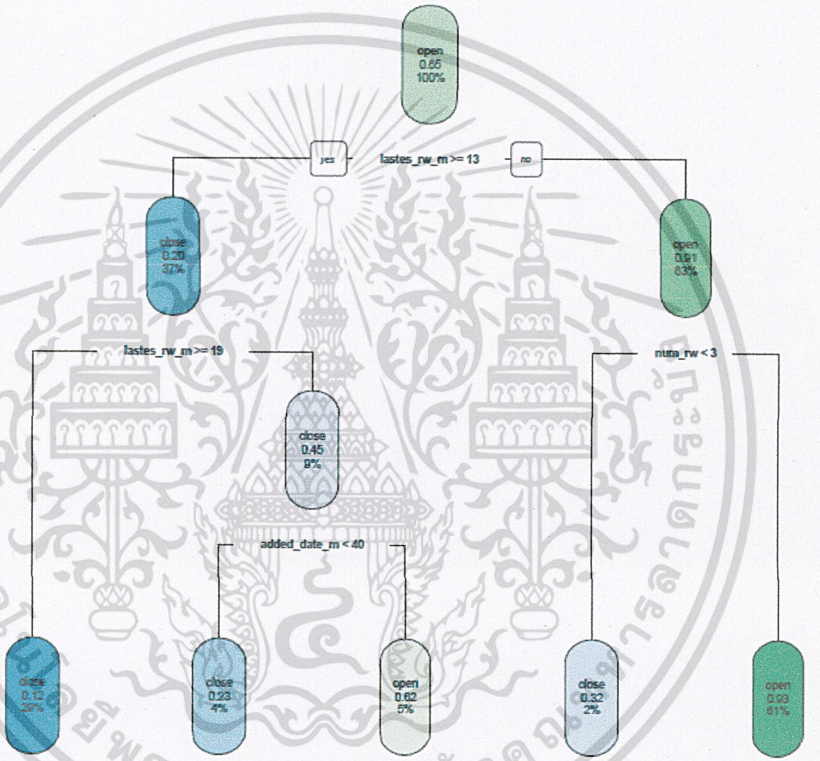
แสดงผลว่าแบบจำลอง classification ออกมาได้ผลลัพ์เป็นอย่างไร และสามารถนำไปตรวจสอบได้อีกครั้งหนึ่งเพื่อความถูกต้องอย่าง100%

ต่อไปจะแสดงถึง Decision Tree ทั้ง 8 แบบที่ได้มาดังนี้

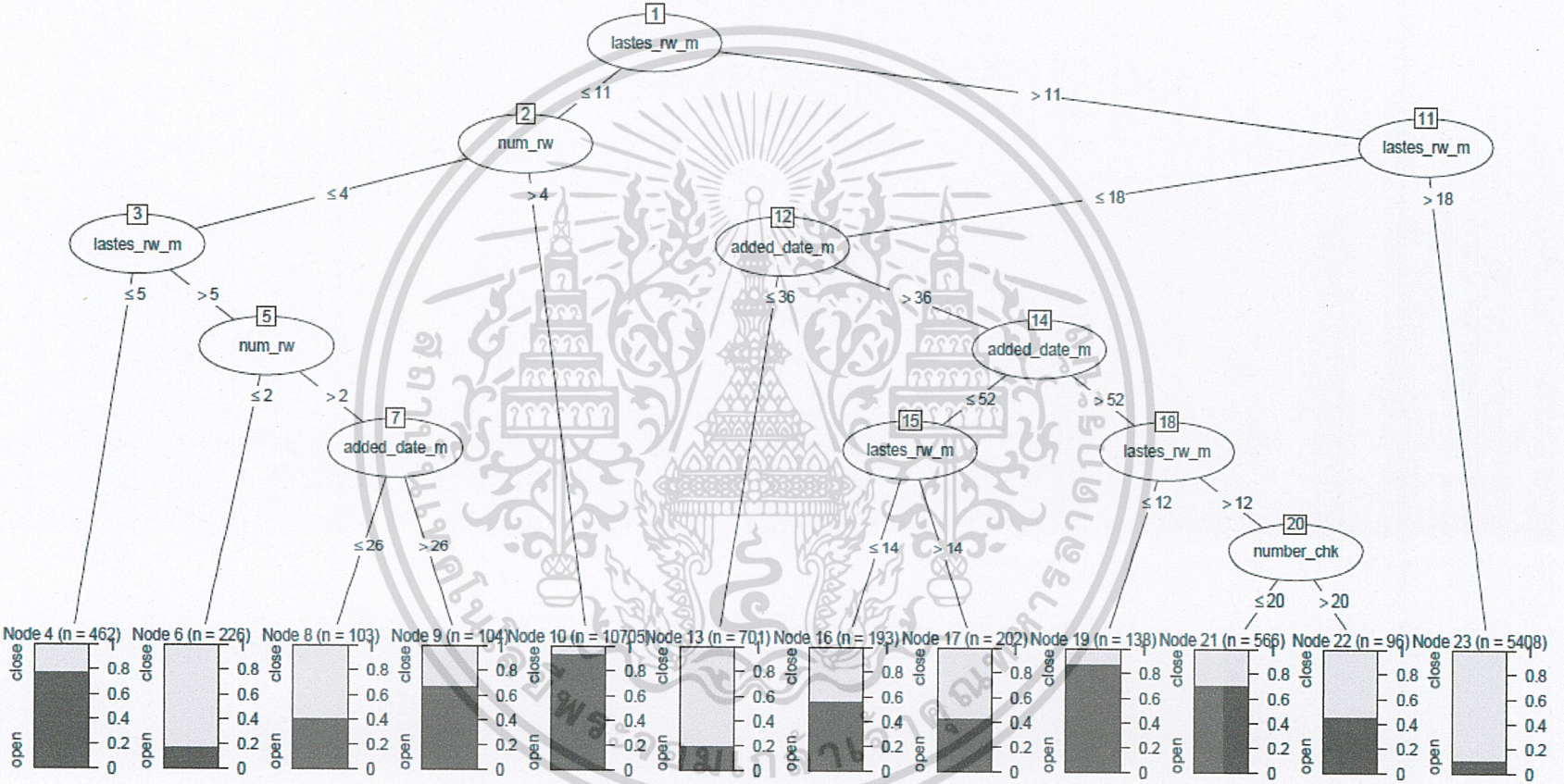
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



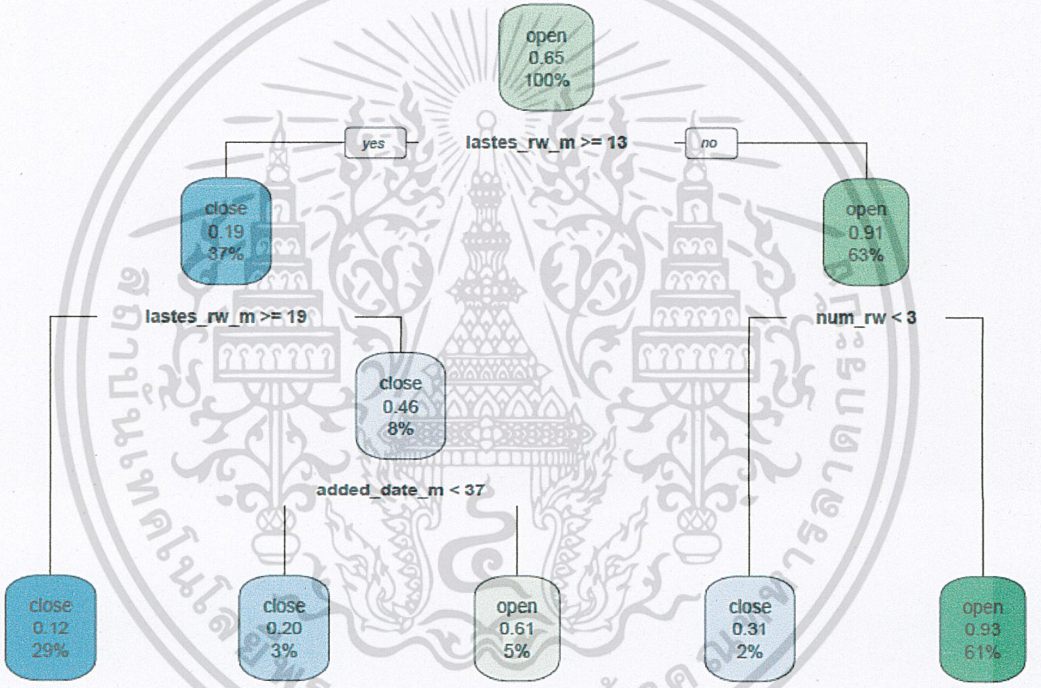
ภาพที่ 14 Decision Tree C5.0 Algorithm (training 80 % and testing 20 %)



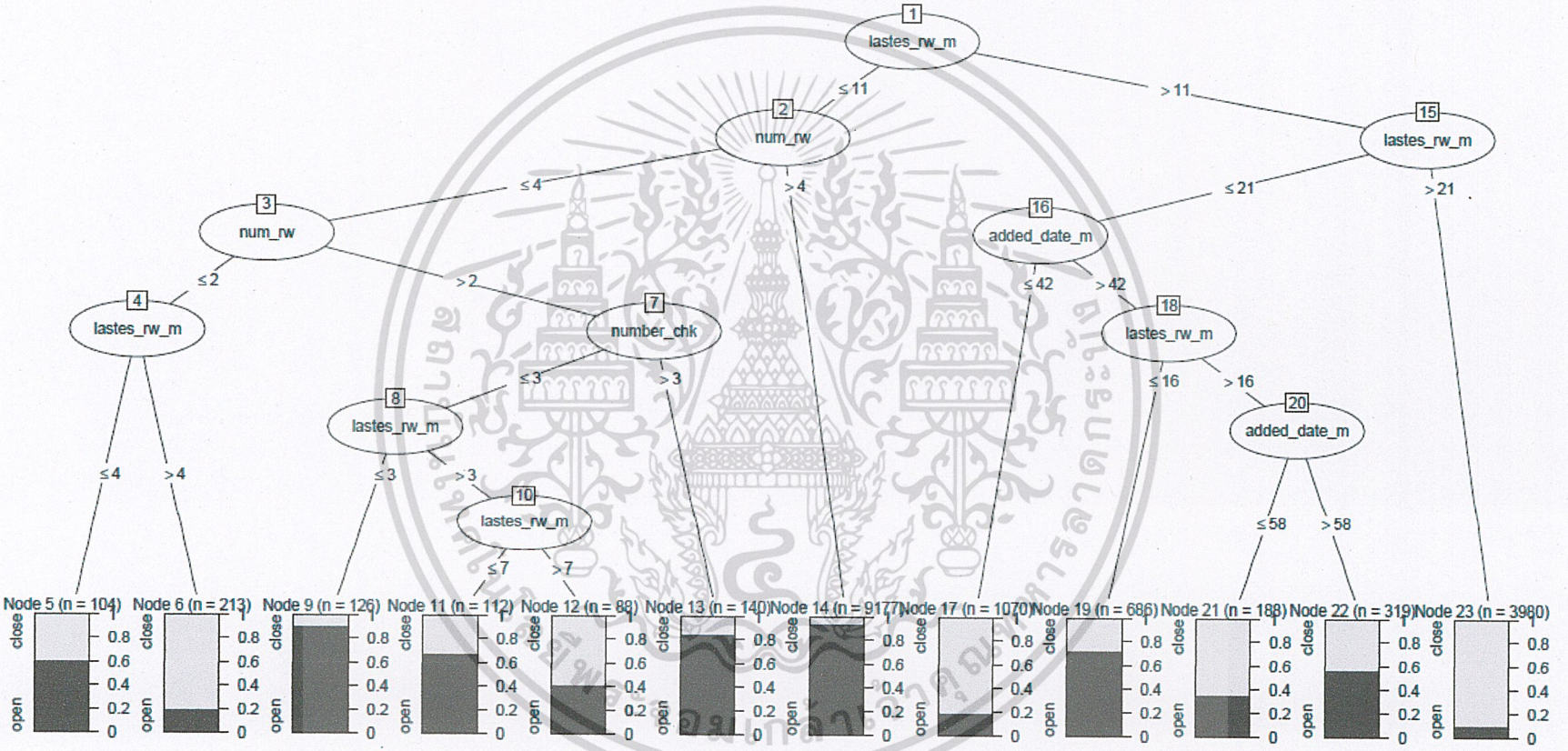
ภาพที่ 15 Decision Tree CART Algorithm (training 80 % and testing 20 %)



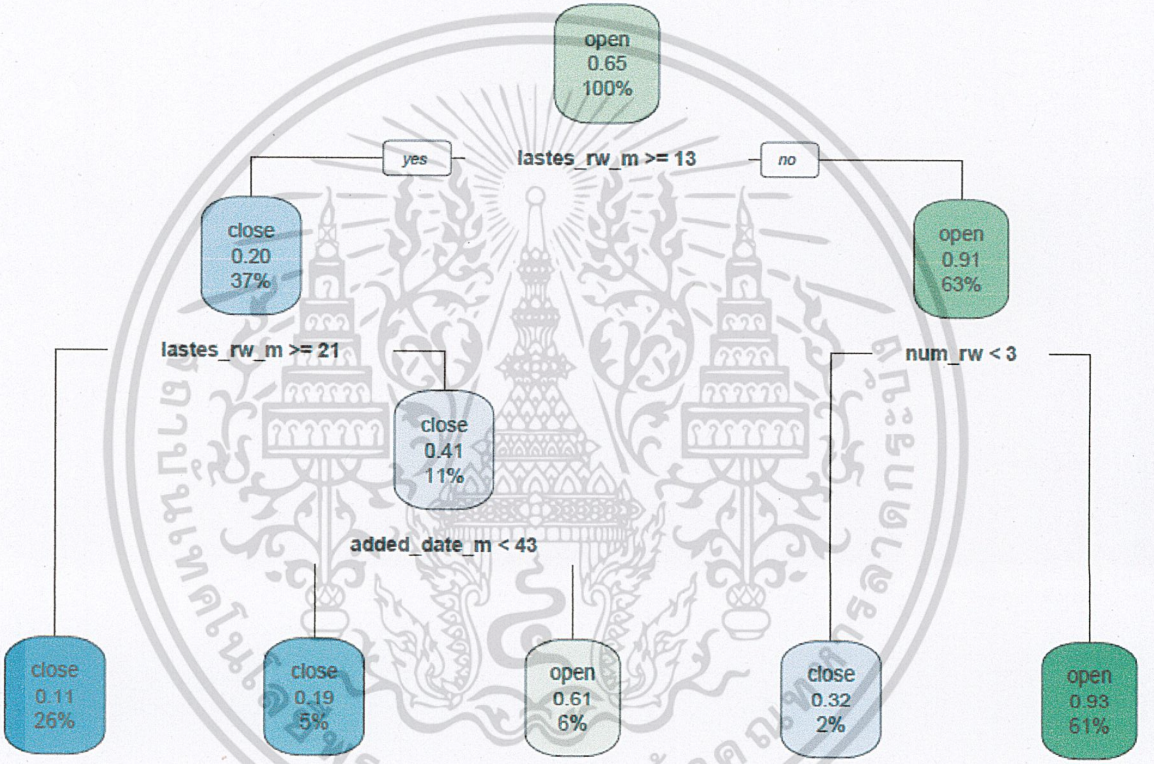
ภาพที่ 16 Decision Tree C5.0 Algorithm (training 70 % and testing 30 %)



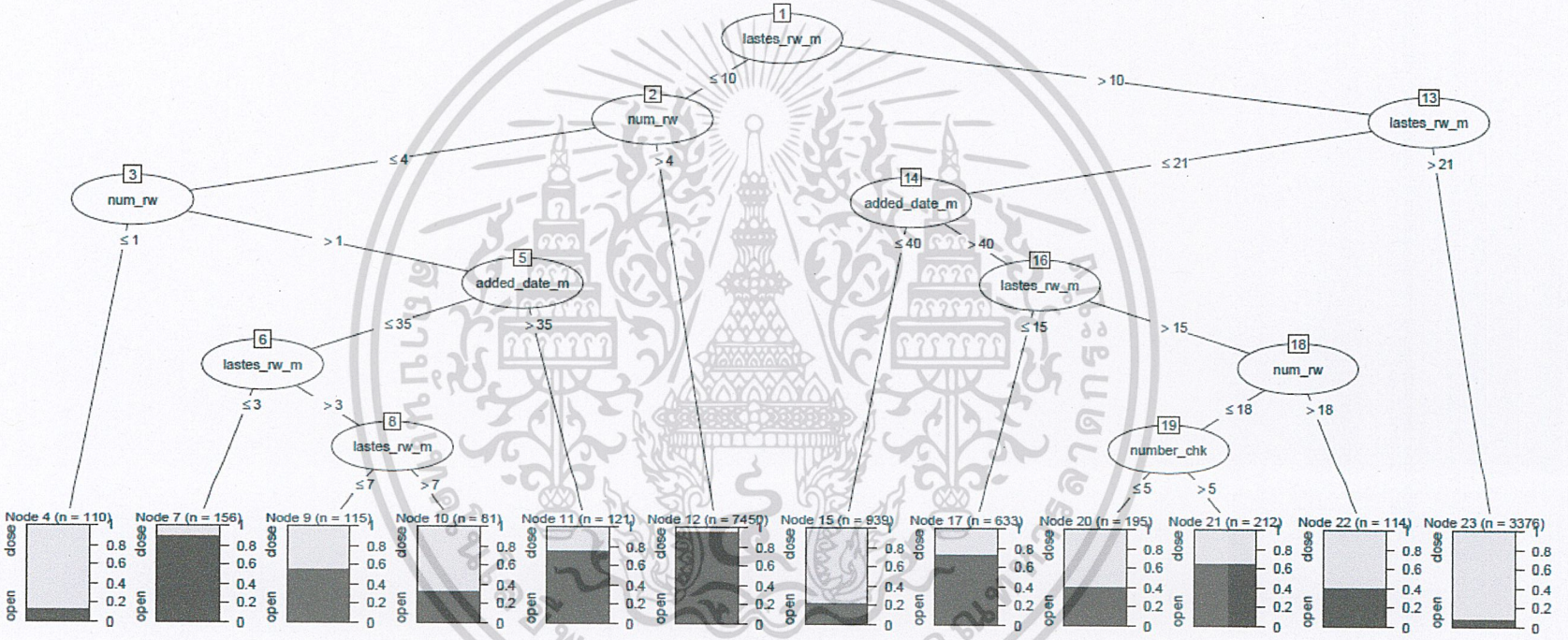
ภาพที่ 17 Decision Tree CART Algorithm (training 70 % and testing 30 %)



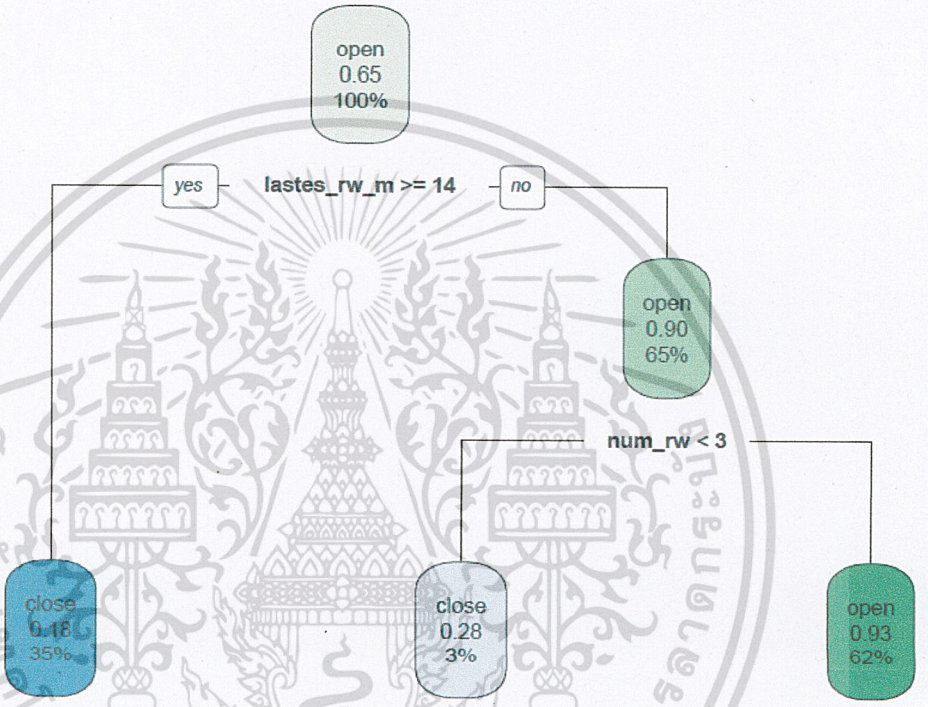
ภาพที่ 18 Decision Tree C5.0 Algorithm (training 60 % and testing 40 %)



ภาพที่ 19 Decision Tree CART Algorithm (training 60 % and testing 40 %)



ภาพที่ 20 Decision Tree C5.0 Algorithm (training 50 % and testing 50 %)



ภาพที่ 21 Decision Tree CART Algorithm (training 50 % and testing 50 %)

3.4 ตรวจสอบค่าPerformanceของแบบจำลองครั้งที่1

จากหัวข้อที่ผ่านมาที่ได้แสดงผลวิธีการดำเนินการไปแล้ว ในหัวข้อนี้จะแสดงผลค่าความแม่นยำ i(Performance) ต่างๆ เพื่อที่จะทราบว่าเราควรจะเลือกแบบจำลองใดมาใช้งาน

ตารางที่ 6 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 80% และ testing data 20 %

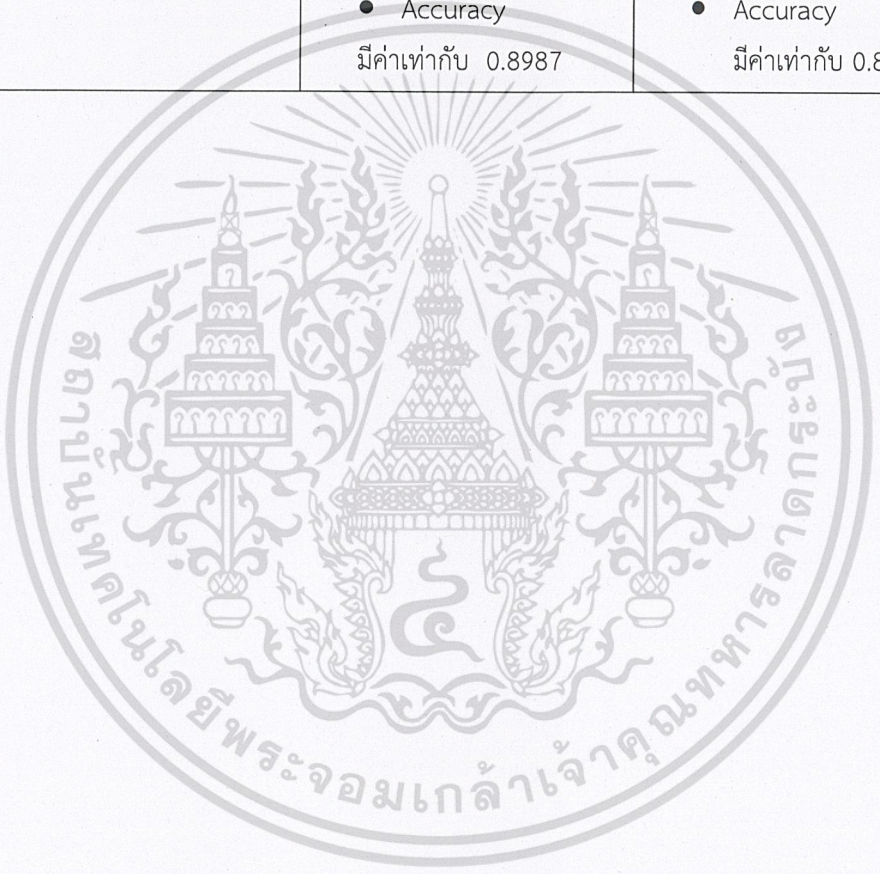
DECISION TREE TYPE / SPLITING	C5.0	CART
TRAINING DATA 80% TESTING DATA 20%	• Confusion matrix close open close 1631 238 open 309 3223 TRUE POSITIVE (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 1631 ร้าน FALSE NEGATIVE (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 238 ร้าน FALSE POSITIVE (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 309 ร้าน TRUE NEGATIVE (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 3223 ร้าน	• Confusion matrix close open close 1577 292 open 265 3267 TRUE POSITIVE (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 1577 ร้าน FALSE NEGATIVE (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 292 ร้าน FALSE POSITIVE (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 265 ร้าน TRUE NEGATIVE (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 3267 ร้าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	• PRECISION โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{1631}{1631+309} = 0.8407$ • RECALL -TRUE POSITIVE RATE $\frac{1631}{1631+238} = 0.8727$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8727เท่าของร้านปิดทั้งหมด -TRUE NEGATIVE RATE $\frac{3223}{3223+309} = 0.9125$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.9125 เท่าของร้านเปิดทั้งหมด -FALSE POSITIVE RATE $\frac{3309}{3309+309} = 0.0875$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.0875เท่าของร้านเปิดทั้งหมด	• PRECISION โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{1577}{1577+265} = 0.8561$ • RECALL -TRUE POSITIVE RATE $\frac{1577}{1577+292} = 0.8438$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8438 เท่าของร้านปิดทั้งหมด -TRUE NEGATIVE RATE $\frac{3267}{3267+265} = 0.925$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.925 เท่าของร้านเปิดทั้งหมด -FALSE POSITIVE RATE $\frac{265}{3267+265} = 0.075$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.075เท่าของร้านเปิดทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	-FALSE NEGATIVE RATE $\frac{238}{1631+238} = 0.1273$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1273 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8987	-FALSE NEGATIVE RATE $\frac{292}{1577+292} = 0.1562$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1562 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8969



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 7 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 70% และ testing data 30 %

DECISION TREE TYPE / SPLITING	C5.0	CART
training data 70% **testing data 30%**	• Confusion matrix Close open close 2389 415 open 477 4820 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 2389 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 415 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 477 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 4820 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{2389}{2389+477} = 0.8336$	• Confusion matrix close open close 2331 473 open 445 4852 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 2331 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 473 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 445 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 4852 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{2331}{2331+445} = 0.8397$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	• Recall -True Positive Rate $\frac{2389}{2389+415} = 0.8520$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8520 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{4820}{4820+477} = 0.9099$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน0.9099 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{477}{4820+477} = 0.0901$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.0901 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{415}{2389+415} = 0.1480$ ค่าที่บอกว่าโปรแกรมทำนาย ว่าร้านเปิด เป็นอันตราส่วน 0.1480 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8899	• Recall -True Positive Rate $\frac{2331}{2331+473} = 0.813$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.813 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{4852}{4852+445} = 0.9160$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.9160 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{445}{4852+445} = 0.0840$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.0840 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{473}{2331+473} = 0.1687$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1687 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8867

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 8 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 60% และ testing data 40 %

DECISION TREE TYPE / SPLITING	C5.0	CART
training data 60% **testing data 40%**	• Confusion matrix Close open close 3229 614 open 535 6424 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 3229 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 614 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 535 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 6424 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{3229}{3229+535} = 0.8579$	• Confusion matrix close open close 3145 698 open 497 6462 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 3145 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 698 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 497 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 6462 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{3145}{3145+497} = 0.8635$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	• Recall -True Positive Rate $\frac{3229}{3229+614} = 0.8402$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8402 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{6424}{6424+535} = 0.9231$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.9231 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{535}{6424+535} = 0.0769$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตรา-ส่วน 0.0769 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{614}{3229+614} = 0.1596$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1596 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8936	• Recall -True Positive Rate $\frac{3145}{3145+698} = 0.8184$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8184 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{6462}{6462+497} = 0.9286$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.9286 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{497}{6462+497} = 0.0714$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตรา-ส่วน 0.0714 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{698}{3145+698} = 0.1816$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1816 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8894

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 9 แสดงค่าความแม่นยำจากการแบ่งโดยtraining data 50% และ testing data 50 %

DECISION TREE TYPE / SPLITING	C5.0	CART
training data 50% **testing data 50%**	• CONFUSION MATRIX close open close 4024 713 open 753 8013 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 4024 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 713 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 753 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 8013 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{4024}{4024+753} = 0.8424$	• CONFUSION MATRIX close open close 4120 617 open 991 7775 True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 4120 ร้าน False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 617 ร้าน False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 991 ร้าน True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 7775 ร้าน • Precision โปรแกรมทำนายว่าร้านปิดถูกต้อง $\frac{4120}{4120+991} = 0.8061$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	• Recall -True Positive Rate $\frac{4024}{4024+713} = 0.8495$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8495 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{8013}{8013+753} = 0.9141$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.9141 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{753}{8013+753} = 0.0859$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.0859 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{713}{4024+713} = 0.1505$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1505 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8914	• Recall -True Positive Rate $\frac{4120}{4120+617} = 0.8697$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.8697 เท่าของร้านปิดทั้งหมด -True Negative Rate $\frac{7775}{7775+991} = 0.8869$ ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.8869 เท่าของร้านเปิดทั้งหมด -False Positive Rate $\frac{991}{7775+991} = 0.1131$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.1131 เท่าของร้านเปิดทั้งหมด -False Negative Rate $\frac{617}{4120+617} = 0.1303$ ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.1303 เท่าของร้านปิดทั้งหมด • Accuracy มีค่าเท่ากับ 0.8809

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

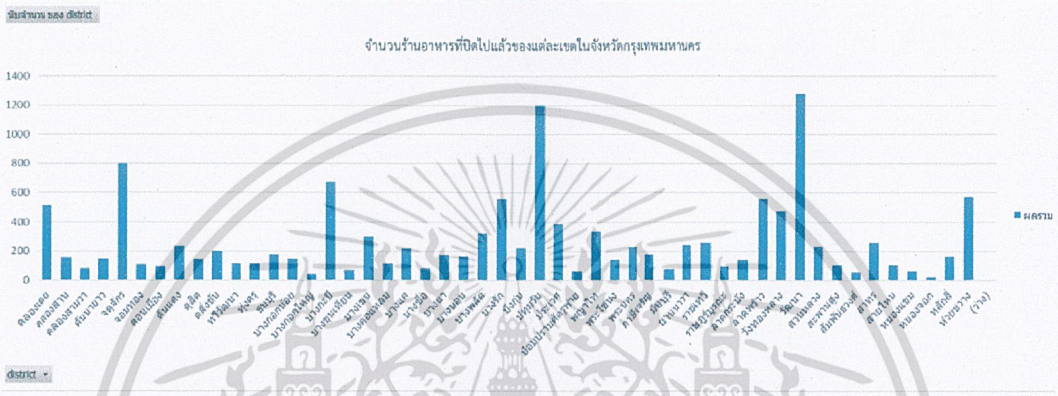
จากตารางที่แสดงค่าความแม่นยำ (Performance) ของผลลัพธ์ที่ดำเนินการทั้งหมดของแบบจำลองครั้งที่1 พบว่า แบบจำลองที่เลือกนำมาใช้คือ แบบจำลอง "***C5.0 Decision tree ที่มีการแบ่ง training data 80% และ testing data 20%***" ซึ่งมีค่า Accuracy สูงที่สุดคือ 0.8987 นั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 การพัฒนาแบบจำลอง

จากการที่ได้ทำแบบจำลองจากครั้งที่1 ในข้างต้นแล้วนั่น ได้มีการพัฒนาเพื่อเพิ่มประสิทธิภาพความแม่นยำของแบบจำลอง โดยมีวิธีการพัฒนาแบบจำลองดังต่อไปนี้

1. นำข้อมูลที่เป็นร้านอาหารในจังหวัดกรุงเทพมหานครที่ทำการปิดกิจการในระบบของเว็บไซต์มาทำการ Pivot เพื่อแสดงผลความถี่ของจำนวนร้านอาหารในเขตต่างๆ



ภาพที่ 22 กราฟแสดงความถี่ของจำนวนร้านอาหารที่ปิดไปแล้วของแต่

จากกราฟที่แสดงความถี่ของจำนวนร้านอาหารที่ปิดไปแล้วของแต่ละเขตในจังหวัดกรุงเทพมหานคร พบว่าร้านที่ปิดไปในระบบแล้ว3อันดับแรก คือ เขตวัฒนา , เขตปทุมวัน และ เขตจตุจักร ตามลำดับ ในการพัฒนาแบบจำลองจึงเลือกนำร้านที่อยู่ในเขตวัฒนามาพิจารณาในการนำมาสร้างแบบจำลองใหม่ โดยเขตนี้เป็น ร้านอาหารเขตในเมือง ซึ่งจากการที่ได้มีการอภิปรายกับพี่ในทีมทำให้ทราบว่าเหตุผลที่ร้านในสองเขตนี้มีการปิดมากนั้นเพราะว่ามีความหลากหลายและการแข่งขันของร้านอาหารอย่างมาก จึงทำให้เกิดการเปลี่ยนแปลงหมุนเวียนของร้านอาหารจำนวนมากนั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ทำการจัดเตรียมข้อมูลอันใหม่เพื่อนำมาพัฒนาแบบจำลอง

id	name	num_chk	num_rw	added_date_m	lastes_rw_m	result
249969	Yellow Spo	0	5	22	8	close
316903	Crepe Box	0	4	5	61	close
345239	ขนมจีนน้าย	0	2	0	0	close
2922	Hip Hut Ca	0	1	94	40	close
7665	iberry	27	40	93	8	close
303004	Kuh Grill &	4	12	8	2	close
23879	ร้านป้าเล็ก	0	1	72	72	close
333091	Motono	0	2	3	2	close
59309	บอนกุสโต้	0	1	69	63	close
60438	เชเรกามิเย่ร์	0	1	69	64	close
3986	Fuji Coffee	0	1	96	75	close
253814	Mine Japar	1	1	21	18	close
307898	Ya Kun Ka	1	2	6	6	close
3887	Coffee Sho	3	2	96	65	close
302741	Romane S	0	1	8	8	close

ภาพที่ 23 ตัวอย่างของร้านอาหารที่ปิดในระบบในเขต

id	name	num_chk	num_rw	added_date_m	lastes_rw_m	result
204034	น้ำเต้าหู้เตาถ	6	8	37	7	open
58385	Bamboo แ	8	8	69	0	open
190	Sophia Re	5	8	94	2	open
295224	Jamie's Bu	4	8	10	2	open
2497	ประจักษ์เบเ	3	8	94	7	open
216431	Daily Roas	8	8	33	3	open
3081	Cabbages	2	8	94	32	open
302187	Shabuku	4	8	8	1	open
325151	บ้านน้ำพริก	6	8	5	2	open
4161	Katsuichi	5	8	96	0	open
262577	Kalpapruek	4	8	19	3	open
259842	Katana Sh	9	8	20	2	open
321932	HAKATA (	7	8	5	2	open
341506	Drip Tea &	4	8	1	0	open
326184	Lerdtip Tho	5	8	4	0	open
217252	Charcoal T	2	8	32	11	open
1893	ชะ ชะ อัน ถ	6	8	94	5	open
294834	Casa Lapir	13	8	10	2	open
306564	เสือใต้	6	8	7	0	open

ภาพที่ 24 ตัวอย่างของร้านอาหารที่ยังเปิดอยู่ในระบบในเขต

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การนำข้อมูลไปทำต้นไม้ตัดสินใจ โดยการใช้ C5.0 Decision Tree Algorithm

```
#preparing data dev
library(caTools) ; library(C50) ; library(dplyr) ; library(readxl) ; library(e1071) .....(1)
#import data
#filename <- read_excel("pathfile/filename.xlsx", sheet = "sheetname").....(2)
filename <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")
DDF <- as.data.frame(filename) .....(3)
######################
set.seed(1)
#80%_20%
#Dtrain80 <- sample_frac(DDF , size = 0.8)    #run for keep train80 first then '#' >
train80 as same as all the time
#Dtest20 <- anti_join(DDF , Dtrain80 , by = "id").....(4)
############################
#C50
names(DDF)
Dtc5_8020 <- C5.0(x = Dtrain80[,c(3:6)], y = as.factor(Dtrain80$result),control =
C5.0Control(noGlobalPruning = TRUE,minCases = seq(1,80,1))) .....(5)
summary(Dtc5_8020) .....(6)
plot(Dtc5_8020) .....(7)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#performanceDc58020
predict.unseenDc58020  <- predict(Dtc5_8020 , Dtest20[,-7] , type = 'class')
confusion.matDc58020 <- table(Dtest20$result,predict.unseenDc58020) .....(8)
confusion.matDc58020
accuract.Dc5test20 <- sum(diag(confusion.matDc58020))/sum(confusion.matDc58020)
accuract.Dc5test20..... .....(9)
library(readxl)
Dtest20 <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")
predict.unseenDc58020 <- predict(Dtc5_8020 , Dtest20 , type = 'class') .....(10)
predict.unseenDc58020.....(11)
#when type is class
Dtest20$class <- predict.unseenDc58020
View(Dtest20) .....(12)
setwd("pathfile/filename.xlsx", sheet = "sheetname")
write.csv(Dtest20 , "Dtester20.csv")
```

อธิบายความหมายของแต่ละขั้นตอนดังนี้

1) library(caTools) ; library(C50) ; library(dplyr) ; library(readxl) ; library(e1071)

   การเรียกใช้ library ที่เกี่ยวข้องในการทำต้นไม้ตัดสินใจ

2) filename <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")

   การเรียกใช้ข้อมูลที่จะนำมาสร้างต้นไม้ตัดสินใจ

3) DDF <- as.data.frame(filename)

   การทำให้ข้อมูลอยู่ในรูปแบบของ data frame ใน R

4) Dtrain80 <- sample_frac(DDF , size = 0.8) และ Dtest20 <- anti_join(DDF , Dtrain80 , by = "id")

   การแบ่งข้อมูลสอน (Training Data) 80% และ การแบ่งข้อมูลชุดทดสอบ (Testing Data) 20 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) Dtc5_8020 <- C5.0การเลือกใช้ Decision Tree C5.0 Algorithm มีการกำหนดตัวควบคุมต่างๆ

6) summary(Dtc5_8020)

แสดงผลการทำต้นไม้การตัดสินใจออกมา

7) plot(Dtc5_8020)

แสดง C5.0 Decision Tree ที่ได้ออกมา

8) predict.unseenDc58020 <- predict(Dtc5_8020 , Dtest20[,-7] , type = 'class')

ทำนายผลข้อมูลชุดทดสอบ (Testing Data) จาก ข้อมูลชุดสอน (Training Data)

9) accuract.Dc5test20

การคำนวณค่าความถูกต้องของการทำนายจากผลของชุดข้อมูลสอนและข้อมูลชุดทดสอบ

10) Dtest20 <- read_excel("pathfile/filename.xlsx", sheet = "sheetname")

นำข้อมูลที่จะ Classification จริงๆนำมาแทนชุดข้อมูลทดสอบ

11) predict.unseenDc58020

ทำนายผลลัพธ์จากข้อมูลจริง

12) View(Dtest20)

เรียกดูผลลัพธ์ที่ได้จากการ classification

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 ตรวจสอบค่าPerformanceของแบบจำลองที่พัฒนาใหม่

<table>
<tr>
<td>

- Confusion matric

predict.unseenDc58020

	close	open
close	171	12
open	18	141

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด และคนบอกว่าร้านปิดจริง ผลลัพธ์ คือ 171 ร้าน

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด แต่คนบอกว่าจริงร้านปิดจริง ผลลัพธ์ คือ 12 ร้าน

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าร้านปิด แต่คนบอกว่าร้านยังเปิดอยู่ ผลลัพธ์ คือ 18 ร้าน

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าร้านเปิด และคนบอกว่าร้านยังเปิดจริง ผลลัพธ์ คือ 141 ร้าน

</td>
<td>

- Recall

True Positive Rate :

$$\frac{171}{171+12} = 0.9344$$

ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านปิด เป็นอัตราส่วน 0.9344 เท่าของร้านปิดทั้งหมด

True Negative Rate :

$$\frac{141}{141+18} = 0.8868$$

ค่าที่บอกว่าโปรแกรมทำนายได้ว่าร้านเปิด เป็นอัตราส่วน 0.8868 เท่าของร้านเปิดทั้งหมด

False Positive Rate :

$$\frac{18}{141+18} = 0.1132$$

ค่าที่บอกว่าโปรแกรมทำนายว่าร้านปิด เป็นอันตราส่วน 0.1132 เท่าของร้านเปิดทั้งหมด

False Negative Rate :

$$\frac{12}{171+12} = 0.0655$$

ค่าที่บอกว่าโปรแกรมทำนายว่าร้านเปิด เป็นอันตราส่วน 0.0655 เท่าของร้านปิดทั้งหมด

</td>
</tr>
<tr>
<td>

- Precision

โปรแกรมทำนายว่าร้านปิดถูกต้อง

$$\frac{171}{171+18} = 0.9048$$

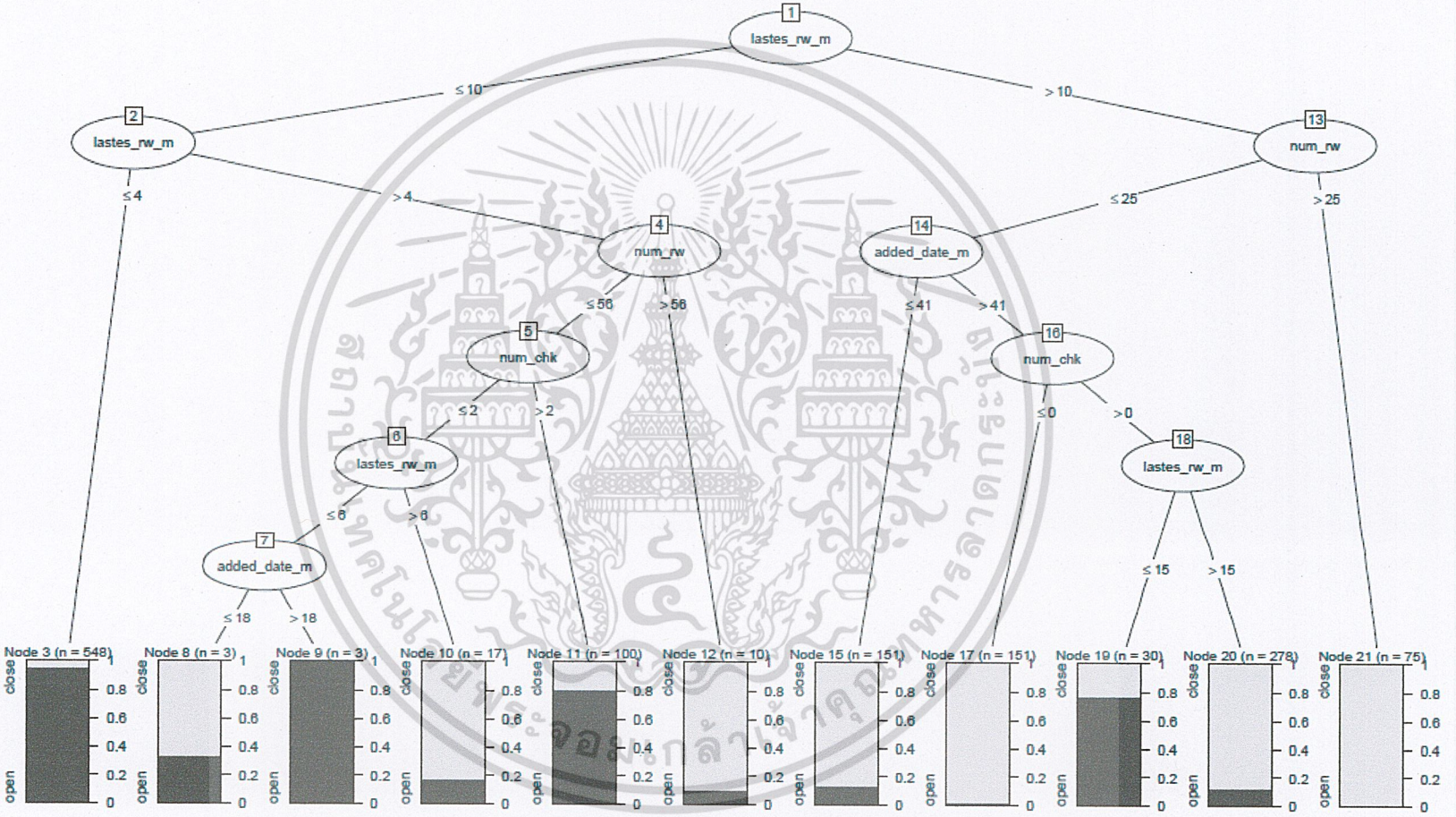
</td>
<td>

- Accuracy

มีค่าเท่ากับ 0.9122807

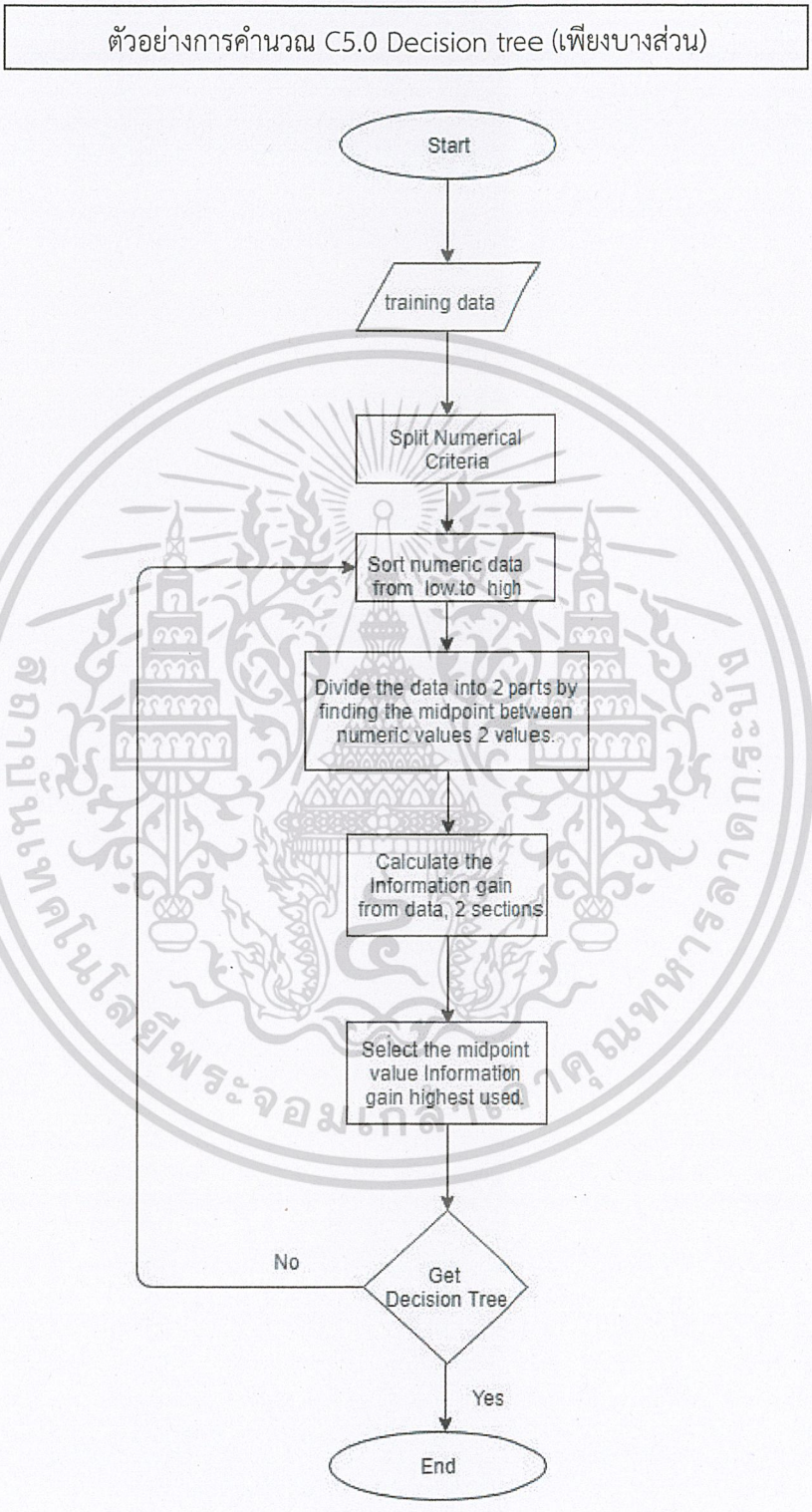
</td>
</tr>
</table>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



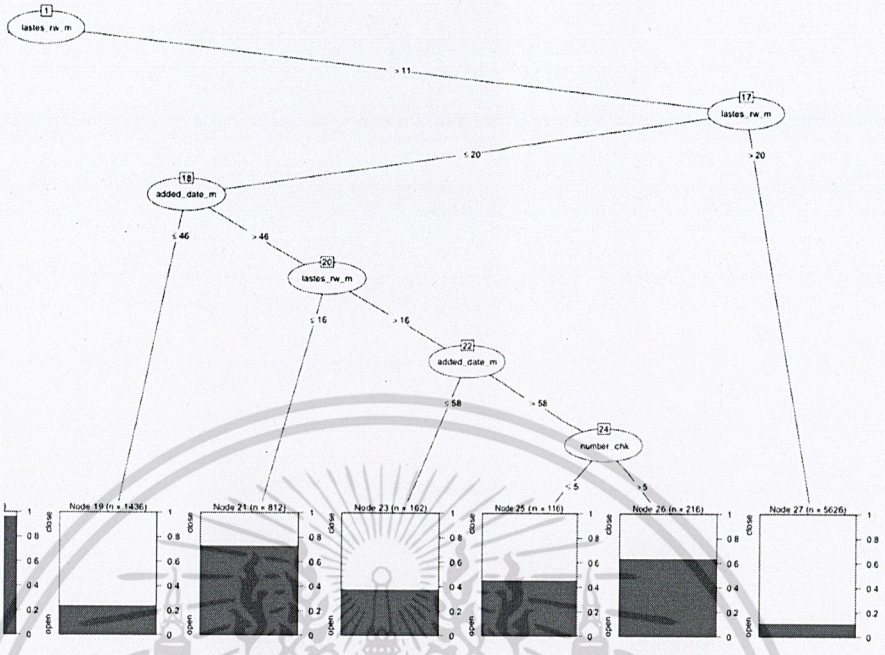
ภาพที่ 25 Decision Tree ของแบบจำลองที่พัฒนาใหม่

เพิ่มเติม



ภาพที่ 26 แสดงขั้นตอนการเลือก Attribute สำหรับตัวเลข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 27 Decision Tree บางส่วนจากการทำแบบจำลองครั้งที่1

- แบ่ง criteria จาก Data Set โดยวิธีการสร้าง Decision Tree สำหรับ atrributeตัวเลข พบว่าค่าInformation Gain ของ lastes_rw_m มีค่ามากสุดจึงนำมาพิจารณาdecision node ต่อไป

lastes_rw_m (≤ 11, > 11) Information Gain =0.39956722	added_date_m (≤ 53, > 53) Information Gain =0.03411338
num_rw (≤ 5, > 5) Information Gain = 0.20542977	number_chk (≤ 3, > 3) Information Gain = 0.16575068

➢ lastes_rw_m (≤ 11) วิธีคิดคำนวณเช่นเดียวกับที่ยกตัวอย่างดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

➢ lastes_rw_m (> 11)

แบ่งข้อมูลที่ lastes_rw_m > 11 โดยวิธีการสร้าง Decision Tree สำหรับ atrribute ตัวเลข

lastes_rw_m (≤ 20, > 20) Information Gain = 0.09613215	added_date_m (≤ 46, > 46) Information Gain = 0.04470011
num_rw (≤ 4, > 4) Information Gain = 0.07463085	number_chk (≤ 2, > 2) Information Gain = 0.06600300

พบว่าค่าInformation Gain ของ lastes_rw_m มีค่ามากสุดจึงนำมาพิจารณาdecision node ต่อไป ซึ่ง lastes_rw_m จะต้องพิจารณาทั้ง ≤ 20 & > 20

➢ lastes_rw_m (≤ 20) ในที่นี่คือเกิดจากการพิจารณาจากlastes_rw_m > 11 แล้วพิจารณาlastes_rw_m <= 20

lastes_rw_m (≤ 16, > 16) Information Gain = 0.02554785	added_date_m (≤ 46, > 46) Information Gain = 0.12712426
num_rw (≤ 4, > 4) Information Gain = 0.10035555	number_chk (≤ 2, > 2) Information Gain = 0.08420140

พบว่าค่าInformation Gain ของ added_date_m มีค่ามากสุดจึงนำมาพิจารณาdecision node ต่อไป

➢ lastes_rw_m (> 20)

โดยที่ minCases ครบกรณีที่กำหนดไว้ จึงทำให้โหนดนี้ไม่ต้องทำการแตกกิ่งออกไปแล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
ผลการทดสอบและอภิปรายผล

4.1 ผลการทดสอบแบบจำลอง

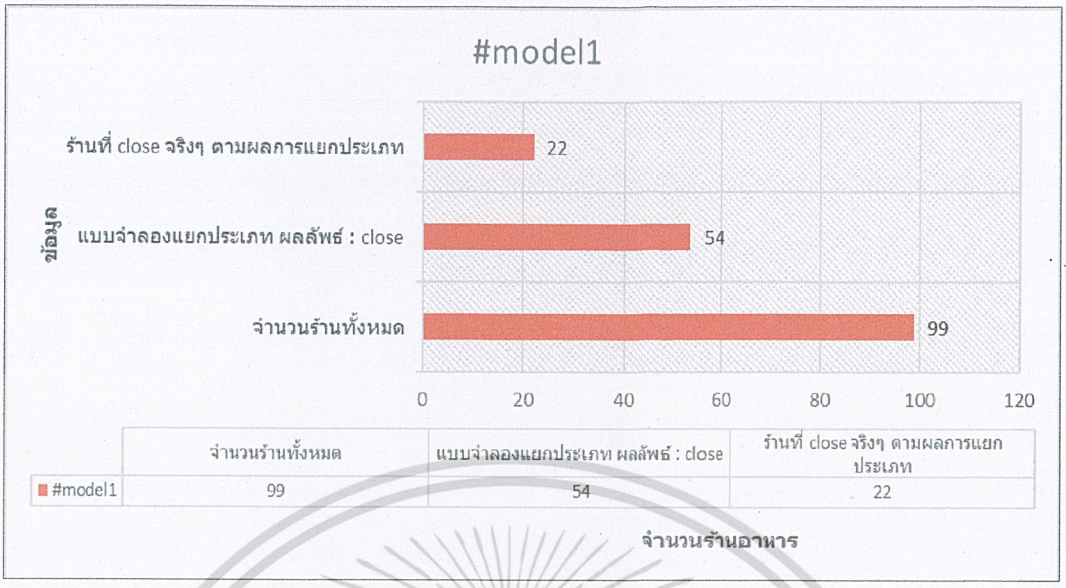
จากการที่นำแบบจำลองที่ได้มาทดลองใช้งานจริงของแบบจำลองจากการทำครั้งที่1 มีการเตรียมข้อมูลโดยการใช้ข้อมูลของร้านอาหารที่ปิดไปแล้วทั้งระบบมาใช้ในการสร้างแบบจำลอง เมื่อนำแบบจำลองที่ได้มาใช้งานจริงกับร้านอาหารในโซนปทุมวัน-วัฒนา จำนวน 99 ร้าน แบบจำลอง Classification ว่าร้านปิดกิจการ 54 ร้าน และพบว่าร้านปิดกิจการจริงๆตามผล Classification จำนวน 22 ร้าน หรือคิดเป็น 40.74 % ของร้านที่จำแนกผลออกมาว่าร้านปิดกิจการทั้งหมดและปิดกิจการจริงๆจากที่นำมาทดลองการใช้งาน

การพัฒนาแบบจำลองยังคงมุ่งเน้นไปใช้งานจริงกับย่านปทุมวัน-วัฒนา (กลุ่มของร้านใจกลางเมือง) เนื่องจากว่าทั้งสองย่านนั้นมีการปิดกิจการของร้านอาหารมากที่สุดในระบบของเว็บไซต์วงในรวมถึงเป็นย่านที่มีความหลากหลายของกลุ่มผู้บริโภค ความหลากหลายของร้านอาหาร ย่านศูนย์การค้า ย่านธุรกิจต่างๆ เป็นต้น ซึ่งก่อนที่จะพัฒนาการสร้างแบบจำลอง ได้มีการจัดกลุ่มของข้อมูลสำหรับการนำมาสร้างแบบจำลองโดยการเลือกร้านในเขตปทุมวันและวัฒนาโดยเฉพาะมาทำการสร้างตัวแบบจำลอง เมื่อนำแบบจำลองที่ได้มาใช้งานจริงกับร้านอาหารในโซนปทุมวัน-วัฒนา จำนวน 62 ร้าน แบบจำลอง Classification ว่าร้านปิดกิจการ 45 ร้าน และพบว่าร้านปิดกิจการจริงๆตามผล Classification จำนวน 28 ร้าน หรือคิดเป็น 62.22% ของร้านที่จำแนกผลออกมาว่าร้านปิดกิจการทั้งหมดและปิดกิจการจริงๆจากที่นำมาทดลองการใช้งาน

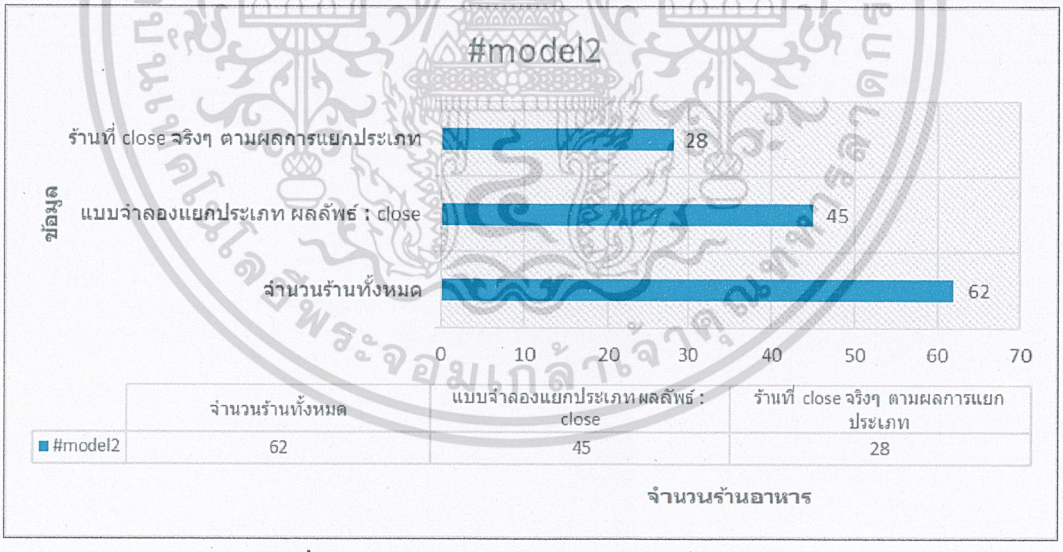
ซึ่งผลลัพธ์ที่เคยได้ทดลองทำโปรเจคแบบจำลองเรื่องการปิดกิจการของร้านอาหารด้วยวิธีการรูปแบบอื่นๆ ได้ผลลัพธ์สำเร็จจากการClassification ร้านที่ปิดกิจการในระบบเป็น 5 % ของร้านที่จำแนกผลออกมาว่าร้านปิดกิจการทั้งหมดและปิดกิจการจริงๆจากที่นำมาทดลองการใช้งาน

ทั้งนี้ ผลการทดสอบที่แสดงผลจากแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วย วิธีการต้นไม้ตัดสินใจ จะได้ผลลัพธ์ที่ดีขึ้นได้นั้นอาจจะต้องมีการศึกษาปัจจัยต่างๆที่หลากหลายในเชิงลึก อาจจะต้องใช้ระยะเวลาในการทำแบบจำลองที่ยาวนานกว่านี้ หรือต้องมีการระดมความคิดจากหลากหลายทีมอาทิ ทีมการตลาด ทีมบริหาร ของในบริษัท เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาพที่ 28 แสดงผลจากการใช้แบบจำลองจากการทำครั้งที่ 1



ภาพที่ 29 แสดงผลจากการใช้แบบจำลองที่พัฒนาใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

reviewed_it	CONCAT("v	name	num_rw	number_ch	max(w_revi	lastes_rw_t	added_date	added_date	X__1	43373	predict	check
35	www.wongnai.c	แป้นปลื้ม ก๋วยเตี๋	11	13	1/1/2017 17:15	20	2/9/2010 0:00	96	NA	NA	close	open
44	www.wongnai.c	Akbar	1	1	21/3/2015 8:30	42	5/10/2010 18:3	95	NA	NA	close	open
45	www.wongnai.c	Al-Hussain Res	2	1	12/1/2016 22:2	32	5/10/2010 18:3	95	NA	NA	close	open
54	www.wongnai.c	Ashibi	3	0	7/1/2015 20:04	44	5/10/2010 18:3	95	NA	NA	close	open
76	www.wongnai.c	Coffee Alley in	13	9	6/10/2015 23:4	35	5/10/2010 18:3	95	NA	NA	close	close
123	www.wongnai.c	La Villa	1	0	31/8/2015 5:42	36	5/10/2010 18:3	95	NA	NA	close	close
144	www.wongnai.c	New Bukhara I	3	7	14/8/2016 9:47	25	5/10/2010 18:3	95	NA	NA	close	open
149	www.wongnai.c	Offshore Fish a	2	0	9/4/2015 20:32	41	5/10/2010 18:3	95	NA	NA	close	close
218	www.wongnai.c	Tropical Max R	7	2	8/1/2017 14:05	20	5/10/2010 18:3	95	NA	NA	close	open
402	www.wongnai.c	Hong Teh Chin	2	0	7/6/2016 19:48	27	5/10/2010 18:3	95	NA	NA	close	open
410	www.wongnai.c	เรือนนาริ สุขุมวิท	2	0	1/10/2014 19:49	47	5/10/2010 18:3	95	NA	NA	close	open
1474	www.wongnai.c	รัมรญา	1	1	30/5/2013 8:54	64	5/10/2010 18:3	95	NA	NA	close	open
1762	www.wongnai.c	คลาสสิค เบลล์	1	1	9/7/2015 20:24	38	5/10/2010 18:3	95	NA	NA	close	open
1793	www.wongnai.c	คายาคิมเกาหลี	3	1	5/7/2014 20:45	50	5/10/2010 18:3	95	NA	NA	close	open
1800	www.wongnai.c	Cassia Cafe &	6	3	7/4/2015 17:57	41	5/10/2010 18:3	95	NA	NA	close	open
1902	www.wongnai.c	ยามไทย	7	6	1/10/2016 16:31	23	5/10/2010 18:3	95	NA	NA	close	open
1943	www.wongnai.c	วัมไทมัลฟอฟี่	1	0	5/8/2013 13:30	61	5/10/2010 18:3	95	NA	NA	close	open
1998	www.wongnai.c	อิน บาร์	2	1	2/12/2016 10:4	21	5/10/2010 18:3	95	NA	NA	close	close
2190	www.wongnai.c	ลำนานไทย	1	1	1/12/2017 0:40	9	5/10/2010 18:3	95	NA	NA	close	open
2420	www.wongnai.c	Burendo	13	4	1/12/2015 1:24	33	5/10/2010 18:3	95	NA	NA	close	close
2447	www.wongnai.c	บ้านตะเกียงไทย	1	0	28/4/2012 8:31	77	5/10/2010 18:3	95	NA	NA	close	close
2470	www.wongnai.c	บ้านสวนอโศก	2	2	2/12/2016 20:54	21	5/10/2010 18:3	95	NA	NA	close	open
2626	www.wongnai.c	ยามาเนะ	1	0	3/9/2016 15:13	24	5/10/2010 18:3	95	NA	NA	close	close
3016	www.wongnai.c	เอ็น หูฉลาม	0	0	3/7/2016 16:02	26	1/10/2010 15:31	95	NA	NA	close	close
3022	www.wongnai.c	Circle Restaura	2	4	6/8/2014 11:02	49	1/10/2010 15:31	95	NA	NA	close	open
3038	www.wongnai.c	เดอะการาจ เบอร์	15	8	0/7/2016 12:31	26	1/10/2010 15:31	95	NA	NA	close	open
3082	www.wongnai.c	Chapter 31	5	2	7/8/2015 23:07	37	1/10/2010 15:31	95	NA	NA	close	close
3086	www.wongnai.c	Trattoria Felice	2	0	9/11/2015 23:0	34	1/10/2010 15:31	95	NA	NA	close	close
3093	www.wongnai.c	แลมปาดัง คาเฟ่	1	1	7/6/2013 13:39	63	1/10/2010 15:31	95	NA	NA	close	close
3105	www.wongnai.c	Korean BBQ St	2	0	25/2/2013 9:07	67	1/10/2010 15:31	95	NA	NA	close	open
3155	www.wongnai.c	Dainoki Restau	1	0	0/2/2018 10:33	7	1/10/2010 15:31	95	NA	NA	close	open
3213	www.wongnai.c	Kitchen Paradis	0	0	0/8/2011 13:03	85	1/10/2010 15:31	95	NA	NA	close	close
3399	www.wongnai.c	สวัสดี โรงแรม In	2	1	2/8/2016 13:06	25	1/10/2010 15:31	95	NA	NA	close	open
3514	www.wongnai.c	Wagyu Chyubo	1	0	8/11/2010 11:3	94	8/11/2010 11:1	94	NA	NA	close	close
3598	www.wongnai.c	Mai Izakaya	1	0	19/11/2017 1:09	10	3/11/2010 15:58	94	NA	NA	close	close
3617	www.wongnai.c	Tessa Homema	1	0	19/4/2011 10:49	89	7/11/2010 16:21	94	NA	NA	close	open
3695	www.wongnai.c	Anego Izakaya	3	0	7/10/2013 10:5	59	1/8/2010 0:00	97	NA	NA	close	open
3703	www.wongnai.c	Aphrodite Inn	1	0	8/1/2016 12:56	32	1/8/2010 0:00	97	NA	NA	close	open
3743	www.wongnai.c	Au Bon Pain	8	8	1/1/2017 19:26	19	1/8/2010 0:00	97	NA	NA	close	open
3759	www.wongnai.c	La Brioche the	8	10	6/11/2016 23:5	22	1/8/2010 0:00	97	NA	NA	close	close
3767	www.wongnai.c	Bar Eleven	1	0	23/5/2013 1:30	64	1/8/2010 0:00	97	NA	NA	close	close
3776	www.wongnai.c	Bawarchi Alif L	2	3	6/10/2014 9:58	47	1/8/2010 0:00	97	NA	NA	close	open
3839	www.wongnai.c	Cafe Fish	11	6	11/10/2015 22:	35	1/8/2010 0:00	97	NA	NA	close	close
3867	www.wongnai.c	Chilli Culture Th	2	1	1/7/2015 22:14	37	1/8/2010 0:00	97	NA	NA	close	open
3873	www.wongnai.c	Cinnamon Rest	2	0	9/7/2015 16:19	38	1/8/2010 0:00	97	NA	NA	close	open
3920	www.wongnai.c	Darak	2	3	11/2/2018 19:2	7	1/8/2010 0:00	97	NA	NA	close	open
3933	www.wongnai.c	Det 5 Restaura	2	2	1/11/2015 19:30	34	1/8/2010 0:00	97	NA	NA	close	open
3940	www.wongnai.c	Dragon Seafoo	1	0	9/3/2018 15:07	6	1/8/2010 0:00	97	NA	NA	close	open
3958	www.wongnai.c	Espresso	1	0	9/9/2017 18:17	12	1/8/2010 0:00	97	NA	NA	close	open
3969	www.wongnai.c	Fifty Five	7	6	3/6/2015 9:02	39	1/8/2010 0:00	97	NA	NA	close	close
4034	www.wongnai.c	Gourmandises	7	4	5/12/2016 8:33	21	1/8/2010 0:00	97	NA	NA	close	close
4067	www.wongnai.c	Hanako 2 Japa	19	12	8/8/2016 21:36	25	1/8/2010 0:00	97	NA	NA	close	close
4100	www.wongnai.co	iberry Cafe Sia	2	0	12/10/2014 0:3	47	1/8/2010 0:00	97	NA	NA	close	close
3845	www.wongnai.c	Cafe Primo	1	0	1/1/2014 23:14	56	1/8/2010 0:00	97	NA	NA	close	close

ภาพที่ 30 ผลลัพธ์จากการใช้แบบจำลองจากการทำครั้งที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

reviewe	name	street_a	hint	rank_sc	CONCAT	average	total_nu	max(w_review.reviewed_dat	class	Open/Cl
191607	Gram Gourr	ถ.สุขุมวิท	ชั้น 7 zone t	1400.4585	www.wong	3.2222222	9	13/6/2016 10:32	close	close
191647	Eighteen be	ถ.สุขุมวิท	ชั้น 6 โซน H	716.21442	www.wong	3.2	5	3/7/2015 0:42	close	close
191654	Bella Ducci	643 ถนน สุข	ชั้น B EmQu	1201.9284	www.wong	3.3333333	6	27/4/2016 19:30	close	close
191662	Billion beef	ถนนสุขุมวิท	ชั้น B	311.67743	www.wong	4	1	10/4/2015 20:58	close	close
191680	Stickhouse	ถนน สุขุมวิท	The Emquar	2294.7485	www.wong	3.5294118	15	16/6/2016 20:51	close	close
192379	สุกี้นัมเบอร์วั	สุขุมวิท	Food Hall ชั้	284.35265	www.wong	2.6666667	3	16/1/2017 23:01	close	open
196240	moonne	ถนน สุขุมวิท	ตู้เค้กบริเวณ	1596.6127	www.wong	3.6	4	15/5/2016 16:59	close	close
197141	fucheer	ถนนสุขุมวิท	ชั้น 7 EmQu	1388.9972	www.wong	3.1538462	13	3/6/2017 18:29	close	close
200267	the brownie	สุขุมวิท	ร้านอยู่ชั้น G	1053.053	www.wong	3.6666667	5	16/10/2015 19:27	close	close
200394	Zaichi	The Emquat	BTSพร้อมพง	3376.3444	www.wong	3.5263158	17	20/4/2017 14:47	close	close
201533	ข้าวมันไก่โก	ถนน สุขุมวิท	ชั้น B Foodh	646.42434	www.wong	3.5	4	6/9/2017 14:26	close	open
203055	PALA	EmQuartier	โซน Take H	274.70275	www.wong	4	1	4/7/2015 16:36	close	close
204948	วลีย์ มังสวิรัติ	ถนน สุขุมวิท	emquartier f	728.69321	www.wong	4	2	30/7/2017 17:09	close	open
212137	ต้มยำกุ้งคุณก	ถนน สุขุมวิท	Emquartier	359.79628	www.wong	4	1	20/9/2016 3:16	close	open
214672	Tokyo Swee	ถนน สุขุมวิท	อยู่บริเวณโซ	662.52717	www.wong	3.3333333	2	20/9/2016 23:32	close	open
218156	Lapin Vanill	ชั้น G Emqu	โซนขนมไซโ	147.20799	www.wong	3	1	15/12/2015 1:01	close	close
222114	Company B	The Emquar	ชั้น G โซนขา	1155.9018	www.wong	4	4	31/7/2016 12:09	close	open
226847	Partytime	ถนน สุขุมวิท	ชั้น B ใกล้ S	58.822961	www.wong	2	1	5/3/2016 23:18	close	close
230814	บะหมี่สว่าง ห	ถนน สุขุมวิท	ใต้ตึกตรงสว	130.65514	www.wong	2.5	1	14/4/2016 15:26	close	open
230879	Love Bar	ถ.สุขุมวิท	1F Waterfall	59.114252	www.wong	4.5	1	13/5/2016 12:47	close	open
233776	ขนมหญ้าหวา	ถนน สุขุมวิท	เดินทางด้วยร	3.2009648	www.wong	5	1	12/5/2016 13:52	close	close
236791	ก๋วยเตี๋ยวเป็ด	ถนน สุขุมวิท	Food court	424.89475	www.wong	4	1	19/7/2016 20:12	close	open
238735	Kiss The Tir	ถนน สุขุมวิท	walking dow	1645.1803	www.wong	3.7333333	11	9/8/2017 22:46	close	close
240123	Minamoto K	ถนน สุขุมวิท	หน้า Gourme	765.86116	www.wong	4.25	1	14/8/2016 2:21	close	close
240703	แซ่บ โดย ส	ถนน สุขุมวิท	Foodhall em	750.97988	www.wong	4	1	12/7/2016 22:38	close	open
244246	Yaki Yaki	ถนน สุขุมวิท	ร้านอยู่ชั้น G	161.05688	www.wong	3	1	11/8/2016 20:13	close	close
244377	Owlvee Soft		ชั้น B Food	720.55736	www.wong	3	5	13/6/2017 19:27	close	close
244853	Cookie Cho	ถนน สุขุมวิท	ชั้น G ร้านเป็	430.6941	www.wong	3.3333333	3	8/10/2016 19:00	close	close
248468	Betagro	ถนน สุขุมวิท	Gourmet ma	371.63878	www.wong	4	1	19/9/2016 17:11	close	open
249281	Royal India	ถนน สุขุมวิท	ชั้น B อยู่ใน	263.77487	www.wong	4	1	28/9/2016 22:36	close	open
258534	จ๊าดล่า	ถนน สุขุมวิท	หน้า Gourme	326.14609	www.wong	3	2	14/3/2017 19:03	close	close
262980	ถ้วยแปดคุณพี่	ถนน สุขุมวิท	ร้านอยู่หน้า g	407.79461	www.wong	3.5	2	30/11/2017 23:42	close	close
273732	Dessert Bar		Floor B @E	140.66719	www.wong	3	1	19/4/2017 20:19	close	close
273818	Kindee	ถนน สุขุมวิท	อยู่ใน Quarti	218.2209	www.wong	3	2	6/6/2017 15:29	close	open
277234	Chicka Che	Emquartier	ชั้นG หน้า Foo	194.83442	www.wong	2.6666667	1	17/5/2017 13:37	close	close
281698	Pomme no	ถนน สุขุมวิท	ชั้น B Food	793.18465	www.wong	3.6	4	4/3/2018 21:42	close	open
282993	Shion - Sus	ถนน สุขุมวิท	ชั้น G, EmQ	1212.0641	www.wong	3.5	2	14/8/2017 7:10	close	open
287165	ขนมครกแม่เ	ถนน สุขุมวิท	Emquartier	1175.3878	www.wong	4	3	17/9/2017 19:33	close	open
287677	made in candy		ชั้นG หน้าทา	143.35406	www.wong	3	1	12/8/2017 17:16	close	close
288435	nomi mono	ถนน สุขุมวิท	ชั้น B Emqu	1429.3114	www.wong	4.25	4	13/9/2017 10:59	close	close
299032	Sweet Powd	ถนน สุขุมวิท	อยู่ชั้น B ติด	158.17727	www.wong	4	1	16/11/2017 4:04	close	close
300531	Homsuwan	The EmQua	ชั้น B โซน Q	328.47985	www.wong	3	2	27/1/2018 10:17	close	open
306334	ไก่ต้มเงินล้าน	ถนน สุขุมวิท	ขายอยู่ในหม่	411.1356	www.wong	4	1	25/1/2018 10:29	close	close
310274	หมูทอดเบรค	ถนน สุขุมวิท	ใน Quartier	179.52641	www.wong	3	1	23/2/2018 18:23	close	close
323874	ขญามิน	ถนน สุขุมวิท	อยู่ชั้น B ตรง	211.12942	www.wong	3	1	23/3/2018 5:25	close	close

ภาพที่ 31 ผลลัพธ์จากการแบบจำลองที่พัฒนาใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การอภิปรายผล

จากการที่ได้ทราบค่าการทดสอบแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหารด้วยวิธีการต้นไม้ตัดสินใจ แล้วนั้น สามารถจำแนกร้านอาหารที่เปิดและปิดได้ดีขึ้นจาก 5% ไปสู่ 40.74% และพัฒนามาได้จนถึง 62.22 % ได้นั่น ทั้งนี้ในอนาคตจะต้องมีการพัฒนา ปรับปรุง แก้ไขแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหาร ให้มีค่าความจำแนกประเภทของร้านอาหารที่เปิดและร้านอาหารที่ปิด ถูกต้องเพิ่มมากขึ้นยิ่งๆขึ้นไป โดยอาจจะต้องคำนึงปัจจัยที่หลากหลายและครอบคลุมให้มากขึ้น ศึกษาเชิงลึกไปถึงกลไกของตลาดในแต่ละช่วง วิเคราะห์ในเชิงลึกที่รวมไปถึงทั้งประเภทร้านอาหาร เขต/ย่านของร้านอาหาร ระดับความนิยมของร้านอาหาร ความนิยมที่เป็นกระแส ณ ตอนนั้นๆ เป็นต้น

ผลลัพธ์ที่ข้าพเจ้าได้จากแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหาร ด้วยวิธีการต้นไม้ตัดสินใจ ในครั้งนี้แม้อาจจะได้ผลลัพธ์ที่ไม่ได้สูงมาก แต่ ผลลัพธ์ที่ได้ก็มากขึ้นจากครั้งก่อนที่เคยได้ทดลองทำ ทำให้แสดงให้เห็นว่าการที่ได้ศึกษาและลงมือปฏิบัติ แบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหาร นั้นมีความเป็นไปได้มากขึ้นกว่าเดิม ไม่ได้เป็นเพียงความคิดแค่ในอุดมคติ แต่เป็นความคิดที่สามารถทำได้จริง โดยหลักการสำคัญที่จะช่วยทำให้ผลลัพธ์ดียิ่งขึ้น คือ คณิตศาสตร์ เป็นศาสตร์ที่สำคัญที่จะช่วยสร้างเครื่องมือต่างๆให้มนุษย์ได้ทำงานได้ง่ายยิ่งขึ้น ช่วยลดระยะเวลาในการทำงาน และ ช่วยให้เกิดประสิทธิภาพการทำงานที่ถูกต้องมายิ่งขึ้นๆไป

ภายหลังจากที่ข้าพเจ้าได้ศึกษาและได้ทำแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหาร ด้วยวิธีการต้นไม้ตัดสินใจ ทำให้รู้และเข้าใจอย่างลึกซึ้งถึงความสำคัญของการทำเรื่องของร้านอาหารที่ปิดบริการไปนั่น เป็นงานที่ละเอียดอ่อนมากๆ เป็นงานที่ต้องใช้ความรอบคอบสูงมากๆ เพราะถ้าหากทำผิดพลาด นั้นจะส่งผลเสียต่อทั้งร้านอาหาร ส่งผลเสียต่อบริษัท ส่งผลเสียต่อผู้ใช้งานที่ไม่สามารถไปใช้บริการได้ตามความต้องการ และถ้าในอนาคตเราพัฒนาในเรื่องของเครื่องมือที่จะมารองรับการใช้งานในเรื่องของร้านอาหารที่ปิดกิจการ ช่วยทำให้ทำงานได้ง่ายขึ้นและมีความถูกต้องมากยิ่งขึ้น แม้ว่าอาจจะต้องใช้กำลังคน ใช้เวลาในการพัฒนา แต่ข้าพเจ้ามั่นใจว่ามันจะต้องส่งผลที่ดีขึ้นมากๆให้กับทุกๆฝ่ายที่เกี่ยวข้องอย่างแน่นอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

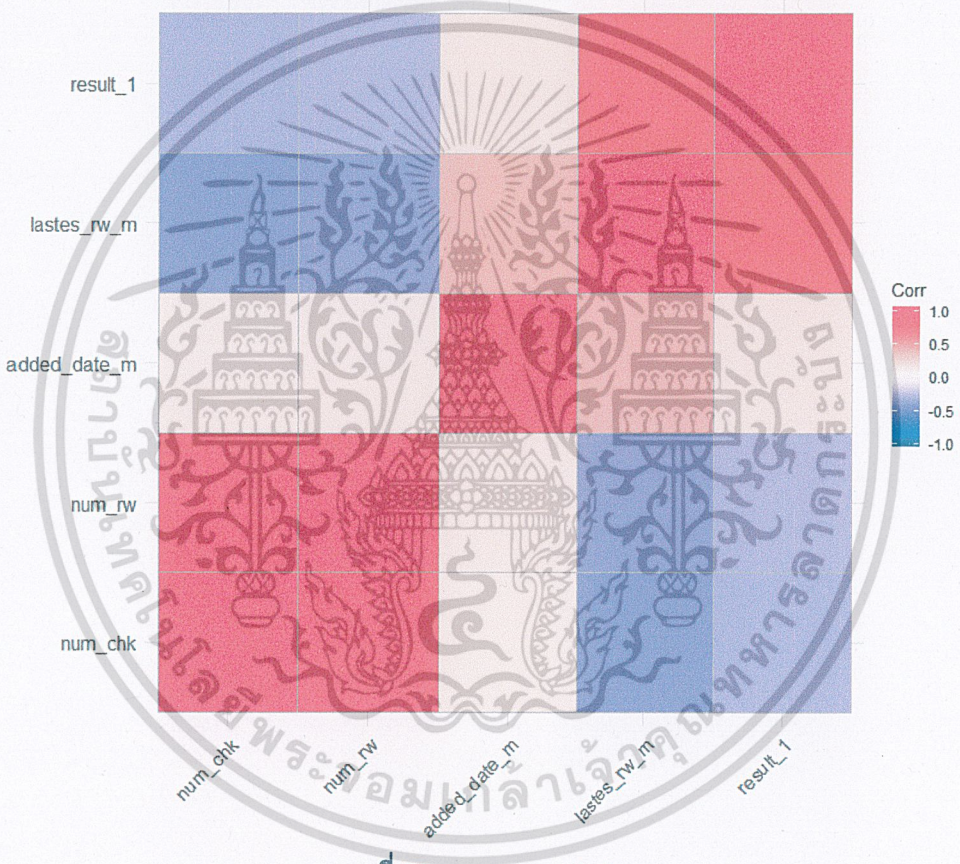
4.3 การอภิปรายผลเพิ่มเติม

ภายหลังจากการทำแบบจำลองแล้วได้รับข้อเสนอแนะจากอาจารย์ที่ปรึกษาให้นำเรื่อง correlation มาประยุกต์ใช้สำหรับงานชิ้นนี้ ได้ผลลัพธ์ดังนี้

ตรวจสอบ correlation ของตัวแปรที่นำมาใช้

ตัวแปร x : num_chk , num_rw , added_date_m , lastes_rw_m

ตัวแปร y : result



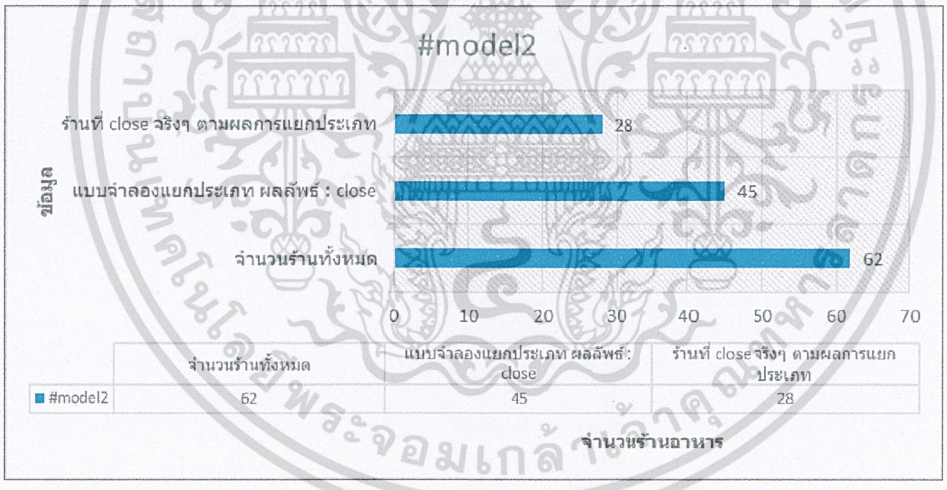
ภาพที่ 32 ตาราง correlation

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นำ added_date_m , lastes_rw_m มาพิจารณา โดยใช้ decision tree ได้ผลลัพธ์ดังนี้

ภาพที่ 33 แสดงผลจากการใช้แบบจำลองที่ตัดตัวแปรจากการตรวจสอบcorrelation

แต่ เมื่อนำมาเปรียบเทียบผลลัพธ์กับอันเดิม



ภาพที่ 34 แสดงผลจากการใช้แบบจำลองที่พัฒนาใหม่

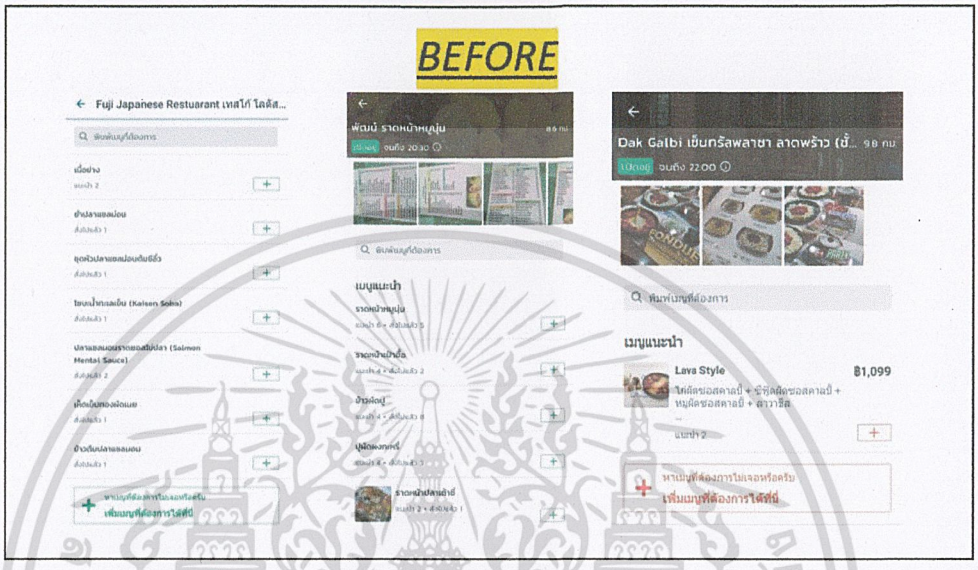
สรุปผลได้ว่า correlation ไม่ได้มีผลต่อการสร้างแบบจำลองด้วยวิธีการต้นไม้ตัดสินใจ เนื่องจากว่า สำหรับการสร้างต้นไม้ตัดสินนั้นได้มีขั้นตอนวิธีในการคัดเลือกความสำคัญของตัวแปรว่าจะนำตัวแปรใดมาพิจารณาก่อนซึ่งเรียกว่า entropy ดังที่ได้กล่าวไปแล้วตามบทที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

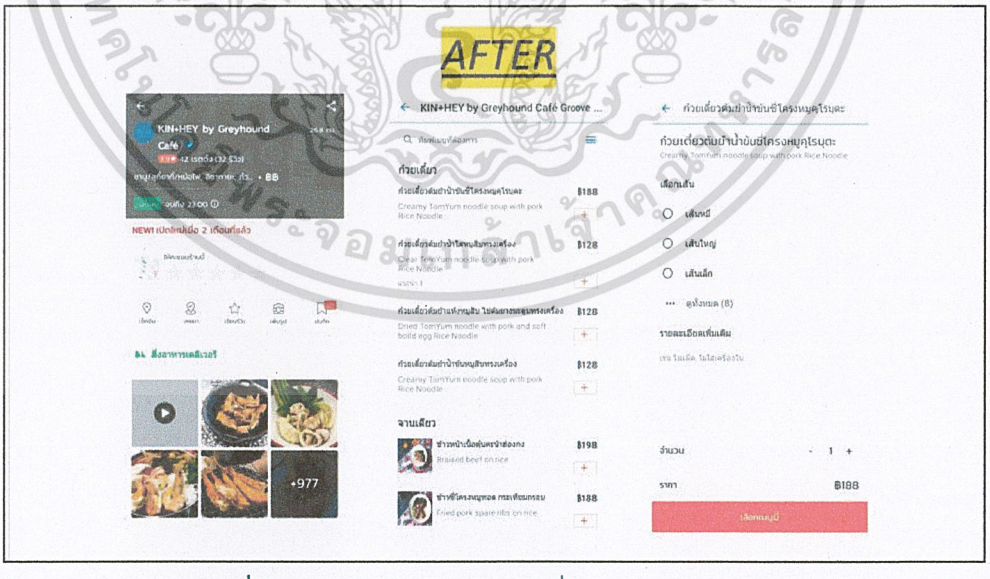
4.4 งานที่ได้รับมอบหมายเพิ่มเติม

- LineMan Squad

ได้ทำe-menu ของบริการ linemanxwongnai ได้ทำไปทั้งสิ้น 10,761 เมนู 130 ร้าน



ภาพที่ 35 หน้าต่าง application ก่อนยังไม่มี e-menu lineman



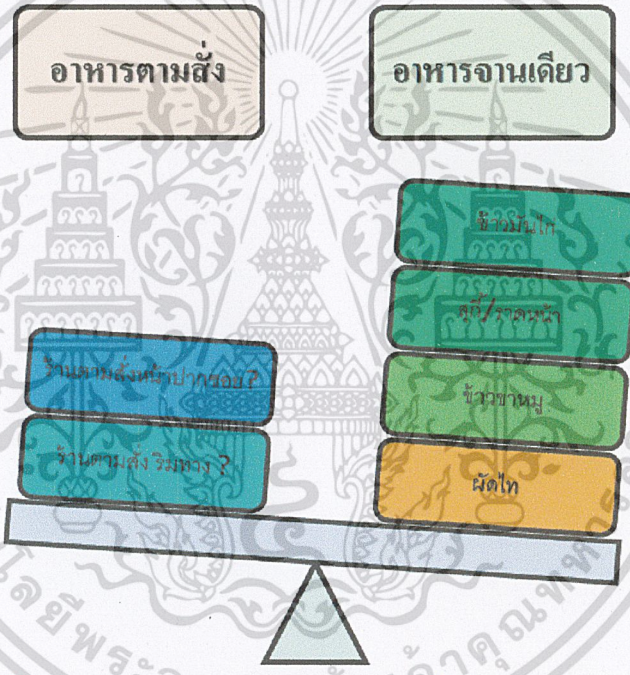
ภาพที่ 36 หน้าต่าง application เมื่อมี e-menu lineman

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- **Search Squard (Development Team & Data Operations Team)**

ได้รับมอบหมายให้จำแนกประเภทข้อมูลของร้านอาหาร ประเภทอาหารจานเดียว/อาหารตามสั่ง เพื่อที่จะได้นำข้อมูลที่จำแนกประเภทไปทำพัฒนาต่อในเรื่องของระบบ search engine เว็บไซต์ wongnai.com

โดยในช่วงระยะเวลาสหกิจศึกษาข้าพเจ้าได้ทำจัดการจำแนกข้อมูลไปทั้งสิ้น 11,300 ร้าน

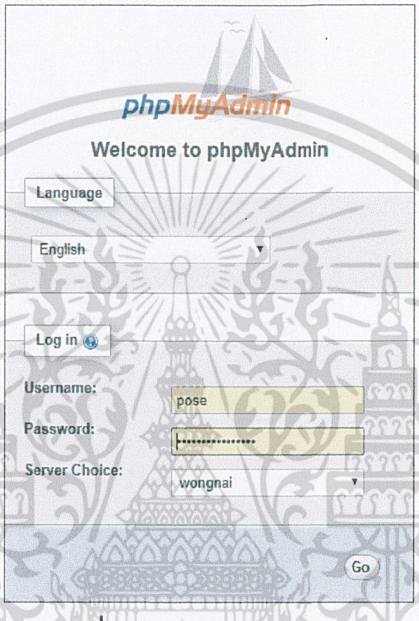


ภาพที่ 37 การคัดแยกประเภทของอาหารตามสั่งและอาหารจานเดียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Approved Reviews

การที่ข้าพเจ้าได้มาทำสหกิจศึกษาใน Data Operations Team ได้ใช้งาน SQL จริงๆ เพราะจะต้องนำเอาข้อมูลของรีวิวที่เข้ามาในระบบของเว็บไซต์มาตรวจสอบและพิจารณาว่ารีวิวมีคุณภาพเหมาะสมที่จะได้ลงเว็บไซต์หรือไม่ และถ้าหากตรวจสอบแล้วพบเจอรีวิวที่ไม่ได้คุณภาพหรือไม่เหมาะสมจะต้องทำการแจ้งผู้รีวิวให้ทราบถ้าหากจะต้องลบรีวิวที่ผู้ใช้งานรีวิวเข้ามา



ภาพที่ 38 หน้าต่าง log in ระบบฐานข้อมูล

1. เข้า MySQL และ query รีวิวทั้งหมดของวันโดยใช้โค้ดด้านล่าง

```
SELECT rv.reviewed_date, u.rank, rv.rating, rv.quality, rv.reviewer_id,
u.name, rv.reviewed_item_id, r.name,
concat("https://www.wongnai.com/reviews/",rv.eid) as link,
rv.summary ,rv.description
FROM w_review rv
inner join w_user u on rv.reviewer_id = u.id
inner join w_restaurant r on r.id = rv.reviewed_item_id
where rv.reviewed_date >= "YYYY-M-DD"
and rv.reviewed_date < "YYYY-M-DD"
order by rv.reviewed_date asc
```

ภาพที่ 39 ตัวอย่างโค้ดในการ query

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการทำสหกิจศึกษาและข้อเสนอแนะ

ในบทนี้จะเป็นการกล่าวสรุปผลการไปทำสหกิจศึกษา และข้อเสนอแนะต่างๆที่สมควรต่อการแก้ไข ซึ่งมีดังต่อไปนี้

5.1 สรุปผลการทำสหกิจศึกษา

การทำสหกิจศึกษาครั้งนี้ผู้จัดทำได้ทำงานใน Data Operations Team ซึ่งภายในระยะเวลา 4 เดือน (ตั้งแต่วันที่ 7 สิงหาคม พ.ศ. 2561 ถึง วันที่ 14 ธันวาคม พ.ศ. 2561) ผู้จัดทำได้รับมอบหมายให้ทำแบบจำลองวิเคราะห์แนวโน้มการเลิกกิจการของร้านอาหาร ด้วยวิธีการต้นไม้ตัดสินใจ เพื่อใช้คัดแยกประเภทของร้านอาหารที่มีการเปิดหรือปิดในระบบและถ้าหากพบข้อมูลของร้านที่ไม่เป็นประโยชน์ต่อผู้ใช้งานจะได้สามารถแก้ไขปรับปรุงข้อมูลให้เป็นประโยชน์ต่อผู้ใช้งานเว็บไซต์สูงสุด

และยังได้รับประโยชน์จากการทำสหกิจศึกษาดังนี้

1. ได้ประสบการณ์ในการทำงานภายในองค์กร ทำให้มีความพร้อมในการเริ่มต้นชีวิตการทำงาน
2. ได้ศึกษาการทำงานจริง กับคนทำงานจริงๆ ให้ได้รู้และเข้าใจการทำงานในสายงานได้มากขึ้น
3. ได้เสริมทักษะการสื่อสารข้อมูลและเรียนรู้การทำงานร่วมกับผู้อื่น
4. ทำให้มีความรับผิดชอบในหน้าที่การงานที่ได้รับมอบหมาย
5. ได้เรียนรู้จัดการเวลา แบ่งการทำงาน การจัดลำดับความสำคัญของงาน ได้ดีมากขึ้น
5. สามารถนำความรู้ความสามารถที่ได้รับจากการทำสหกิจศึกษาไปต่อยอดในอนาคต
6. ได้รับค่าตอบแทนจากสถานประกอบการตามที่สถานประกอบการเห็นสมควร
7. ได้เห็นถึงความสำคัญของความรู้ที่เรียนมา ว่าสามารถนำมาประยุกต์ใช้ได้จริงๆกับงานที่ได้รับมอบหมาย
8. ได้รู้จักกับปัญหา และ สามารถจัดการแก้ไขปัญหาได้อย่างถูกต้อง ให้ผ่านพ้นไปได้ด้วยดี
9. ได้พบเพื่อนใหม่ๆ ต่างสถาบัน ต่างคณะ แลกปลี่ยนเรื่องราวต่างๆที่แต่ละคนเคยพบเจอมา
10. สำหรับ บริษัท วงใน มีเดีย จำกัด ได้มีการจัดกิจกรรม Weshare ทำให้ได้ฟังบรรยายจากคนดังคนเก่ง จากสายงานต่างๆ อาทิเช่น วุฒิธร มิลินทจินดา (Woody) , ดวงฤทธิ์ บุนนาค(พี่ด้วง) , ยุทธนา บุญอ้อม(ป๊าเต็ด) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ปัญหาอุปสรรค

5.2.1 การยื่นเอกสารเกี่ยวกับสหกิจศึกษา น่าจะมีการระบุขั้นตอนไว้อย่างชัดเจน มีแบบแผนว่านักศึกษาที่ต้องการไปสหกิจศึกษาควรปฏิบัติตามกฎ ข้อบังคับ อย่างไร

5.2.2 เนื่องจากตัวข้าพเจ้าเองไม่เคยลงเรียนวิชา database จึงทำให้ต้องมาเรียนรู้การเขียน SQL ใหม่จากสถานประกอบการ

5.2.3 ในบางเรื่องบางส่วนของ การทำและพัฒนาแบบจำลอง เป็นเรื่องที่ใหม่และซับซ้อนต่างจากที่เรียนมาในมหาวิทยาลัย ทำให้ต้องใช้เวลาศึกษาและทำความเข้าใจสักระยะหนึ่ง

5.3 แนวทางแก้ปัญหา

5.3.1 ติดตามข่าวสารเกี่ยวกับการทำสหกิจศึกษาอยู่เสมอเพื่อเตรียมยื่นเอกสารตามระยะเวลาที่กำหนดและควรดำเนินการให้เสร็จสิ้นก่อนหนึ่งถึงสองเดือนที่เริ่มทำงานจริง

5.3.2 ได้มีโอกาสเข้าคลาสเรียนพื้นฐานการเขียน SQL และ มีพี่ๆในทีมช่วยสอนช่วยอธิบายส่วนต่างๆใน database ของสถานประกอบการ

5.3.2 เมื่อผู้จัดทำมีเวลาว่างจะค้นคว้าหาข้อมูลเพิ่มเติมเพื่อให้มีความรู้ความเข้าใจเพียงพอในการทำงาน

5.4 ข้อเสนอแนะ

ทางภาควิชาควรจะทำแบบแผนข้อปฏิบัติ ขั้นตอน การทำสหกิจศึกษาไว้อย่างแน่ชัด เพื่อที่จะให้นักศึกษาที่สนใจการสหกิจศึกษาได้ปฏิบัติได้ถูกต้อง รวมถึงมีการแนะนำ/แนะแนวบริษัทที่เปิดรับสหกิจศึกษาไว้รองรับนักศึกษา หากมีปัญหาข้อผิดพลาดที่ไม่มีสถานประกอบการติดต่อรับนักศึกษาในการทำสหกิจศึกษา ที่สำคัญความรู้บางวิชาหรือเนื้อหาบทเรียนการสอนที่นำมาใช้ประโยชน์จริงๆกับการทำงาน ภาควิชาน่าจะสอนให้เข้มข้นมากขึ้นว่าเดิม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] Life@Wongnai.**เรื่องราวการทำงานของชาววงใน**. [Online].เข้าถึงเมื่อวันที่ 1 ก.ย.2561 https://life.wongnai.com/

[2] นนทวัฒน์ บุญบา.**classification 1 decision tree and rule based classification**. [Online].เข้าถึงเมื่อวันที่ 5 ก.ย. 2561.https://www.slideshare.net/NontawatB/05-classification-1-decision-tree-and-rule-based-classification.

[3] Simplilearn. **Decision Tree Algorithm With Example | Decision Tree In Machine Learning | Data Science |Simplilearn**. [Online]. เข้าถึงเมื่อวันที่ 7 ก.ย. 2561. https://www.slideshare.net/Simplilearn/decision-tree-algorithm-with-example-decision-tree-in-machine-learning-data-science-simplilearn?qid=659ef5a6-d11e-4860-9acf-997c64667094&v=&b=&from_search=8.

[4] Vithan Minaphinant. **Machine Learning คืออะไร?**. [Online].เข้าถึงเมื่อวันที่ 9 ก.ย. 2561. https://blog.finnomena.com/machine-learning-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-fa8bf6663c07.

[5] Nuttavut Thongjor. **[Machine Learning#1] Machine Learning คืออะไร? รู้จักประเภทต่างๆของ Machine Learning**. [Online].เข้าถึงเมื่อวันที่ 10 ก.ย.2561. https://www.babelcoder.com/blog/posts/machine-learning-introduction

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[6] Shashank Gupta. **Ten Machine Learning Algorithms You Should Know to Become a Data Scientist**. [Online].เข้าถึงเมื่อวันที่ 10 ก.ย.2561. https://towardsdatascience.com/ten-machine-learning-algorithms-you-should-know-to-become-a-data-scientist-8dc93d8ca52e.

[7] Jason Brownlee, PhD. **Your First Machine Learning Project in R Step-By-Step (tutorial and template for future projects)**.[Online].เข้าถึงเมื่อวันที่ 10 ก.ย.2561. https://machinelearningmastery.com/machine-learning-in-r-step-by-step/.

[8] Michail Alifierakis. **Using Yelp Data to Predict Restaurant Closure.**.[Online].เข้าถึงเมื่อวันที่ 10 ก.ย.2561. https://towardsdatascience.com/using-yelp-data-to-predict-restaurant-closure-8aafa4f72ad6

[9] David Cassel. **Can Yelp Data Predict Restaurant Closures?**. [Online].เข้าถึงเมื่อวันที่ 11 ก.ย.2561. https://thenewstack.io/can-yelp-data-predict-restaurant-closures/

[10] Komal Arunjeet Kaur, Dr. Lalita Bhutani. **A Review on Classification Using Decision Tree**. [Online].เข้าถึงเมื่อวันที่ 11 ก.ย.2561. https://pdfs.semanticscholar.org/767a/b3801f9dbf4b03ec891e276c443fbd4029fd.pdf

[11] NYC Data Science Academy. **Predicting Michelin Stars from Yelp Reviews in San Francisco** . [Online].เข้าถึงเมื่อวันที่ 11 ก.ย.2561. https://www.datasciencecentral.com/profiles/blogs/predicting-michelin-stars-from-yelp-reviews-in-san-francisco

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[12] Pang Taithong.**การสร้าง Classification Modelโดยใช้ภาษา R**.. [Online].เข้าถึงเมื่อวันที่ 15 ก.ย.2561.
https://purinko.wordpress.com/2016/03/20/%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%AA%E0%B8%A3%E0%B9%89%E0%B8%B2%E0%B8%87-classification-model%E0%B9%82%E0%B8%94%E0%B8%A2%E0%B9%83%E0%B8%8A%E0%B9%89%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2-r/

[13] Pang Taithong.**การสร้าง Classification Modelโดยใช้ภาษา R ตอนที่2**. . [Online].เข้าถึงเมื่อวันที่ 15 ก.ย.2561.
https://purinko.wordpress.com/2016/03/20/%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%AA%E0%B8%A3%E0%B9%89%E0%B8%B2%E0%B8%87-classification-model%E0%B9%82%E0%B8%94%E0%B8%A2%E0%B9%83%E0%B8%8A%E0%B9%89%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2-r-%E0%B8%95/

[14] Pang Taithong.**คำสั่งภาษา R เบื้องต้น (Decision Tree, Random Forest)**. . [Online].เข้าถึงเมื่อวันที่ 16 ก.ย.2561.
https://purinko.wordpress.com/2016/03/26/%E0%B8%84%E0%B8%B3%E0%B8%AA%E0%B8%B1%E0%B9%88%E0%B8%87%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2-r-%E0%B9%80%E0%B8%9A%E0%B8%B7%E0%B9%89%E0%B8%AD%E0%B8%87%E0%B8%95%E0%B9%89%E0%B8%99-decision-tree-random-fo/

[15] guru99. **Decision Trees in R with Example**. [Online].เข้าถึงเมื่อวันที่ 16 ก.ย.2561.
https://www.guru99.com/r-decision-trees.html

[16] DAVO. **Building a classification tree in R**.[Online].เข้าถึงเมื่อวันที่ 17 ก.ย.2561.
https://davetang.org/muse/2013/03/12/building-a-classification-tree-in-r/

เอกสารนี้สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[17] Madhu Sanjeevi (Mady). **Chapter 4: Decision Trees Algorithms**. [Online].เข้าถึงเมื่อวันที่ 18 ก.ย.2561. https://medium.com/deep-math-machine-learning-ai/chapter-4-decision-trees-algorithms-b93975f7a1f1

[18] ดร. เอกสิทธิ์ พัชรวงศ์ศักดา.**การคัดเลือก feature (feature selection) ด้วยวิธี Information Gain**. [Online].เข้าถึงเมื่อวันที่ 18 ก.ย.2561. http://dataminingtrend.com/2014/data-mining-techniques/feature-selection-information-gain/

[19] Kendal Wong. **Chapter 24: Decision Trees**. [Online].เข้าถึงเมื่อวันที่ 19 ก.ย.2561. https://ademos.people.uic.edu/Chapter24.html

[20] Rahul Saxena . **DECISION TREE CLASSIFIER IMPLEMENTATION IN R**. [Online].เข้าถึงเมื่อวันที่ 19 ก.ย.2561. http://dataaspirant.com/2017/02/03/decision-tree-classifier-implementation-in-r/

[21] Ben Gorman . **Decision Trees in R using rpart**. [Online].เข้าถึงเมื่อวันที่ 20 ก.ย.2561. https://gormanalysis.com/decision-trees-in-r-using-rpart/

[22] Wikibooks. **Data Mining Algorithms In R/Dimensionality Reduction/Feature Selection**. [Online].เข้าถึงเมื่อวันที่ 20 ก.ย.2561. https://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Dimensionality_Reduction/Feature_Selection

[23] KOR'S BLOG. **Machine learning #4 | การนำ Machine learning มาใช้งานต้องทำอะไรบ้าง**.[Online].เข้าถึงเมื่อวันที่ 21 ก.ย.2561. https://khasathan.in.th/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[24] Edwin de Jonge,Mark van der Loo.**An introduction to data cleaning with R** . [Online].เข้าถึงเมื่อวันที่ 21 ก.ย.2561. https://cran.rproject.org/doc/contrib/de_Jonge+van_der_Loo-Introduction_to_data_cleaning_with_R.pdf

[25] Susie Humby**. How to split data into training/testing sets using sample function.** [Online].เข้าถึงเมื่อวันที่ 21 ก.ย.2561. https://stackoverflow.com/questions/17200114/how-to-split-data-into-training-testing-sets-using-sample-function

[26] Max Kuhn . **confusionMatrix** . [Online].เข้าถึงเมื่อวันที่ 22 ก.ย.2561. https://www.rdocumentation.org/packages/caret/versions/3.45/topics/confusionMatrix

[27] KOR'S BLOG. **Machine Learning #5 | ทำความเข้าใจ Entropy กันสักหน่อย** .[Online].เข้าถึงเมื่อวันที่ 22 ก.ย.2561. https://khasathan.in.th/

[28] Michelle. **Can two or more splits in a binary decision tree be made on the same variable?** .[Online].เข้าถึงเมื่อวันที่ 25 ก.ย.2561. https://stats.stackexchange.com/questions/41105/can-two-or-more-splits-in-a-binary-decision-tree-be-made-on-the-same-variable

[29] webfocusinfocenter. **Explanation of the Decision Tree Model**. [Online].เข้าถึงเมื่อวันที่ 26 ก.ย.2561. https://webfocusinfocenter.informationbuilders.com/wfappent/TLs/TL_rstat/source/DecisionTree47.htm

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[30] Max Kuhn, Quinlan Ross. **C5.0 Classification Models**. [Online].เข้าถึงเมื่อวันที่ 27 ก.ย.2561. https://topepo.github.io/C5.0/articles/C5.0.html

[31] cjohnson318. **DECISION TREES IN R USING THE C50 PACKAGE**. [Online].เข้าถึงเมื่อวันที่ 27 ก.ย.2561. http://connor-johnson.com/2014/08/29/decision-trees-in-r-using-the-c50-package/

[32] Alvin Nguyen. **Comparative Study of C5.0 and CART algorithms**. [Online].เข้าถึงเมื่อวันที่ 28 ก.ย.2561. http://mercury.webster.edu/aleshunas/Support%20Materials/C4.5/Nguyen-Presentation%20Data%20mining.pdf

[33] Max Kuhn . **Classification Using C5.0 UseR! 2013** . [Online].เข้าถึงเมื่อวันที่ 30 ก.ย. 2561. https://static1.squarespace.com/static/51156277e4b0b8b2ffe11c00/t/51e7e42ce4b0fd2e32684bca/1374151724529/user_C5.0.pdf

[34] Vladimir Vysotsky . **Classification using Decision Trees in R** . [Online].เข้าถึงเมื่อวันที่ 30 ก.ย.2561. https://en.proft.me/2016/11/9/classification-using-decision-trees-r/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

เอกสารทำสหกิจศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่ ศธ 0524.08/ 4042

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เลขที่ 1 ซอยฉลองกรุง 1 ถนนฉลองกรุง แขวงลาดกระบัง
เขตลาดกระบัง กรุงเทพฯ 10520

16 กรกฎาคม 2561

เรื่อง ขอส่งนักศึกษาเข้าร่วมโครงการสหกิจศึกษา

เรียน คุณวันจักร ปาณะเจริญ (Data Operations Manager)
Wongnai Media Co., Ltd.

ตามที่ Wongnai Media Co., Ltd. ได้ให้ความอนุเคราะห์รับนักศึกษา ภาควิชาคณิตศาสตร์ สาขาวิชาคณิตศาสตร์ประยุกต์ ชั้นปีที่ 4 คณะวิทยาศาสตร์ เข้าร่วมโครงการสหกิจศึกษาที่หน่วยงานของท่าน ในภาคการศึกษาที่ 1 ปีการศึกษา 2561 ในระหว่างวันที่ 6 สิงหาคม 2561 ถึงวันที่ 14 ธันวาคม 2561 นั้น

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง จึงขอส่งตัวนักศึกษาเข้าร่วมโครงการสหกิจศึกษา จำนวน 1 คน ดังมีรายนามต่อไปนี้

นางสาวสิริกร ศิโรรัตน์ณัฐโชค

จึงเรียนมาเพื่อทราบ และขอขอบพระคุณเป็นอย่างยิ่งในการให้ความอนุเคราะห์

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร. อนุรักษ์ โพธิ์เอียม)
รองคณบดี
ปฏิบัติการแทนคณบดี

งานกิจการนักศึกษา (คุณปิยาภรณ์ โพธิโยธะ) ผู้ประสานงาน
โทรศัพท์สายตรง 02-3298417
โทรศัพท์ 02-3298400-8411 ต่อ 271
โทรสาร 02-3298414 หรือ 02-3298412

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รับเอกสาร

ชื่อผู้รับ.....

วันที่.....

แบบคำขอหนังสือ
ขออนุญาตไปนิเทศนักศึกษา สหกิจศึกษา
ระดับปริญญาตรี
งานกิจการนักศึกษา คณะวิทยาศาสตร์

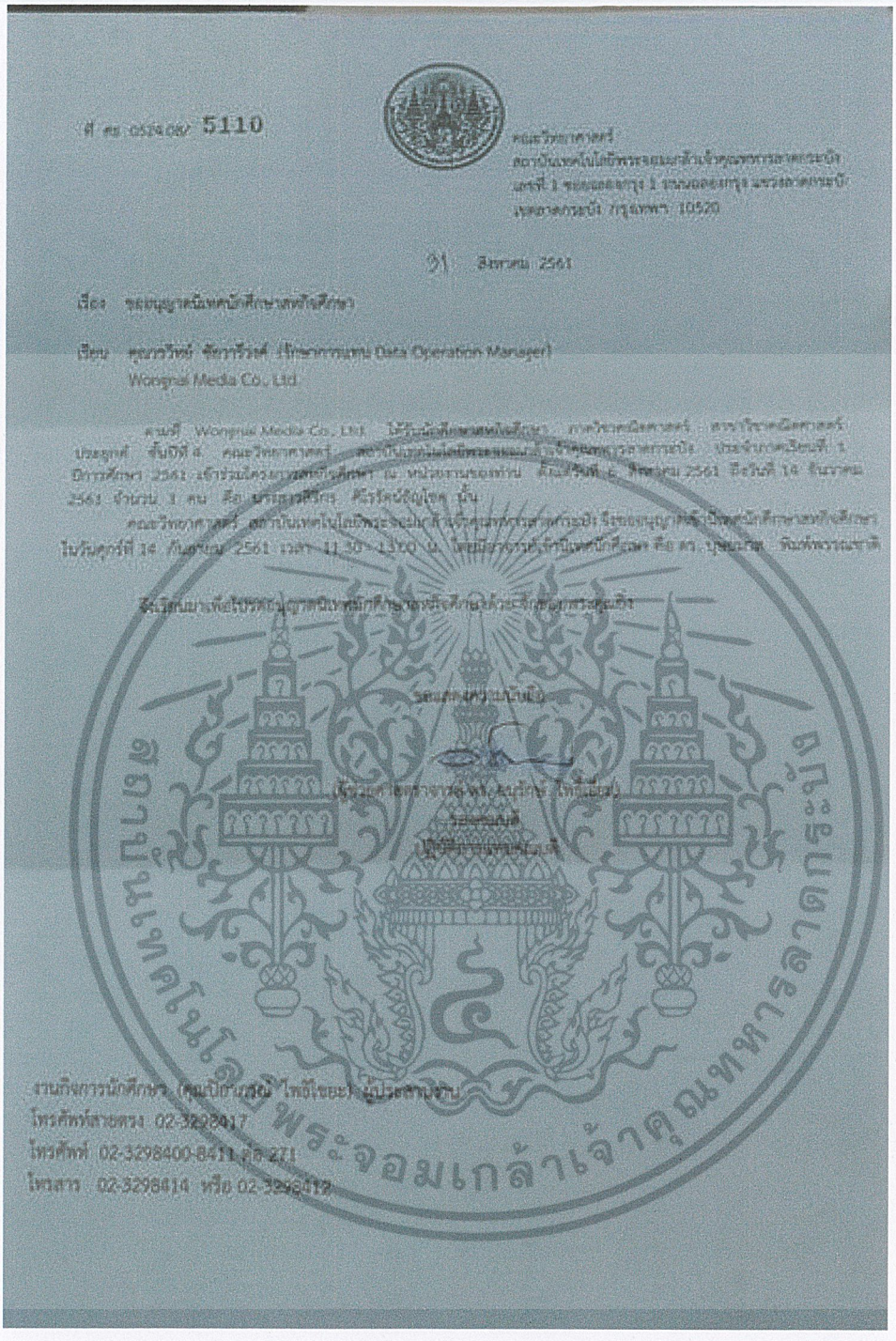
☐ ภาควิชาเคมี ○ ป.ตรี ☐ ภาควิชาชีววิทยา ○ ป.ตรี
☐ ภาควิชาฟิสิกส์ ○ ป.ตรี ☐ ภาควิชาวิทยาการคอมพิวเตอร์ ○ ป.ตรี
☑ ภาควิชาคณิตศาสตร์ ⊘ ป.ตรี ☐ สถิติประยุกต์ ○ ป.ตรี

ชื่อสถานประกอบการ
ที่อยู่สถานประกอบการ
ชื่อหน่วยงาน/บริษัท Sangngai Media Co.,Ltd.
ที่อยู่หน่วยงาน/บริษัท
ชื่อ-นามสกุล
ตำแหน่ง
ความคิดเห็นของหัวหน้าภาควิชา
หมายเหตุ

☐ ส่งเอกสารถึงอาจารย์ที่ปรึกษา ☐ ส่งเอกสารให้นักศึกษาเพื่อนำส่งในภายหลัง ☐ อื่นๆ.....

หมายเหตุ กรุณาส่งแบบคำขอนี้ที่งานกิจการนักศึกษาคณะวิทยาศาสตร์ก่อนที่จะไปนิเทศนักศึกษาอย่างน้อย 1 สัปดาห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ที่ ศธ 0526.08/ 5110

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
เลขที่ 1 ซอยฉลองกรุง 1 แขวงลาดกระบัง เขตลาดกระบัง กรุงเทพฯ 10520

21 สิงหาคม 2561

เรื่อง ขอหนังสือตอบรับนักศึกษาสหกิจศึกษา

เรียน คุณวรวิทย์ ชัยวารีวงศ์ (ในฐานะ Data Operation Manager)
Wongnai Media Co., Ltd.

ตามที่ Wongnai Media Co., Ltd. ได้รับนักศึกษาสหกิจศึกษา ภาควิชาคณิตศาสตร์ สาขาวิชาคณิตศาสตร์ ประยุกต์ ชั้นปีที่ 4 คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ประจำภาคเรียนที่ 1 การศึกษา 2561 เข้าร่วมโครงการสหกิจศึกษา ณ หน่วยงานของท่าน ตั้งแต่วันที่ 6 พฤษภาคม 2561 ถึงวันที่ 14 ตุลาคม 2561 จำนวน 1 คน คือ นายธนวัฒน์ ศิริโชติปัญโญ นั้น

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง จึงขอหนังสือตอบรับนักศึกษาสหกิจศึกษา ในวันศุกร์ที่ 14 กันยายน 2561 เวลา 11.30-13.00 น. โดยมีอาจารย์เจ้าที่เทคนิคนักศึกษา คือ ดร. บุญธนวัฒ พิมพ์ทวนยศ

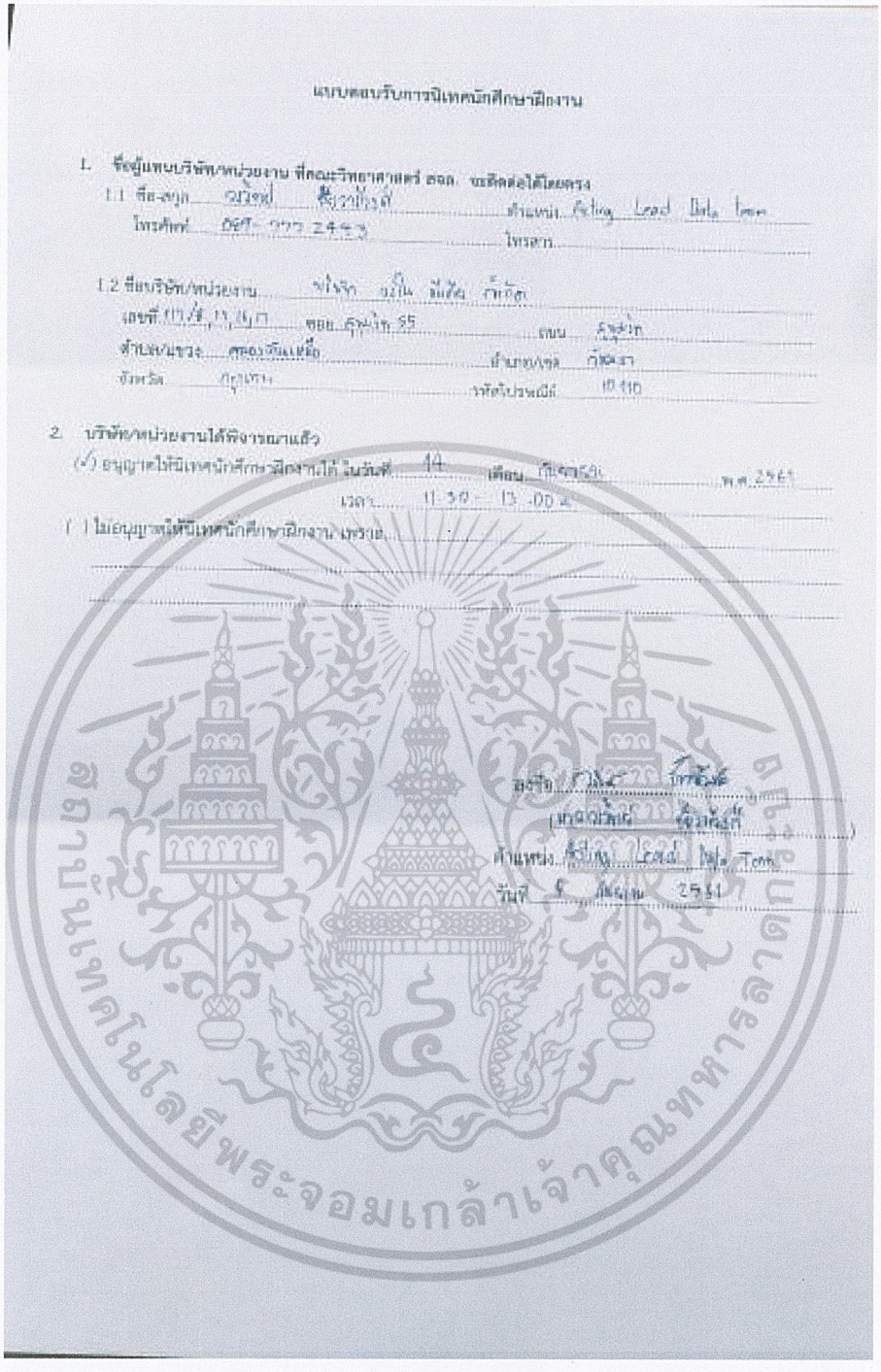
จึงเรียนมาเพื่อโปรดพิจารณาตอบรับนักศึกษาสหกิจศึกษาด้วยจักเป็นพระคุณยิ่ง

รองศาสตราจารย์

(ผู้ช่วยศาสตราจารย์ ดร.ณรงค์ โพธิ์ประสม)
รองคณบดี
ปฏิบัติราชการแทนคณบดี

งานกิจการนักศึกษา (คุณปิยานุช ไพบูลย์ขยะ) ผู้ประสานงาน
โทรศัพท์สายตรง 02-3298417
โทรศัพท์ 02-3298400-8411 ต่อ 271
โทรสาร 02-3298414 หรือ 02-3298412

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



แบบตอบรับการนิเทศนักศึกษาฝึกงาน

1. ชื่อผู้แทนบริษัท/หน่วยงาน ที่คณะวิทยาศาสตร์ มจล. จะติดต่อได้โดยตรง
1.1 ชื่อ-สกุล [illegible] ตำแหน่ง Acting Lead Data Team
โทรศัพท์ 086-[illegible] โทรสาร
1.2 ชื่อบริษัท/หน่วยงาน [illegible]
เลขที่ [illegible] ซอย [illegible] ถนน [illegible]
ตำบล/แขวง [illegible] อำเภอ/เขต [illegible]
จังหวัด [illegible] รหัสไปรษณีย์ 10110

2. บริษัท/หน่วยงานได้พิจารณาแล้ว
(/) อนุญาตให้นิเทศนักศึกษาฝึกงานได้ ในวันที่ 14 เดือน [illegible] พ.ศ. 2564
เวลา 11:30 - 13:00 น.
( ) ไม่อนุญาตให้นิเทศนักศึกษาฝึกงาน เพราะ

ลงชื่อ [illegible]
(นาย[illegible])
ตำแหน่ง Acting Lead Data Team
วันที่ 5 กันยายน 2564

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข
ประวัติอย่างย่อของผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติอย่างย่อของผู้จัดทำ

ชื่อ: นางสาว สิริกร ศิโรรัตน์ธัญโชค รหัสประจำตัวนักศึกษา 58050168

วันเกิด: 7 ธันวาคม 2539

ที่อยู่: 569/13 ถ.แสงชูโต ต.ท่าเรือ อ.ท่ามะกา จ.กาญจนบุรี 71130

เบอร์โทรศัพท์: 087-1699893

E-mail: peppermintsiri@gmail.com

ประวัติการศึกษา: โรงเรียน วิสุทธรังษี จังหวัด กาญจนบุรี

ปริญญาตรี: สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง คณะวิทยาศาสตร์ สาขาคณิตศาสตร์ประยุกต์

ประสบการณ์การทำงาน

- Data Operations Team บริษัท วงใน มีเดีย จำกัด (Internship 2 เดือนและ Cooperative 4 เดือน)

พี่เลี้ยง

นายวรวิทย์ ชัยวารีวงศ์ ตำแหน่ง Acting Team Lead Data Operations Team

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

ภาพกิจกรรมในช่วงสหกิจศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในทุกๆชั้น ของสำนักงานวงใน จะมีมุมเครื่องดื่ม/ขนม กินทั้งวัน ไม่อั้น

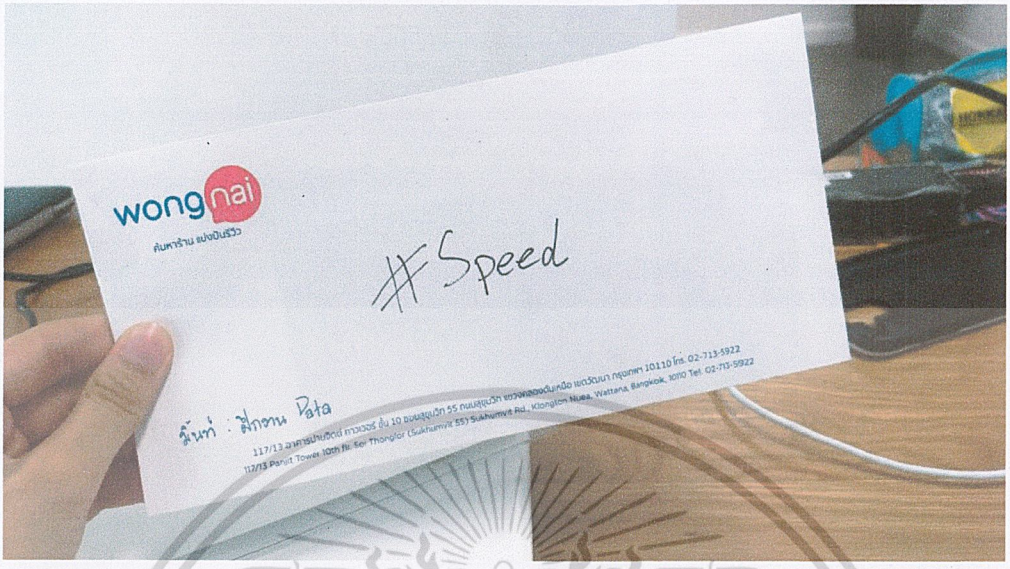
ภาพวัน Wongnai Townhall Q3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

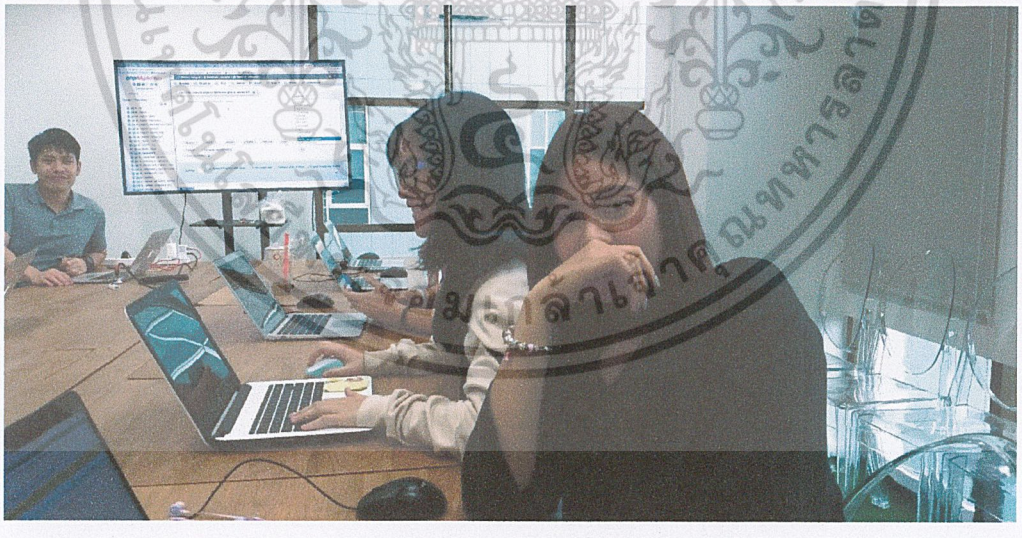
บรรยากาศ Weshare ในตอนที่ P'Woody ที่มา แบ่งปันเล่าประสบการณ์

DATA OPERATIONS TEAM

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

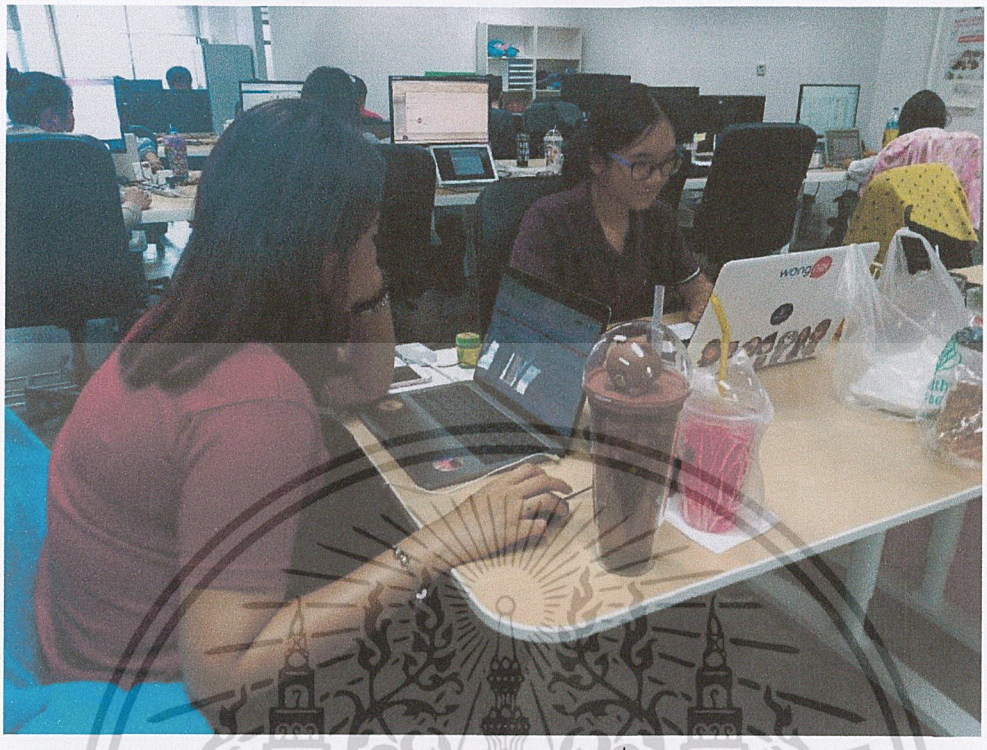


ข้าพเจ้าได้รับ Weekly bonus #speed ซึ่งเป็น 1 ใน 4 core value ของพนักงาน
วงใน ได้รับการเสนอชื่อจาก AE Team

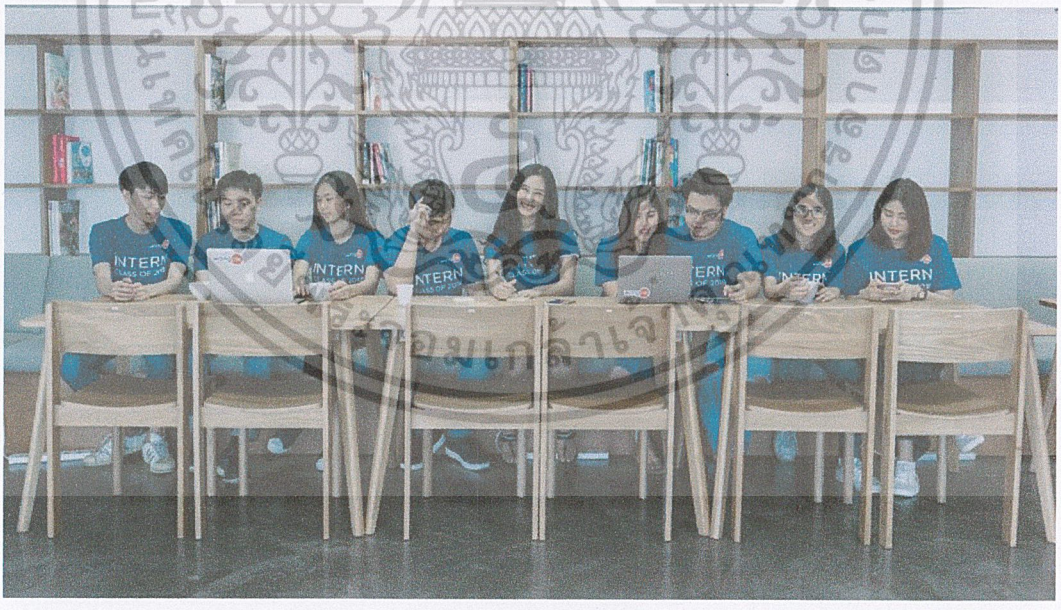


SQL fundamental class by it-support team

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



บรรยากาศวันทำงานทั่วไป



เพื่อนๆ Internship 2018

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ดร.บุษยมาส พิมพ์พรรณชาติ อาจารย์ที่ปรึกษา มานิเทศสหกิจศึกษาที่บริษัท และ

พูดคุยกับ พี่เลี้ยงหัวหน้าทีมที่ดูแลข้าพเจ้าตลอดการสหกิจศึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้