การพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซ

วิทยานิพนธ์

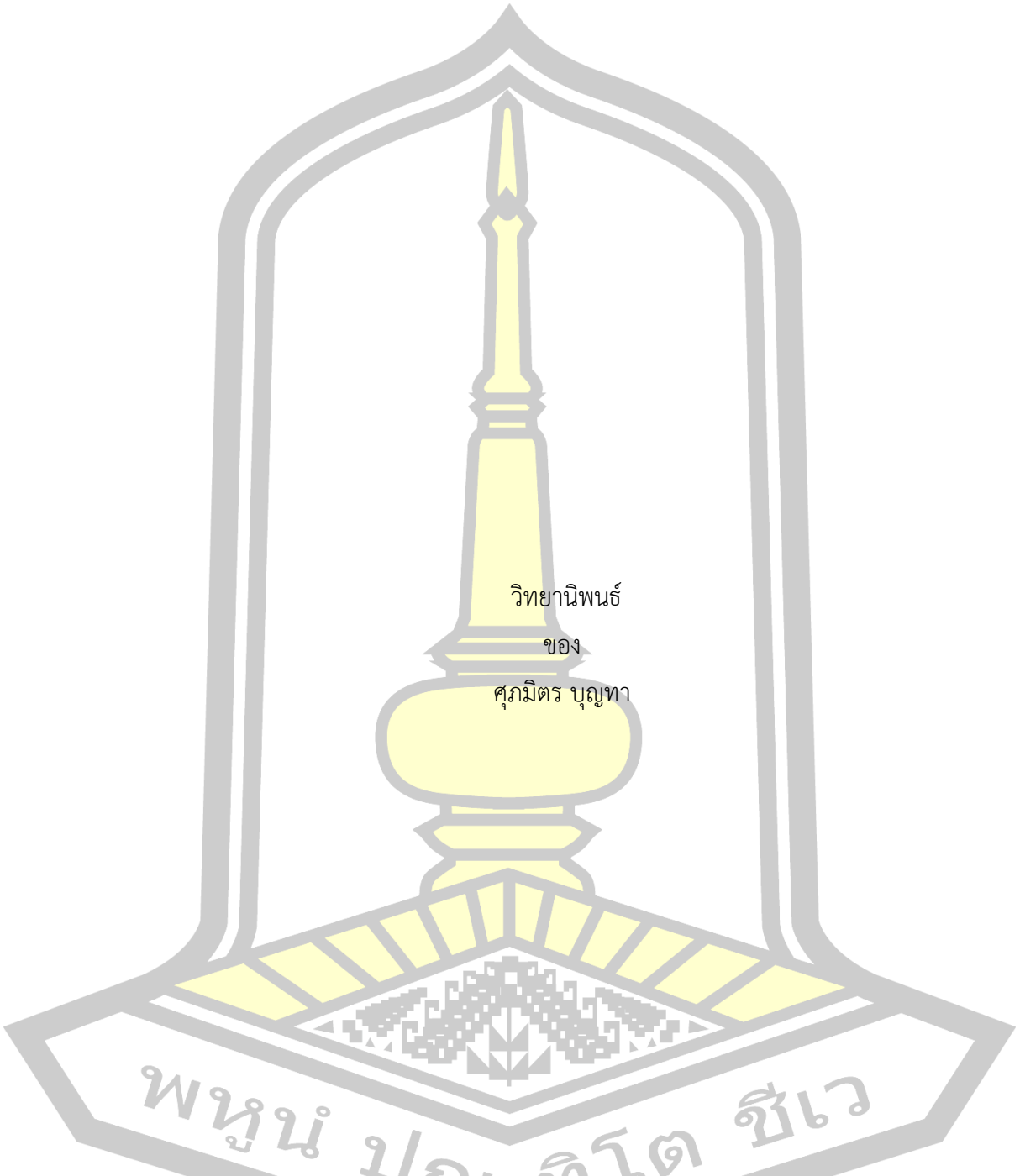
ของ

ศุภมิตร บุญทา

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล
พฤศจิกายน 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซ



วิทยานิพนธ์

ของ

ศุภมิตร บุญทา

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล

พฤศจิกายน 2566

ลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Forecasting Business Performance of Restaurant on Social Commerce

Supamit Boonta

A Thesis Submitted in Partial Fulfillment of Requirements

for Doctor of Philosophy (Business Administration and Digital Innovation)

November 2023

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายศุภมิตร บุญทา แล้ว เห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(รศ. ดร. ธีระวัฒน์ เจริญราษฎร์)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ดร. กาญจนา หินเธาว์)

.....กรรมการ
(รศ. ดร. กิตติพล วิแสง)

.....กรรมการ
(ดร. พงศธร ตันตระบัณฑิตย์)

.....กรรมการ
(ดร. การันต์ กิจระการ)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา ปรัชญาดุษฎีบัณฑิต สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล ของมหาวิทยาลัยมหาสารคาม

.....
(ดร. ชลธิชา ธรรมวิญญู)
คณบดีคณะการบัญชีและการจัดการ

.....
(รศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง การพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซ
ผู้วิจัย ศุภมิตร บุญทา
อาจารย์ที่ปรึกษา ดร. กาญจนา หินเธาว์
ปริญญา ปรัชญาดุษฎีบัณฑิต **สาขาวิชา** บริหารธุรกิจและนวัตกรรมดิจิทัล
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2566

บทคัดย่อ

ธุรกิจร้านอาหารเป็นหนึ่งในธุรกิจที่มีความสำคัญอย่างยิ่งต่อเศรษฐกิจของประเทศไทย เนื่องจากในปี 2565 มีจำนวนร้านอาหาร คิดเป็น 2.14% ของธุรกิจทั้งหมด แต่ขณะเดียวกันร้านอาหารยังเป็นธุรกิจที่มีการเลิกกิจการสูงที่สุดเป็นอันดับ 3 คิดเป็นสัดส่วนราว 2% ของธุรกิจที่เลิกกิจการทั้งหมด ในปัจจุบันโซเชียลคอมเมิร์ซได้เข้ามามีบทบาทสำคัญในการทำธุรกิจร้านอาหาร ดังนั้นการศึกษาปัจจัยที่สัมพันธ์กับความสำเร็จของธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซจึงเป็นสิ่งที่สำคัญเพื่อลดอัตราความล้มเหลวและเพิ่มอัตราความสำเร็จของธุรกิจ

การศึกษาวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ เพื่อเปรียบเทียบประสิทธิภาพอัลกอริธึมในการพยากรณ์ด้วยเทคนิคแมชชีนเลิร์นนิ่ง สำหรับธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ และเพื่อพัฒนาแบบจำลองสำหรับการวิเคราะห์ปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

จากผลการวิจัยพบว่า เทคนิคโครงข่ายประสาทเทียม (Neural Networks) สามารถให้ประสิทธิภาพดีที่สุดในการพยากรณ์ จึงสามารถนำมาใช้ในการสร้างตัวแบบในการพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซได้

ด้านปัจจัยภายใน ระดับราคามีผลต่อระดับความพึงพอใจของลูกค้า โดยเฉพาะกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลางและสูง ความหนาแน่นของประชากรในพื้นเป็นและจำนวนที่นั่งส่งผลต่อกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำและปานกลาง

ด้านการให้บริการอื่นๆ เพื่อเพิ่มความสะดวกสบายแก่ลูกค้า การมีบริการ Wi-Fi เป็นปัจจัยที่สามารถดึงดูดลูกค้าให้เข้ามาใช้บริการในร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ ส่วนร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงควรพิจารณาการรับบัตรเครดิต และการมีบริการเครื่องดื่ม

แอลกอฮอล์

ด้านปัจจัยที่เกี่ยวข้องกับการตลาดผ่านโซเชียลมีเดีย ระดับความคิดเห็นของบทวิจารณ์ออนไลน์ส่งผลมากที่สุดต่อการตัดสินใจใช้บริการของลูกค้าในทุกกลุ่มร้านอาหาร และจำนวนผู้ถูกใจเพจซึ่งเป็นข้อมูลจาก Facebook Fanpage มีผลต่อระดับความพึงพอใจของลูกค้ากลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง ซึ่งแสดงให้เห็นว่าการเลือกใช้ช่องทางแพลตฟอร์มโซเชียลคอมเมิร์ซหลายช่องทางสามารถส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซ

คำสำคัญ : โซเชียลคอมเมิร์ซ, การเรียนรู้ของเครื่อง, การพยากรณ์ประสิทธิภาพของธุรกิจ

พหูน ปณ ฑิโต ชีเว

TITLE Forecasting Business Performance of Restaurant on Social Commerce

AUTHOR Supamit Boonta

ADVISORS Kanjana Hinthaw , Ph.D

DEGREE Doctor of Philosophy **MAJOR** Business Administration and Digital Innovation

UNIVERSITY Mahasarakham University **YEAR** 2023

ABSTRACT

The restaurant business is one of the most important businesses in the Thai economy; in 2022, the number of restaurants accounted for 2.14% of all businesses. However, at the same time, restaurants were also the third-highest go out of business, accounting for about 2% of all go out of business. Presently, social commerce has played an essential role in the restaurant business. Therefore, it is crucial to study the factors related to the success of a restaurant business on social commerce platforms to reduce the failure rate and increase the success rate of the business.

This research study aimed to examine the factors that affect restaurant performance on social commerce platforms and compare the algorithmic performance of forecasting with machine learning techniques. For restaurant businesses on social commerce platforms and to develop a model for analyzing factors affecting restaurant performance on social commerce platforms.

According to the research, found that Neural Networks techniques can provide the best performance in forecasting. Therefore, it can be used to create a model to predict the performance of a restaurant business in social commerce.

Internal factors the price level affects customer satisfaction—especially in the restaurant category with medium and high price ranges. The population density

in the area and the number of seats affects the restaurant category with low and medium price ranges.

Other services to increase convenience for customers include Wi-Fi service, which is a factor that can attract customers to use the service in restaurants with a low price range. Restaurants with a higher price range should consider accepting credit cards and serving alcoholic beverages.

Factors related to social media marketing the level of opinion expressed in online reviews has the most significant influence on customer service decisions across restaurant chains. Furthermore, the number of page likes, according to the data from Facebook Fan page, affects the customer satisfaction level of restaurants with high price ranges. Thus, the level of customer satisfaction shows that choosing to use multiple social commerce platforms can impact the performance of a restaurant business on social commerce.

Keyword : Social commerce, Machine learning, Forecasting business performance

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จสมบูรณ์ได้ด้วยความกรุณาและความช่วยเหลืออย่างสูงยิ่งจาก ดร. กาญจนา หินเธาว์ ที่ปรึกษาวิทยานิพนธ์หลัก ที่เสียสละเวลาอันมีค่าในการให้คำปรึกษา ทั้งในเวลาราชการและนอกเวลาราชการในการให้คำแนะนำ ให้กำลังใจ ติดตาม และผลักดัน จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

ขอขอบพระคุณ รองศาสตราจารย์ ดร. ธีระวัฒน์ เจริญราษฎร์ ประธานกรรมการควบคุมวิทยานิพนธ์ ดร. พงศธร ตันตระบัณฑิตย์ ดร. การันต์ กิจระการ กรรมการสอบวิทยานิพนธ์ และขอขอบพระคุณ รองศาสตราจารย์ ดร. กิตติพล วิแสง ประธานหลักสูตรบริหารธุรกิจและนวัตกรรมดิจิทัลที่กรุณาให้แนวทางในการทำงานวิจัยในการศึกษาระดับดุษฎีบัณฑิตตั้งแต่วันแรกที่เข้ารับการศึกษา ตลอดจนการปรับปรุงงานค้นคว้าวิจัยให้มีคุณภาพ

ขอขอบคุณผู้บริหาร บุคลากร และเพื่อนร่วมงานในสาขาวิชาคอมพิวเตอร์ธุรกิจมหาวิทยาลัยราชภัฏสกลนครที่ช่วยเหลือ ให้โอกาส และทุนสนับสนุนในการศึกษา

สุดท้ายนี้ขอขอบคุณครอบครัวของข้าพเจ้าที่สนับสนุน และเป็นกำลังใจในการศึกษาจนข้าพเจ้าสามารถจัดทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปด้วยดี

ศุภมิตร บุญทา

พหูน ปณฺฑิโต ชีเว

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... ฉ

กิตติกรรมประกาศ ..... ช

สารบัญ ..... ฌ

บัญชีตาราง ..... ฑ

บัญชีภาพประกอบ ..... ท

บทที่ 1 บทนำ 1

ภูมิหลัง 1

ขั้นตอนในการทำวิจัย 5

ความมุ่งหมายของการวิจัย 5

ความสำคัญของการวิจัย 6

ขอบเขตของการวิจัย 6

นิยามศัพท์เฉพาะ 6

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 8

แนวคิดเกี่ยวกับประสิทธิภาพของธุรกิจร้านอาหาร 8

แนวคิดเกี่ยวกับโซเชียลคอมเมิร์ซ 9

แนวคิดเกี่ยวกับการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์ (eWOM Marketing) 10

แนวคิดเกี่ยวกับความเข้มข้นทางอารมณ์ (Emotional intensity) 10

ความตั้งใจซื้อบนแพลตฟอร์มโซเชียลคอมเมิร์ซ 11

การเรียนรู้ของเครื่อง (Machine learning) 12

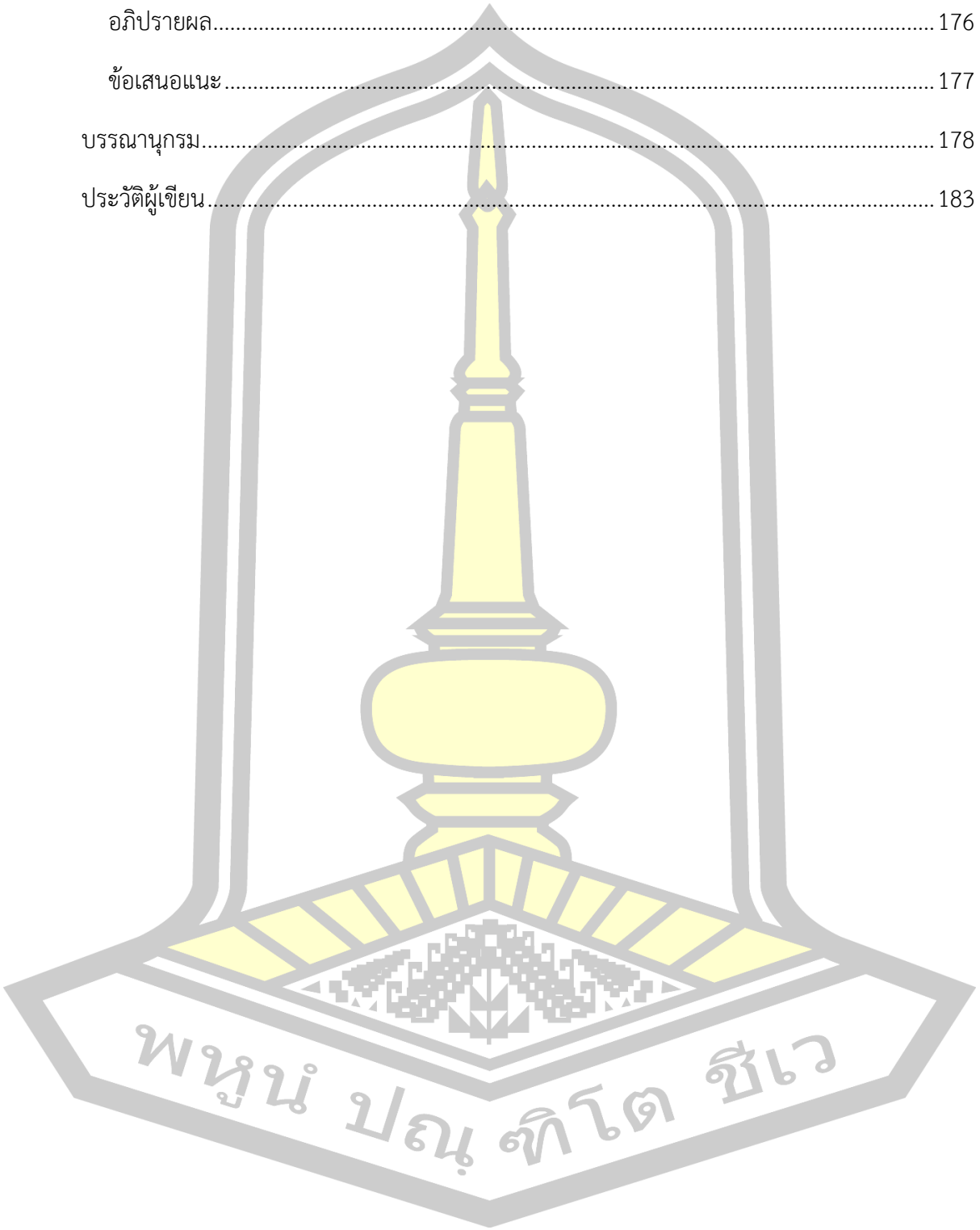
1. การเรียนรู้โดยมีผู้สอน (Supervised Learning) 12

2. การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning) 12
3. การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) 12
ต้นไม้ตัดสินใจ (Decision Tree) 13
โครงข่ายประสาทเทียม (Neural Network) 14
การวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis) 15
ซัพพอร์ตเวกเตอร์แมชชีน (SVM) 15
การคัดเลือกคุณสมบัติ (Feature Selection) 16
1. การคัดเลือกตัวแปรโดยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 16
2. การเลือกตัวแปรโดยวิธีเพิ่มตัวแปร (Forward Selection) 17
3. การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward Elimination) 17
การวัดประสิทธิภาพของแบบจำลอง (Method Evaluation) 17
1. ค่าความถูกต้อง (Accuracy) 18
2. ค่าความระลึก (Class Recall) 18
3. ค่าความแม่นยำ (Class Precision) 19
4. ค่าพื้นที่ใต้กราฟ ROC (Area Under the Curve; AUC) 19
งานวิจัยที่เกี่ยวข้อง 19
บทที่ 3 วิธีดำเนินการวิจัย 28
ประชากรและกลุ่มตัวอย่าง 28
เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล 28
การเก็บรวบรวมข้อมูล 29
การจัดกระทำกับข้อมูลและการวิเคราะห์ข้อมูล 29
การทำความสะอาดข้อมูล (Data Cleaning) 30
การแปลงข้อมูล (Data Transformation) 30
การวิเคราะห์ความคิดเห็น (sentiment analysis) 32

การลดมิติข้อมูล (Dimensionality Reduction) 33
การแบ่งกลุ่มข้อมูล (Clustering) 34
การวิเคราะห์ข้อมูล (Data Analysis) 34
การวัดประสิทธิภาพของแบบจำลอง (Method Evaluation) 34
บทที่ 4 ผลการวิเคราะห์ข้อมูล 36
การนำเข้าข้อมูล 36
การแบ่งกลุ่มข้อมูล (Clustering) 41
การคัดเลือกตัวแปร (Feature Selection) 43
1. ผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) 44
2. ผลการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) 46
3. ผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 49
การสร้างตัวแบบพยากรณ์ 51
1. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) 52
1.1 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) 52
1.2 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) 61
1.3 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 70
1.4 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 79
2. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) 88
2.1 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection) 88

2.2 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) 96
2.3 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและเพิ่มและลดตัวแปร (Optimize selection) 105
2.4 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 112
3. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) 122
3.1 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) 122
3.2 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) 131
3.3 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 143
3.4 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 153
การประเมินประสิทธิภาพตัวแบบพยากรณ์ 167
1. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) 167
2. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 169
3. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 171
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ 174
ความมุ่งหมายของการวิจัย 174
เครื่องมือที่ใช้ในการวิจัย 174
การเก็บรวบรวมข้อมูล 175

การวิเคราะห์ข้อมูล (Data Analysis) 175
อภิปรายผล 176
ข้อเสนอแนะ 177
บรรณานุกรม 178
ประวัติผู้เขียน 183



บัญชีตาราง

หน้า

ตาราง 1 เมทริกซ์ความสับสน (Confusion matrix) 18

ตาราง 2 งานวิจัยที่เกี่ยวข้องด้วยเทคโนโลยีการเรียนรู้ของเครื่อง 26

ตาราง 3 เกณฑ์การแปลงข้อมูล 31

ตาราง 4 ข้อมูลการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering 42

ตาราง 5 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 56

ตาราง 6 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 58

ตาราง 7 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 60

ตาราง 8 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 65

ตาราง 9 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 67

ตาราง 10 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 69

ตาราง 11 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 74

ตาราง 12 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 76

ตาราง 13 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 78

ตาราง 14 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมดในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 82

ตาราง 15 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 85

ตาราง 16 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 87

ตาราง 17 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 91

ตาราง 18 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 91

ตาราง 19 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 92

ตาราง 20 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 93

ตาราง 21 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 94

ตาราง 22 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง..... 95

ตาราง 23 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 99

ตาราง 24 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 99

ตาราง 25 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 101

ตาราง 26 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 102

ตาราง 27 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 103

ตาราง 28 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 104

ตาราง 29 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 107

ตาราง 30 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 108

ตาราง 31 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 109

ตาราง 32 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 109

ตาราง 33 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 111

ตาราง 34 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 111

ตาราง 35 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 115

ตาราง 36 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 116

ตาราง 37 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 117

ตาราง 38 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 118

ตาราง 39 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 119

ตาราง 40 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 120

ตาราง 41 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 125

ตาราง 42 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 126

ตาราง 43 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 127

ตาราง 44 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 128

ตาราง 45 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 130

ตาราง 46 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง..... 130

ตาราง 47 ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 135

ตาราง 48 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 136

ตาราง 49 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 138

ตาราง 50 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 139

ตาราง 51 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 141

ตาราง 52 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 142

ตาราง 53 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 146

ตาราง 54 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 147

ตาราง 55 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 149

ตาราง 56 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 149

ตาราง 57 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 151

ตาราง 58 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 152

ตาราง 59 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 157

ตาราง 60 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 158

ตาราง 61 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 161

ตาราง 62 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 163

ตาราง 63 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 165

ตาราง 64 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 166

ตาราง 65 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) 167

ตาราง 66 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 169

ตาราง 67 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 170

ตาราง 68 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 171

ตาราง 69 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 172

ตาราง 70 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 173

พหูน ปณ ฑิโต ชีเว

บัญชีภาพประกอบ

หน้า

ภาพประกอบ 1 รายงานสถิติและการคาดการณ์ผู้ใช้โซเชียลคอมเมิร์ซ 2020-2026 2

ภาพประกอบ 2 สถิติช่องทางโซเชียลมีเดียที่ผู้บริโภคใช้ในการเข้าถึงเว็บไซต์อื่น 3

ภาพประกอบ 3 เว็บไซต์ Wongnai.com ซึ่งเป็นเว็บไซต์รีวิวร้านอาหาร 4

ภาพประกอบ 4 ขั้นตอนในการทำวิจัย..... 5

ภาพประกอบ 5 โครงสร้างต้นไม้ตัดสินใจ 13

ภาพประกอบ 6 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น 14

ภาพประกอบ 7 หลักการทำงานของเทคนิค Support Vector Machine (SVM) 16

ภาพประกอบ 8 ขั้นตอนการจัดกระทำกับข้อมูลและการวิเคราะห์ข้อมูล 30

ภาพประกอบ 9 เว็บไซต์ Cloud Natural Language ของบริษัท Google 33

ภาพประกอบ 10 ตัวอย่างหน้าจอโปรแกรมสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai 37

ภาพประกอบ 11 ตัวอย่างโค๊ดโปรแกรมสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai 38

ภาพประกอบ 12 ตัวอย่างหน้าจอข้อมูลร้านอาหาร 39

ภาพประกอบ 13 หน้าจอ Natural Language AI..... 40

ภาพประกอบ 14 ตัวอย่างหน้าจอข้อมูลคะแนนรีวิว 40

ภาพประกอบ 15 ขั้นตอนการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering 41

ภาพประกอบ 16 ผลการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering 42

ภาพประกอบ 17 ตัวอย่างข้อมูลที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูลด้วยเทคนิค K-means Clustering..... 43

ภาพประกอบ 18 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) 44

ภาพประกอบ 19 ขั้นตอนการทดสอบประสิทธิภาพภายใน Forward Selection โดยใช้ Cross Validation 45

ภาพประกอบ 20 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือดตัวแปรด้วยวิธี Forward Selection 45

ภาพประกอบ 21 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) 46

ภาพประกอบ 22 ขั้นตอนการทดสอบประสิทธิภาพภายใน Backward Elimination โดยใช้ Cross Validation 47

ภาพประกอบ 23 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือดตัวแปรด้วยวิธี Backward Elimination 48

ภาพประกอบ 24 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 49

ภาพประกอบ 25 ขั้นตอนการทดสอบประสิทธิภาพภายใน Optimize Selection โดยใช้ Cross Validation 50

ภาพประกอบ 26 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือดตัวแปรด้วยวิธี Optimize Selection 50

ภาพประกอบ 27 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) 52

ภาพประกอบ 28 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection) 53

ภาพประกอบ 29 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีเพิ่มตัวแปร (forward selection) 54

ภาพประกอบ 30 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 55

ภาพประกอบ 31 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ ... 56

ภาพประกอบ 32 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 57

ภาพประกอบ 33 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 58

ภาพประกอบ 34 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 59

ภาพประกอบ 35 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 60

ภาพประกอบ 36 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) 61

ภาพประกอบ 37 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination) 62

ภาพประกอบ 38 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีลดตัวแปร(Backward Elimination) 63

ภาพประกอบ 39 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 64

ภาพประกอบ 40 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 65

ภาพประกอบ 41 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 66

ภาพประกอบ 42 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 67

ภาพประกอบ 43 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 68

ภาพประกอบ 44 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 69

ภาพประกอบ 45 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 70

ภาพประกอบ 46 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)..... 71

ภาพประกอบ 47 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)..... 72

ภาพประกอบ 48 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 73

ภาพประกอบ 49 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 74

ภาพประกอบ 50 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 75

ภาพประกอบ 51 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 76

ภาพประกอบ 52 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง..... 77

ภาพประกอบ 53 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง..... 78

ภาพประกอบ 54 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)..... 79

ภาพประกอบ 55 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)..... 80

ภาพประกอบ 56 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)..... 81

ภาพประกอบ 57 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 82

ภาพประกอบ 58 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมดในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 83

ภาพประกอบ 59 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 84

ภาพประกอบ 60 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 85

ภาพประกอบ 61 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 86

ภาพประกอบ 62 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 87

ภาพประกอบ 63 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection) 88

ภาพประกอบ 64 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection) 89

ภาพประกอบ 65 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีเพิ่มตัวแปร (forward selection) 90

ภาพประกอบ 66 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 92

ภาพประกอบ 67 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 94

ภาพประกอบ 68 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 95

ภาพประกอบ 69 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination)..... 96

ภาพประกอบ 70 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination)..... 97

ภาพประกอบ 71 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีลดตัวแปร (Backward Elimination)..... 98

ภาพประกอบ 72 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 100

ภาพประกอบ 73 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 102

ภาพประกอบ 74 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง..... 104

ภาพประกอบ 75 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)..... 105

ภาพประกอบ 76 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)..... 106

ภาพประกอบ 77 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)..... 106

ภาพประกอบ 78 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ..... 108

ภาพประกอบ 79 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง..... 110

ภาพประกอบ 80 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 112

ภาพประกอบ 81 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 113

ภาพประกอบ 82 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 114

ภาพประกอบ 83 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 114

ภาพประกอบ 84 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 117

ภาพประกอบ 85 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 119

ภาพประกอบ 86 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 121

ภาพประกอบ 87 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) 122

ภาพประกอบ 88 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (Forward selection) 123

ภาพประกอบ 89 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีเพิ่มตัวแปร (Forward selection) 124

ภาพประกอบ 90 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 125

ภาพประกอบ 91 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 126

ภาพประกอบ 92 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 127

ภาพประกอบ 93 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 129

ภาพประกอบ 94 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 129

ภาพประกอบ 95 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 131

ภาพประกอบ 96 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) 132

ภาพประกอบ 97 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination) 133

ภาพประกอบ 98 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีลดตัวแปร (Backward Elimination) 133

ภาพประกอบ 99 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 134

ภาพประกอบ 100 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 136

ภาพประกอบ 101 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 137

ภาพประกอบ 102 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 139

ภาพประกอบ 103 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 140

ภาพประกอบ 104 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 142

ภาพประกอบ 105 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 143

ภาพประกอบ 106 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 144

ภาพประกอบ 107 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) 145

ภาพประกอบ 108 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 146

ภาพประกอบ 109 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 148

ภาพประกอบ 110 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 148

ภาพประกอบ 111 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 150

ภาพประกอบ 112 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 151

ภาพประกอบ 113 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 152

ภาพประกอบ 114 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 153

ภาพประกอบ 115 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 154

ภาพประกอบ 116 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) 155

ภาพประกอบ 117 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 156

ภาพประกอบ 118 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 159

ภาพประกอบ 119 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 160

ภาพประกอบ 120 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 163

ภาพประกอบ 121 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 164

ภาพประกอบ 122 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 167

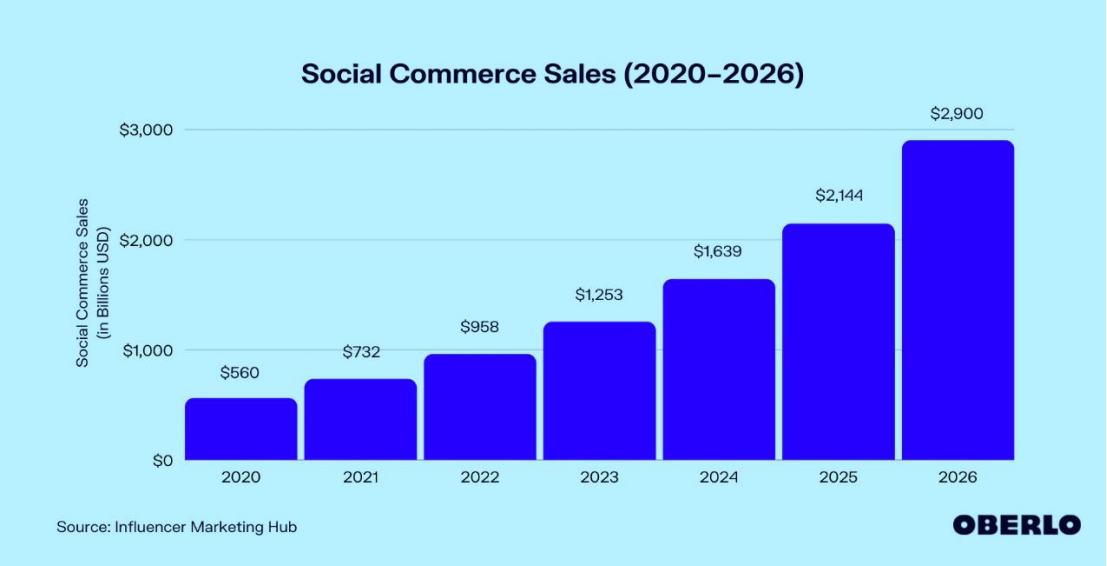
บทที่ 1
บทนำ

ภูมิหลัง

ธุรกิจร้านอาหารเป็นหนึ่งในธุรกิจที่มีความสำคัญอย่างยิ่งต่อเศรษฐกิจของประเทศ แต่ในขณะเดียวกันก็เป็นธุรกิจที่มีการแข่งขันสูงและมีความไม่แน่นอนอย่างมาก แม้ว่าต้องเผชิญกับความผันผวนอย่างรุนแรงในช่วงปีที่ผ่านมา เนื่องจากเศรฐกิจโลกที่ซบเซาและสถานการณ์การแพร่ระบาดของไวรัส Covid-19 เมื่อสถานการณ์เริ่มคลี่คลายธุรกิจร้านอาหารจึงเป็นธุรกิจที่เติบโตขึ้นอย่างต่อเนื่องจากการขยายตัวจากปัจจัยโครงสร้างประชากรที่เปลี่ยนแปลงไปทั้งจากขนาดครัวเรือนที่เล็กลง การขยายตัวของเมืองและไลฟ์สไตล์ผู้บริโภคที่เปลี่ยนแปลงไป ประกอบกับพฤติกรรมผู้บริโภคยุคใหม่ที่ใช้ชีวิตด้วยความเร่งรีบ มีข้อจำกัดด้านเวลา ต้องการความสะดวกสบาย รวมถึงต้องการหลีกเลี่ยงปัญหาจราจร ส่งผลให้มีความต้องการบริการอาหารดิลิเวอรีเพิ่มมากขึ้น ธุรกิจร้านอาหารที่ดำเนินกิจการอยู่ ณ 30 กันยายน 2565 มีจำนวน 18,096 ราย คิดเป็น 2.14% ของธุรกิจทั้งหมด ส่วนใหญ่ตั้งอยู่ในกรุงเทพมหานครจำนวน 7,516 ราย แต่ขณะเดียวกัน ร้านอาหารยังเป็นธุรกิจที่มีการเลิกกิจการสูงที่สุดเป็นอันดับ 3 คิดเป็นสัดส่วนราว 2% ของธุรกิจที่เลิกกิจการทั้งหมด (กรมพัฒนาธุรกิจการค้า, 2565)

ปัจจัยที่นำไปสู่ล้มเหลวหรือความสำเร็จของร้านอาหารสามารถแบ่งออกได้เป็นปัจจัยมหาภาคและจุลภาค ปัจจัยมหภาค ได้แก่ สภาพเศรษฐกิจปัจจุบัน การออกกฎหมาย สภาพอากาศและเหตุการณ์ทางธรรมชาติ การวางผังเมือง ปัจจัยทางวัฒนธรรมและการแข่งขันใหม่ ๆ ปัจจัยย่อยประกอบด้วยลักษณะของร้านอาหาร ได้แก่ ขนาด ชื่อ ประเภท คุณภาพอาหาร บริการสถานที่การออกแบบ การจัดวาง และคุณสมบัติของเจ้าของร้าน รวมถึงความสามารถและประสบการณ์ในการเป็นผู้ประกอบการและการบริหารจัดการ การสร้างแบรนด์ และเงินลงทุน (Parsa, 2010) ร้านอาหารประเภทเดียวกันที่อยู่ใกล้เคียงกันยังมีผลที่ชัดเจนต่อธุรกิจ ตำแหน่งที่ตั้งทางภูมิศาสตร์ที่ดีสามารถลดต้นทุนและดึงดูดผู้บริโภคได้มากขึ้น ดังนั้นการเลือกตำแหน่งที่ตั้งทางภูมิศาสตร์จึงถือเป็นการตัดสินใจเชิงกลยุทธ์ที่สำคัญ (Y. Wang, 2021) นอกจากนี้การทำการตลาดโดยการประชาสัมพันธ์สินค้าหรือบริการอันเป็นส่วนหนึ่งของการส่งเสริมการขาย (Promotion) และช่องทางขาย (Place) ซึ่งนับเป็น 2 P ใน 4 P ของการตลาดผ่านโซเชียลมีเดียที่เรียกว่าโซเชียลคอมเมิร์ซ ซึ่งได้เข้ามามีบทบาทสำคัญต่อการตัดสินใจเลือกใช้สินค้าหรือบริการของผู้บริโภคในยุคปัจจุบัน (Mehrbakhsh Nilashi, 2021) โดยการเผยแพร่ข้อมูลในรูปแบบของการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์ (eWOM

Marketing) ซึ่งผู้บริโภคสามารถใช้สื่อโซเชียลมีเดียต่างๆ เช่น Facebook Line Instagram หรือ twitter ซึ่งสามารถแบ่งปันความคิดเห็นเกี่ยวกับคุณภาพและบริการต่างๆ ของร้านอาหารได้มากขึ้น



ที่มา : https://www.oberlo.com/statistics/social-commerce-market-size

ภาพประกอบ 1 รายงานสถิติและการคาดการณ์ผู้ใช้โซเชียลคอมเมิร์ซ 2020-2026

จากภาพประกอบ 1 ซึ่งแสดงสถิติขนาดของตลาดโซเชียลคอมเมิร์ซตั้งแต่ปี 2020-2023 และการคาดการณ์จนถึงปี 2026 พบว่าในปี 2023 มีมูลค่าในตลาดโลกรวม 1,253 ล้านดอลลาร์สหรัฐ ซึ่งสูงกว่าปีก่อนหน้าถึง 30.8% ซึ่งนับว่าเป็นการเติบโตอย่างรวดเร็วและต่อเนื่อง โดยคาดการณ์ว่าจะมีมูลค่าประมาณ 2,900 ล้านดอลลาร์สหรัฐภายในปี 2026 (oberlo.com, 2023)

วิจัยกรุงศรีซึ่งเป็นหน่วยงานภายใต้ธนาคารกรุงศรีอยุธยา จำกัด (มหาชน) ได้สำรวจความคิดเห็นของผู้บริโภคทั่วไปในประเทศไทยจำนวน 522 คน ผ่านช่องทางออนไลน์ระหว่างวันที่ 29 มีนาคม – 1 พฤษภาคม 2564 เพื่อให้เข้าใจพฤติกรรมของผู้บริโภคที่ซื้อสินค้าทางโซเชียลคอมเมิร์ซ ผู้บริโภคส่วนใหญ่มักซื้อสินค้าผ่าน Facebook เป็นหลัก (กรุงศรีวิจัย, 2021)

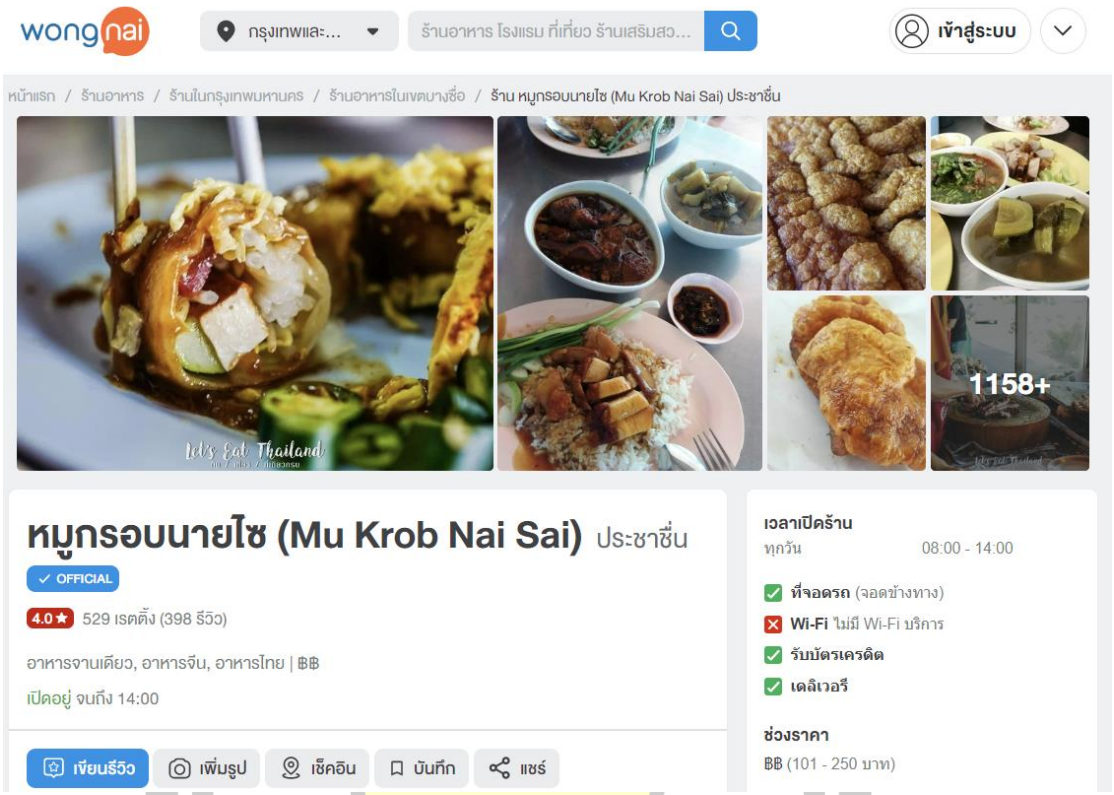
ที่มา : https://datareportal.com/reports/digital-2023-thailand

ภาพประกอบ 2 สถิติช่องทางโซเชียลมีเดียที่ผู้บริโภคใช้ในการเข้าถึงเว็บไซต์อื่น

เว็บไซต์ wearesocial.com ซึ่งเป็นแพลตฟอร์มด้านการบริหารจัดการสื่อสังคมออนไลน์ที่ทำการสำรวจข้อมูลและพฤติกรรมการใช้งานสื่อดิจิทัลของผู้คนทั่วโลกได้รายงานว่าในปี 2023 Facebook คือช่องทางโซเชียลมีเดียที่ผู้บริโภคใช้ในการเข้าถึงเว็บไซต์อื่นมากที่สุดถึง 67.13% รองลงมาคือ Twitter 10.38% และ Instragram 9.65% ซึ่งแสดงให้เห็นว่า Facebook สามารถเป็นช่องทางในการเข้าถึงและประชาสัมพันธ์เว็บไซต์หรือแพลตฟอร์มอื่นได้อย่างมีประสิทธิภาพ

เว็บไซต์ wongnai.com ซึ่งเป็นเว็บไซต์ด้านการรีวิวร้านอาหารซึ่งได้รับความนิยมอย่างมากในประเทศไทย ซึ่งให้ข้อมูลเกี่ยวกับ เมนูอาหาร ตำแหน่งที่ตั้ง ประเภทร้านอาหาร ช่วงราคา จำนวนที่นั่ง เวลาเปิด-ปิด ของร้าน สิ่งอำนวยความสะดวก เช่น ที่จอดรถ การรับบัตรเครดิต บริการดิลิเวอรี เป็นต้น และผู้ใช้ยังสามารถเข้าถึง Facebook Fanpage และโซเชียลมีเดียอื่น ๆ เช่น Line Instagram หรือเว็บไซต์ของร้านอาหารผ่านเว็บไซต์ wongnai.com ได้อีกด้วย รวมถึงสามารถแบ่งปันข้อมูลการใช้บริการโดยการเช็คอิน การให้คะแนนรีวิว การเขียนบทวิจารณ์ออนไลน์ซึ่งผู้ใช้สามารถแสดงออกถึงความคิดเห็นและความรู้สึกที่มีต่อการใช้บริการร้านอาหารทั้งในเชิงบวกและเชิงลบ โดยข้อมูลต่างๆ เหล่านี้ผู้บริโภคสามารถใช้เป็นข้อมูลอ้างอิงในการตัดสินใจใช้บริการตามแนวคิดการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์ (eWOM Marketing) ซึ่งมีผลอย่างมากต่อธุรกิจทั้งออนไลน์และออฟไลน์ โดยธุรกิจสามารถใช้ประโยชน์จากบทวิจารณ์ออนไลน์เหล่านี้ซึ่งเป็นข้อมูลที่ไม่มีโครงสร้างโดยอาศัยการวิเคราะห์ด้วยโมเดลการประมวลผลภาษาธรรมชาติ (Natural language

Processing) และเหมืองข้อความ (Text Mining) เพื่อให้ได้ข้อมูลเชิงลึก เช่น การคาดการณ์ความเชื่อมั่นของผู้บริโภคเพื่อนำไปใช้ในการตัดสินใจเชิงธุรกิจ



ที่มา : https://www.wongnai.com/

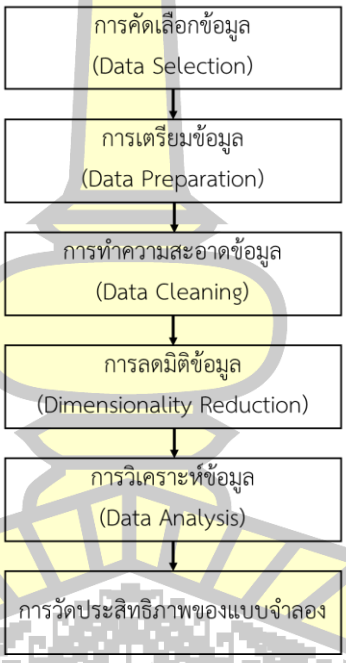
ภาพประกอบ 3 เว็บไซต์ Wongnai.com ซึ่งเป็นเว็บไซต์รีวิวร้านอาหาร

จากการเปลี่ยนแปลงอย่างรวดเร็วของเทคโนโลยีในปัจจุบันทำให้เกิดปัจจัยใหม่ที่มีความสำคัญในการตัดสินใจดำเนินธุรกิจร้านอาหาร เช่น บริการดิลิเวอรี หรือความหลากหลายของช่องทางในการประชาสัมพันธ์ผ่านสื่อโซเชียลมีเดียต่างๆ ทำให้ธุรกิจร้านอาหารจำเป็นต้องปรับตัวเพื่อให้สามารถแข่งขันได้ ดังนั้นผู้วิจัยจึงมีความสนใจศึกษาถึงปัจจัยที่ส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ ซึ่งประสิทธิภาพของธุรกิจร้านอาหารเป็นผลมาจากความพึงพอใจของผู้บริโภค สำหรับโซเชียลคอมเมิร์ซสามารถระบุความพึงพอใจของผู้บริโภคได้ในแง่ของการให้คะแนนวิจารณ์ (Susskind, 2019) และการเพิ่มช่องทางในการประชาสัมพันธ์ผ่านสื่อโซเชียลมีเดียโดยการใช้ Facebook Fanpage ซึ่งเป็นโซเชียลมีเดียที่ได้รับความนิยมมากที่สุดในประเทศไทย จะส่งผลหรือมีปัจจัยใดที่มีความสัมพันธ์เชื่อมโยงกับประสิทธิภาพของธุรกิจร้านอาหารบนเว็บไซต์

wongnai.com ซึ่งเป็นเว็บไซต์ด้านการรีวิวร้านอาหารซึ่งได้รับความนิยมอย่างมากในประเทศไทยหรือไม่ เพื่อให้ได้ข้อมูลเชิงลึกสำหรับใช้ในการตัดสินใจเชิงกลยุทธ์ในการดำเนินธุรกิจร้านอาหารในปัจจุบัน

ขั้นตอนในการทำวิจัย

งานวิจัยนี้มุ่งเน้นการหาปัจจัยที่ส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารซึ่งเป็นผลมาจากจากความพึงพอใจของผู้บริโภคโดยสามารถวัดได้จากคะแนนวิจารณ์ออนไลน์บนโซเชียลคอมเมิร์ซ และการเปรียบเทียบประสิทธิภาพความแม่นยำของตัวแบบการพยากรณ์ (Prediction Model) ที่สร้างขึ้นโดยใช้อัลกอริธึมแมชชีนเลิร์นนิ่ง โดยมีขั้นตอนในการดำเนินการวิจัยดังภาพประกอบ 4



ภาพประกอบ 4 ขั้นตอนในการทำวิจัย

ความมุ่งหมายของการวิจัย

1. เพื่อศึกษาปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

2. เพื่อเปรียบเทียบประสิทธิภาพอัลกอริธึมในการพยากรณ์ด้วยเทคนิคแมชชีนเลิร์นนิ่งสำหรับธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

3. เพื่อพัฒนาแบบจำลองสำหรับการวิเคราะห์ปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

ความสำคัญของการวิจัย

1. ทำให้ทราบถึงปัจจัยที่มีผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

2. ได้แบบจำลองที่เหมาะสมสำหรับการพยากรณ์ประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

3. เพื่อเป็นแนวทางในการตัดสินใจสำหรับผู้ประกอบธุรกิจร้านอาหารโดยใช้แพลตฟอร์มโซเชียลคอมเมิร์ซ

ขอบเขตของการวิจัย

การศึกษาและการวิจัยมุ่งเน้นการศึกษาปัจจัยที่มีผลต่อประสิทธิภาพของธุรกิจร้านอาหารโดยใช้เทคนิคเหมืองข้อมูล บนแพลตฟอร์มโซเชียลคอมเมิร์ซที่ได้จาก wongnai ซึ่งเป็นเว็บไซต์บริการฐานข้อมูลร้านอาหาร โดยเลือกร้านอาหารที่อยู่ในเขตกรุงเทพมหานครและเป็นร้านอาหารที่มี facebook fanpage เพื่อประชาสัมพันธ์ข้อมูล และให้ผู้บริโภคได้แสดงความคิดเห็นหรือเช็คอิน

นิยามศัพท์เฉพาะ

การพยากรณ์ หมายถึง การศึกษาข้อมูลและรูปแบบต่าง ๆ ในอดีต เพื่อคาดการณ์แนวโน้มในอนาคต เพื่อช่วยในการตัดสินใจ

โซเชียลคอมเมิร์ซ หมายถึง การผสมผสานการช็อปปิ้งออนไลน์แบบดั้งเดิมและการตลาดแบบปากต่อปากซึ่งเป็นกระบวนทัศน์ใหม่ของการทำการค้าโดยใช้โซเชียลมีเดียเพื่อเข้าถึงผู้บริโภคในเครือข่าย

ประสิทธิภาพ หมายถึง ความพึงพอใจของผู้บริโภคสำหรับโซเชียลคอมเมิร์ซซึ่งสามารถระบุได้ในแง่ของการให้คะแนนวิจารณ์ออนไลน์

แมชชีนเลิร์นนิ่ง หมายถึง การนำข้อมูลต่าง ๆ มาฝึกให้เครื่องจักรหรือคอมพิวเตอร์ทำการเรียนรู้จากชุดข้อมูลเรียนรู้ (Training Data set) เพื่อหาคำตอบ จากนั้นจึงนำวิธีการหาคำตอบที่ได้ไปใช้ในการพยากรณ์ผลของข้อมูลชุดใหม่

ร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ หมายถึง ร้านอาหารในเขตพื้นที่กรุงเทพมหานครที่มีข้อมูลอยู่บนเว็บไซต์ wongnai.com และมี facebook fanpage เพื่อเป็นช่องทางการติดต่อและการประชาสัมพันธ์

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิจัยนี้มีวัตถุประสงค์เพื่อค้นหาองค์ความรู้โดยการใช้เทคนิคการทำเหมืองข้อมูลเพื่อค้นหาปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ เพื่อให้ได้ข้อมูลเชิงลึกสำหรับใช้ในการตัดสินใจเชิงกลยุทธ์ในการดำเนินงานของธุรกิจร้านอาหาร ดังนั้น ผู้วิจัยจึงได้ทำการรวบรวมเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

1. แนวคิดเกี่ยวกับประสิทธิภาพของร้านอาหาร
2. แนวคิดเกี่ยวกับโซเชียลคอมเมิร์ซ
3. แนวคิดเกี่ยวกับการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์
4. แนวคิดเกี่ยวกับความเข้มข้นทางอารมณ์
5. ความตั้งใจซื้อบนแพลตฟอร์มโซเชียลคอมเมิร์ซ
6. การเรียนรู้ของเครื่อง
7. โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น
8. การวิเคราะห์ถดถอยเชิงเส้น
9. ซัพพอร์ตเวกเตอร์แมชชีน
10. การคัดเลือกคุณสมบัติ
11. การวัดประสิทธิภาพของแบบจำลอง
12. งานวิจัยที่เกี่ยวข้อง

แนวคิดเกี่ยวกับประสิทธิภาพของธุรกิจร้านอาหาร

ประสิทธิภาพของธุรกิจร้านอาหารเป็นผลมาจากความพึงพอใจของผู้บริโภค ซึ่งส่งผลต่อความตั้งใจในการกลับมาใช้บริการซ้ำ และนำไปสู่ผลกำไรของธุรกิจ ปัจจัยที่มีผลต่อความพึงพอใจของผู้บริโภคมีหลายปัจจัย ได้แก่ การต้อนรับของพนักงาน บรรยากาศภายในร้าน ความรวดเร็วในการให้บริการ ความสะดวกสบาย คุณภาพของการบริการ ความหลากหลายของเมนู ราคา คุณภาพของอาหาร โดยความสำคัญของแต่ละปัจจัยขึ้นอยู่กับประเภทของร้านอาหารและประเภทของผู้บริโภค เช่น คุณภาพของอาหารคือปัจจัยที่สำคัญที่สุดสำหรับร้านอาหารแบบ Full-Service ในขณะที่ความรวดเร็วในการให้บริการคือปัจจัยที่สำคัญที่สุดสำหรับร้านอาหารแบบ Quick-Service (Susskind,

2019) ความรู้ความสามารถของเจ้าของกิจการ และกิจกรรมด้านนวัตกรรม เป็นส่วนสำคัญในการเพิ่มประสิทธิภาพของธุรกิจร้านอาหาร รวมถึงกิจกรรมที่เกี่ยวข้องกับผลิตภัณฑ์ การบริการ การจัดการและการตลาด (Lee, 2018) ตำแหน่งที่ตั้งและความหลากหลายทางภูมิศาสตร์เป็นปัจจัยแทรกแซงที่สำคัญในการดำเนินธุรกิจร้านอาหารซึ่งส่งผลต่อประสิทธิภาพทางการเงินและการบริหารจัดการเนื่องจากทรัพยากรหรือปัญหาเฉพาะถิ่น เมื่อเข้าสู่ตลาดทางภูมิศาสตร์ใหม่ผู้ประกอบธุรกิจร้านอาหารจึงควรศึกษาความรู้และข้อมูลที่เกี่ยวข้องเพื่อให้สามารถบริหารจัดการกับปัญหาที่เกิดขึ้นในพื้นที่ทางภูมิศาสตร์ที่แตกต่างกันได้อย่างมีประสิทธิภาพ (Song, 2021) การพิจารณาทำแลที่ตั้งของร้านอาหารเป็นสิ่งสำคัญโดยเฉพาะอย่างยิ่งร้านอาหารที่ตั้งอยู่ในเขตเมืองใหญ่ควรพิจารณาสถานที่ที่มีถนนหรือรถไฟใต้ดินเข้าถึงได้และมีสิ่งอำนวยความสะดวกเชิงพาณิชย์โดยรอบและจำนวนร้านอาหารที่อยู่ในบริเวณใกล้กัน (Wu, 2021) นอกจากความพึงพอใจของผู้บริโภคแล้วประสิทธิภาพยังเกิดจากความสามารถในการแข่งขัน ซึ่งเป็นผลจากการบริหารจัดการเพื่อให้เกิดความยั่งยืนในธุรกิจร้านอาหาร ได้แก่ แนวปฏิบัติทางสังคมและเศรษฐกิจต่อผู้มีส่วนได้ส่วนเสีย เช่น พนักงาน ผู้บริโภค ซัพพลายเออร์ และชุมชน (Cantele, 2020)

แนวคิดเกี่ยวกับโซเชียลคอมเมิร์ซ

โซเชียลคอมเมิร์ซหมายถึงการรวมกันระหว่าง e-commerce และ e-marketing โดยมีพื้นฐานการทำงานอยู่บนเทคโนโลยี Web 2.0 ที่ผู้ใช้เป็นผู้สร้างเนื้อหาทำให้กลุ่มผู้ใช้งานสามารถมีปฏิสัมพันธ์กันในลักษณะของสื่อสังคมออนไลน์ ซึ่งมีทฤษฎีต่างๆ ที่เกี่ยวข้องและถูกนำมาใช้ ได้แก่ ทุนทางสังคม (social capital) จิตวิทยาสังคม (Social Psychology) พฤติกรรมผู้บริโภค (Consumer Behavior) และ การทำงานร่วมกันแบบออนไลน์ (Online Collaboration) ทำให้เกิดกิจกรรมทางสังคมที่เกี่ยวข้องกับการซื้อขายแลกเปลี่ยน การโฆษณาประชาสัมพันธ์ การแบ่งปันข้อมูลและประสบการณ์ (Turban, 2016) กล่าวอีกนัยหนึ่ง โซเชียลคอมเมิร์ซเป็นแนวคิดใหม่ที่เกิดจากการผสมผสานการช็อปปิ้งออนไลน์แบบดั้งเดิมและการตลาดแบบปากต่อปากซึ่งเป็นกระบวนทัศน์ใหม่ของการทำการค้าโดยใช้โซเชียลมีเดียเพื่อเข้าถึงผู้บริโภคและกลุ่มเพื่อนในเครือข่าย (Sohn, 2020) โซเชียลมีเดียออนไลน์สามารถสร้างชุมชนเสมือนจริง และด้วยแพลตฟอร์มโซเชียลเน็ตเวิร์ก โซเชียลคอมเมิร์ซยังสามารถส่งเสริมการขายสินค้าหรือบริการออนไลน์ (Liao, 2021) โซเชียลคอมเมิร์ซยังส่งเสริมให้เกิดการมีส่วนร่วมในการเข้าสู่ E-marketplace ผ่านโซเชียลมีเดียด้วยการแบ่งปันอย่างเท่าเทียม (Boardman, 2019)

ปัจจุบันผู้บริโภคใช้โซเชียลมีเดียต่างๆ เพื่อแสดงความคิดเห็นเกี่ยวกับคุณภาพ การบริการของร้านอาหารเพิ่มมากขึ้น เนื่องจากผู้บริโภคต้องการทบทวนความคิดเห็นของอีกฝ่ายก่อน

กระบวนการตัดสินใจ (Gao, 2018) ความคิดเห็นของผู้บริโภคเกี่ยวกับผลิตภัณฑ์เป็นตัวกำหนดว่าผลิตภัณฑ์นี้สามารถตอบสนองความต้องการของผู้บริโภคได้มากน้อยเพียงใด (Hussain, 2018) ข้อมูลที่รวบรวมจากโซเชียลมีเดียเป็นแพลตฟอร์มข้อมูลเพื่อปรับปรุงคุณภาพการบริการของร้านอาหารผ่านการรีวิวออนไลน์ของผู้บริโภคซึ่งเป็นแหล่งข้อมูลที่เชื่อถือได้ ซึ่งช่วยให้ผู้บริโภคประเมินคุณภาพอาหารได้ การพัฒนาวิธีการประเมินผลการรีวิวคุณภาพและบริการร้านอาหารที่ผู้บริโภคสร้างขึ้นอย่างมีประสิทธิภาพจึงเป็นสิ่งสำคัญ

แนวคิดเกี่ยวกับการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์ (eWOM Marketing)

การตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์ หมายถึงข้อมูลของผู้บริโภคที่ได้รับจากแหล่งข้อมูลที่เกิดจากความสัมพันธ์ส่วนบุคคล เช่น เพื่อน หรือครอบครัว ซึ่งมีอิทธิพลต่อความเร็วและอัตราการแพร่ของนวัตกรรม และเป็นกุญแจสำคัญในการนำผลิตภัณฑ์ไปใช้โดยส่งผลกระทบต่อกระแสข้อมูลในสังคม (Chu, 2019) eWOM เป็นตลาดที่ซับซ้อน ในการค้นหา eWOM ผู้บริโภคสามารถเลือกจากหลายแพลตฟอร์ม เช่น เว็บไซต์ตลาดกลางออนไลน์หรือเว็บไซต์อิสระ และจากหลากหลายรูปแบบ เช่น บทวิจารณ์ออนไลน์ บล็อก วิดีโอ เป็นต้น ซึ่งผู้บริโภคสามารถแสดงความคิดเห็นเกี่ยวกับผลิตภัณฑ์หรือบริการได้ตามข้อกำหนดของเว็บไซต์นั้นๆ เช่น การพิมพ์ข้อความ รูปภาพ หรือการให้คะแนน (Pyle, 2021) ด้วยการเพิ่มขึ้นของโซเชียลคอมเมิร์ซ ผู้บริโภคจึงมักจะพึ่งพาตัวชี้นำเพื่อลดความเสี่ยงจากความไม่แน่นอนโดยธรรมชาติอันเป็นผลมาจากการไม่สามารถจับต้องได้ของผลิตภัณฑ์ การบอกปากต่อปากทางอิเล็กทรอนิกส์ (eWOM) จึงกลายเป็นข้อมูลอ้างอิงที่สำคัญสำหรับผู้ใช้ในการตัดสินใจซื้อ (Zhao, 2020) ในภาคการบริการ ผู้บริโภคอาศัยการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์มากขึ้นในการค้นหาข้อมูลเกี่ยวกับผู้ให้บริการและแบ่งปันประสบการณ์ส่วนตัวในการใช้บริการ (Golmohammadi, 2020) เนื่องจากอารมณ์การบริโภคมีบทบาทสำคัญในการอธิบายพฤติกรรมผู้บริโภคในระยะหลังการซื้อ ดังนั้นการตลาดแบบปากต่อปากผ่านทางอิเล็กทรอนิกส์จึงเป็นส่วนสำคัญในการตอบสนองต่อพฤติกรรมหลังการซื้อของผู้บริโภคในยุคดิจิทัล (Yan, 2018)

แนวคิดเกี่ยวกับความเข้มข้นทางอารมณ์ (Emotional intensity)

อารมณ์สามารถเสริมสร้างความจำได้ ตัวอย่างเช่น ในชีวิตประจำวัน เหตุการณ์ไม่พึงประสงค์มักจะทิ้งความประทับใจไว้ เมื่อเทียบกับเหตุการณ์ที่เป็นกลาง เหตุการณ์ทางอารมณ์จะมีรายละเอียดเกี่ยวกับการรับรู้และประสาทสัมผัสมากกว่า ส่งผลให้เกิดความทรงจำได้มากขึ้น

ปรากฏการณ์นี้เรียกว่าการเสริมสร้างอารมณ์ในการจดจำ (Emotional Enhancement of Remembering) (Y. Wang, 2021) การรับรู้และอารมณ์ของผู้บริโภคจึงเป็นปัจจัยที่มีอิทธิพลต่อความตั้งใจในการซื้อของผู้บริโภค

การแสดงอารมณ์ความรู้สึกในการแสดงความคิดเห็นหรือบทวิจารณ์ออนไลน์เกี่ยวกับผลิตภัณฑ์หรือบริการบนสื่อที่มีรูปแบบแตกต่างกันระหว่างโซเชียลเน็ตเวิร์กกับเว็บไซต์รีวิว โดยในสื่อโซเชียลมีเดียผู้บริโภคจะมีอคติเชิงบวกในการแบ่งปัน eWOM เนื่องจากสภาพแวดล้อมออนไลน์มักเชื่อมโยงกับโลกออฟไลน์ ผู้บริโภคจึงมักให้ความสำคัญกับภาพลักษณ์ที่ดีของตนเองกับสมาชิกคนอื่นๆ ในโซเชียลเน็ตเวิร์ก ต่างจากการแสดงความคิดเห็นในเว็บไซต์รีวิวซึ่งผู้บริโภคและผู้เผยแพร่ข้อมูลต่างไม่รู้จักกัน การให้ eWOM จึงมุ่งเน้นไปที่ผลิตภัณฑ์ บริการ หรือแบรนด์ แทนที่จะเป็นตัวผู้บริโภคเอง ส่งผลให้การแสดงความคิดเห็นอิงตามความรู้สึกจริงของผู้บริโภค (H. Liu, 2021) ความเป็นประโยชน์ของบทวิจารณ์ออนไลน์หรือการผสมผสานระหว่างข้อมูลและการโน้มน้าวใจของบทวิจารณ์ออนไลน์สามารถเพิ่มความมั่นใจในการตัดสินใจ นอกจากนั้นแล้วความขัดแย้งกันในเชิงบวกหรือเชิงลบของบทวิจารณ์ออนไลน์จะทำหน้าที่เป็นสัญญาณช่วยให้ผู้บริโภคสามารถตรวจจับได้ว่าสถานการณ์นั้นต้องการความรู้ความเข้าใจที่มากขึ้นหรือไม่ (Ruiz-Mafe, 2018)

ความตั้งใจซื้อบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

ด้วยการผสมผสานระหว่างพาณิชย์อิเล็กทรอนิกส์ (e-Commerce) และสื่อสังคมออนไลน์ (Social Media) โซเชียลคอมเมิร์ซคือแพลตฟอร์มที่มีศักยภาพซึ่งมีอิทธิพลต่อความตั้งใจในการซื้อของผู้บริโภคอย่างมีนัยสำคัญ บนแพลตฟอร์มโซเชียลคอมเมิร์ซ ผู้บริโภคสามารถแบ่งปันประสบการณ์การช็อปปิ้ง การให้คะแนน การให้ความคิดเห็น และคำแนะนำจากผู้บริโภครายอื่น ซึ่งสามารถกระตุ้นความตั้งใจในการซื้อ (P. Liu, 2021) ความตั้งใจในการซื้อซ้ำและการบอกต่อแบบปากต่อปากได้รับผลกระทบเชิงบวกมาจากความไว้วางใจและความพึงพอใจ ซึ่งทั้งความไว้วางใจและความพึงพอใจได้รับผลกระทบเชิงบวกมาจากชื่อเสียงและคุณภาพของข้อมูล (Meilatinova, 2021) ความเชื่อถือในไซต์โซเชียลคอมเมิร์ซจะเพิ่มการสืบค้นข้อมูลภายในไซต์โซเชียลคอมเมิร์ซผ่านช่องทางต่างๆ เช่น ฟอรัม/ชุมชน บทวิจารณ์และการให้คะแนน และคำแนะนำเกี่ยวกับผลิตภัณฑ์ เพื่อรับข้อมูลที่เป็นประโยชน์สำหรับการซื้อ (Hajli, 2017) เศรษฐกิจ ความจำเป็น ความน่าเชื่อถือ ปฏิสัมพันธ์ และการส่งเสริมการขาย ส่งเหล่านี้ล้วนแต่ส่งผลต่อความตั้งใจในการซื้อบนแพลตฟอร์มโซเชียลคอมเมิร์ซ (Sohn, 2020)

การเรียนรู้ของเครื่อง (Machine learning)

การเรียนรู้ของเครื่อง (Machine learning; ML) คือการนำข้อมูลต่าง ๆ มาฝึกให้เครื่องจักรหรือคอมพิวเตอร์ทำการเรียนรู้จากชุดข้อมูลเรียนรู้ (Training Data set) เพื่อหาคำตอบ จากนั้นจึงนำวิธีการหาคำตอบที่ได้ไปใช้ในการพยากรณ์ผลของข้อมูลชุดใหม่ (Sebastian Raschka, 2019) การเรียนรู้ของเครื่องแบ่งออกเป็น 3 ประเภทได้แก่

1. การเรียนรู้โดยมีผู้สอน (Supervised Learning)

มีวัตถุประสงค์เพื่อพยากรณ์ข้อมูลที่ยังไม่ทราบค่าหรือข้อมูลในอนาคต โดยการนำข้อมูลหรือตัวแปรในอดีตที่ทราบคำตอบแล้วมาใช้ฝึกหัด (training data) โดยหากตัวแปรดังกล่าวมีค่าไม่ต่อเนื่องจะเรียกกระบวนการนี้ว่าการจำแนกกลุ่ม (classification) แต่หากข้อมูลหรือตัวแปรมีค่าต่อเนื่องจะเรียกกระบวนการนี้ว่าการถดถอย (regression) หรือการพยากรณ์ (prediction) ตัวอย่างเช่น การจำแนกด้วยแผนภูมิต้นไม้ (decision tree) หรือโครงข่ายประสาทเทียม (neural network) เป็นต้น

2. การเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning)

เป็นการนำข้อมูลที่ยังไม่ทราบคำตอบหรือความสัมพันธ์ระหว่างตัวแปรต่าง ๆ นำมาศึกษาหาความสัมพันธ์ (association) หรือการจัดกลุ่ม (clustering) โดยไม่ได้มีจุดมุ่งหมายเพื่อการทำนายหรือพยากรณ์คำตอบในอนาคตแต่เป็นการพยายามหาความสัมพันธ์ของข้อมูล เช่น การหากฎความสัมพันธ์ หรือการแบ่งกลุ่มข้อมูลที่มีลักษณะใกล้เคียงกัน เป็นต้น

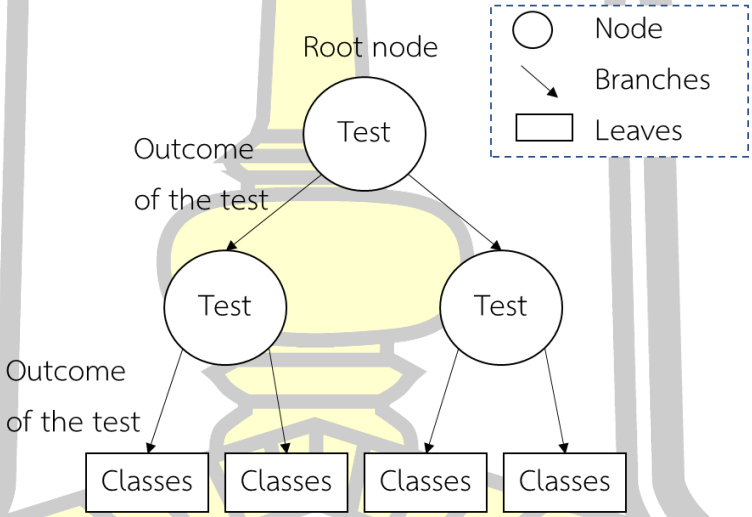
3. การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)

มีวัตถุประสงค์เพื่อสร้างระบบซึ่งสามารถปรับปรุงประสิทธิภาพบนพื้นฐานของการปฏิสัมพันธ์กันสภาพแวดล้อมจากการลองผิดลองถูกภายใต้สถานการณ์ต่าง ๆ ซึ่งมีองค์ประกอบหลักได้แก่ ผู้กระทำ (Agent) ระบบที่ Agent มีปฏิสัมพันธ์ด้วย (Environment) การกระทำของ Agent ซึ่งส่งผลต่อ Environment (Action) สถานการณ์ของ Environment ที่ทาง Agent สามารถรับรู้ได้ (State) สิ่งที่กำหนดให้ Agent ตัดสินใจเลือก Action หลังจากการประเมินสถานการณ์แล้วโดยอาศัย Markov Decision Process ซึ่งเป็น Framework สำหรับการตัดสินใจ (Decision Making)

(Policy) และสิ่งที่ประเมินผลลัพธ์ที่เกิดจากการกระทำของ Agent (Reward) เช่น คะแนน ผลแพ้ชนะ เป็นต้น ตัวอย่างของการเรียนรู้แบบ Reinforcement Learning เช่น AlphaGO เป็นต้น

ต้นไม้ตัดสินใจ (Decision Tree)

Decision tree เป็นอัลกอริธึมที่เป็นที่นิยมอย่างกว้างขวาง ใช้งานง่าย เข้าใจง่าย และมีประสิทธิภาพ Decision Tree มีลักษณะคล้ายกับต้นไม้กลับหัวซึ่งประกอบด้วยราก (Root node) ซึ่งเป็นจุดเริ่มต้นของลำดับการตัดสินใจ ถัดลงมาคือกิ่ง (Branch node) เป็นตัวเชื่อมระหว่างโหนดที่เป็นเงื่อนไขการตัดสินใจซึ่งจะแบ่งข้อมูลออกเป็น 2 กิ่งหรือมากกว่าขึ้นอยู่กับเงื่อนไขของ Child node และสุดท้ายคือใบ (Leaf node) ซึ่งเป็นโหนดสิ้นสุด ซึ่งเป็นผลลัพธ์ที่ได้จากการจัดหมวดหมู่ข้อมูลหลังจากการทำ Classification (ปริญญา สงวนสัตย์, 2019)



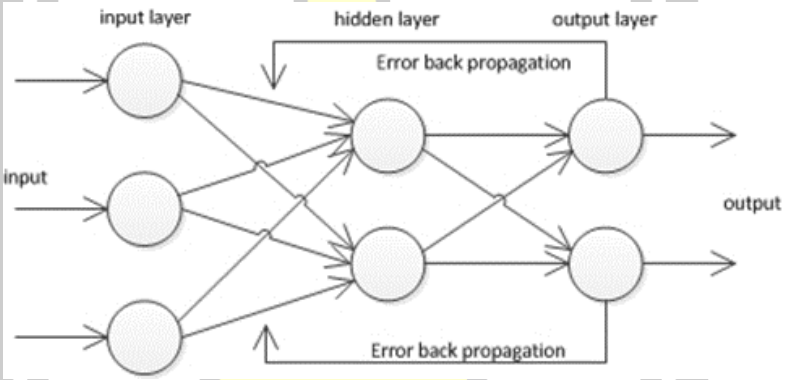
ภาพประกอบ 5 โครงสร้างต้นไม้ตัดสินใจ

Decision Tree คือแบบจำลองที่มีลักษณะเป็น rule-based คือ การสร้างกฎ if-else จากค่าของแต่ละตัวแปรอิสระโดยไม่มีสมการมากำกับความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม ซึ่งวิธีนี้มีจุดเด่นในเรื่องของความชัดเจนและง่ายต่อการนำไปประยุกต์ใช้

โครงข่ายประสาทเทียม (Neural Network)

โครงข่ายประสาทเทียม (ANN) คือระบบที่ซับซ้อนโดยสามารถเรียนรู้ด้วยตนเอง ซึ่งเลียนแบบการทำงานของระบบโครงข่ายประสาทในสมองของมนุษย์ ซึ่งประกอบด้วยหน่วยของการปฏิบัติการและเชื่อมโยงกันด้วยโครงข่ายจำนวนมาก (Desai, 2021) ประสิทธิภาพของโครงข่ายประสาทเทียมขึ้นอยู่กับขั้นตอนการเรียนรู้ โดยผู้ใช้จะต้องกำหนดสถาปัตยกรรมของเครือข่ายและกำหนดพารามิเตอร์ที่เกี่ยวข้อง (Du, 2019)

สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น ประกอบด้วย ชั้นข้อมูลนำเข้า (input layer) ชั้นซ่อน (hidden layer) และชั้นส่งออก (output layer) ดังภาพประกอบ 6



ภาพประกอบ 6 สถาปัตยกรรมของโครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น

แต่ละชั้นประกอบด้วยเซลล์ประสาทจำนวนหนึ่ง โดยข้อมูลจะไหลไปข้างหน้าอย่างต่อเนื่องและค่าผิดพลาดของเครือข่ายจะแพร่กระจายไปข้างหลัง การเชื่อมต่อค่าน้ำหนักและเกณฑ์ของเครือข่ายจะเริ่มต้นแบบสุ่มเป็นครั้งแรก จากนั้นจะใช้กระบวนการฝึกสอนโดยทำการเปรียบเทียบค่าความผิดพลาดของผลลัพธ์กับค่าจริง ค่าน้ำหนักและเกณฑ์ของเครือข่ายจะถูกปรับเพื่อลดค่าความผิดพลาด mean square error (MSE) จนกระทั่งจะได้ผลลัพธ์ที่ใกล้เคียงกับผลลัพธ์ที่ต้องการ (L. Wang, 2015)

การวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis)

คือเทคนิคเชิงเส้นใช้กันอย่างแพร่หลายในการทำนายตัวแปรตามเพียงตัวเดียว (ตัวแปรผลลัพธ์) กับตัวแปรอิสระ (ตัวแปรพยากรณ์) หนึ่งตัวหรือหลายตัวแปร โดยการกำหนดความสัมพันธ์เชิงเส้นระหว่างตัวแปรผลลัพธ์กับตัวแปรอิสระที่ปรากฏในรูปแบบที่ไม่เป็นเชิงเส้น (Chiu, 2008)

กระบวนการพื้นฐานของการถดถอยเชิงเส้นคือการประมาณค่าสัมประสิทธิ์สำหรับตัวแปรทำนายแต่ละตัว (X) เพื่อแสดงถึงผลที่ตัวแปรมีต่อตัวแปรผลลัพธ์ (Y) โดยสมการทั่วไปของการวิเคราะห์การถดถอยแสดงดังสมการที่ (1)

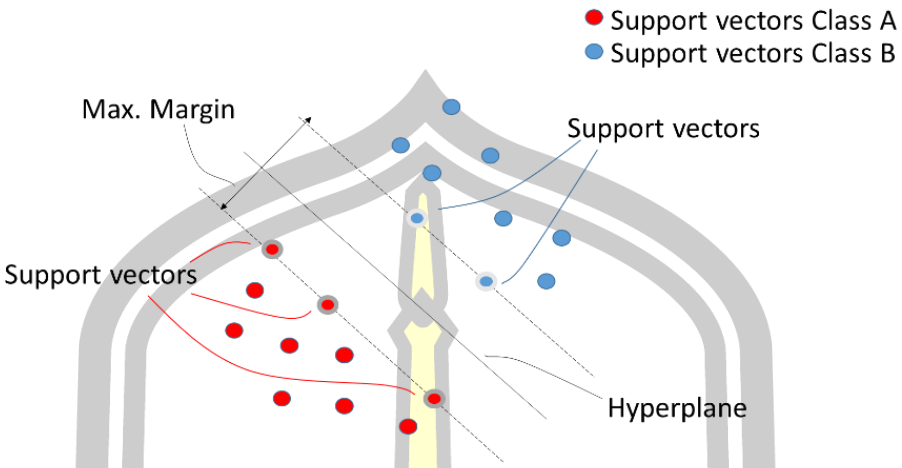
$$\gamma=a+b_1X_1+b_2X_2+b_3+\ldots+b_nX_n \qquad (1)$$

X แทน ตัวแปรทำนาย
Y แทน ตัวแปรผลลัพธ์
a แทน Y เมื่อ X เท่ากับ 0
b แทน ค่าสัมประสิทธิ์การถดถอย

การถดถอยเชิงเส้นที่มีตัวแปรอิสระเพียงตัวเดียวเรียกว่า Simple Linear Regression และการถดถอยเชิงเส้นที่มีตัวแปรอิสระหลายตัวเรียกว่า Multiple Linear Regression (Nisbet, 2009)

ซัพพอร์ตเวกเตอร์แมชชีน (SVM)

เป็นการเรียนรู้แบบมีผู้สอน ที่ได้รับความนิยมมากที่สุดเทคนิคหนึ่ง โดยเฉพาะอย่างยิ่งกับการสร้างแบบจำลองที่ไม่เป็นเชิงเส้นสำหรับระบบและกระบวนการที่ซับซ้อน (David L. Olson, 2008) สามารถใช้ได้กับการจำแนกประเภท (Classify) และการถดถอย (Regression) การทำงานของซัพพอร์ตเวกเตอร์แมชชีน จะทำการแบ่งข้อมูลออกเป็น 2 กลุ่มใน Feature Space ข้อมูลจะถูกแบ่งออกเป็นสองกลุ่มโดยเส้นตรงที่เรียกว่า Hyperplane ซึ่งเส้นตรงที่มีระยะห่างจากแต่ละกลุ่มน้อยที่สุดคือเส้นที่เหมาะสมที่สุด (X.-Y. Wang, 2016) โดยหลักการทำงานของเทคนิค Support Vector Machine (SVM) แสดงได้ดังภาพประกอบ 7



ภาพประกอบ 7 หลักการทำงานของเทคนิค Support Vector Machine (SVM)

การคัดเลือกคุณสมบัติ (Feature Selection)

ในยุคปัจจุบันที่การขับเคลื่อนด้วยข้อมูล (data-driven) เป็นสิ่งสำคัญในการดำเนินธุรกิจ ทำให้เกิดการเพิ่มขึ้นของข้อมูลและมิติของข้อมูลอย่างมหาศาล การลดขนาดข้อมูลโดยการการคัดเลือกคุณสมบัติจึงมักถูกนำมาใช้เป็นขั้นตอนก่อนการประมวลผลข้อมูลที่สำคัญในการออกแบบระบบที่มีประสิทธิภาพ โดยยังรักษาคุณภาพข้อมูลที่เหมาะสมที่สุดเท่าที่จะทำได้

การคัดเลือกคุณสมบัติสำหรับการพยากรณ์ด้วยการจำแนกกลุ่ม (Classification) มีจุดมุ่งหมายเพื่อเพิ่มประสิทธิภาพของการพยากรณ์ โดยการคัดเลือกเฉพาะตัวแปรที่ส่งผลต่อการเพิ่มขึ้นของประสิทธิภาพของตัวแบบพยากรณ์จากจำนวนตัวแปรอิสระทั้งหมด (Xu, Lu, Li, & Yuan, 2022) ประโยชน์อีกประการหนึ่งคือช่วยลดความซับซ้อนของกระบวนการเรียนรู้และลดการเกิด Overfitting (Ba, Wang, Yang, Yu, & Yu, 2023) ซึ่งเกิดขึ้นเมื่อรูปแบบการเรียนรู้ของเครื่องสามารถพยากรณ์ได้อย่างถูกต้องสำหรับชุดข้อมูลเรียนรู้ (Training Data set) แต่ไม่สามารถให้ผลการพยากรณ์ที่ถูกต้องสำหรับข้อมูลใหม่

1. การคัดเลือกตัวแปรโดยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

เป็นการนำตัวแปรอิสระทุกตัวที่ศึกษาที่มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญทางสถิติและไม่มีนัยสำคัญทางสถิติเข้าไปวิเคราะห์ในสมการถดถอยโดยไม่มีการคัดเลือกพร้อมกันทุก

ตัว ถึงแม้ว่าตัวแปรอิสระจะไม่ทำให้ค่าการพยากรณ์ดีขึ้นก็ตามวิธีนี้มักจะใช้ในกรณีที่ต้องการทราบว่าตัวแปรแต่ละตัวที่ทำการศึกษาจะสามารถพยากรณ์ตัวแปรตามได้หรือไม่มากน้อยเพียงใด ข้อด้อยคือเป็นการวิเคราะห์ที่ไม่ได้คัดเลือกตัวแปรเข้าสู่สมการถดถอยที่เหมาะสมทำให้เป็นวิธีที่สิ้นเปลืองเวลาและทรัพยากร

2. การเลือกตัวแปรโดยวิธีเพิ่มตัวแปร (Forward Selection)

เป็นวิธีการคัดเลือกตัวแปรอิสระเข้าสู่สมการทีละตัว โดยใช้เกณฑ์ในการพิจารณาคือค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (Simple correlation coefficient) โดยตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมากที่สุดจะถูกคัดเลือกเข้าก่อน จากนั้นจะทำการทดสอบว่าตัวแปรอิสระนั้นสามารถพยากรณ์ตัวแปรตามได้อย่างมีนัยสำคัญทางสถิติหรือไม่ จากนั้นจะทำการคัดเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอันดับถัดไป ซึ่งเมื่อรวมกับตัวแปรอิสระที่เลือกไว้ก่อนหน้าต้องให้คะแนนเกณฑ์ที่ดีที่สุดในแต่ละรอบ กระบวนการดำเนินไปจนกว่าเกณฑ์จะดีขึ้นเรื่อย ๆ ทำเช่นนี้ไปจนกว่าจะไม่สามารถคัดเลือกตัวแปรอิสระใดเข้าสู่สมการได้อีกจึงหยุดการคัดเลือกตัวแปรอิสระและถือว่าสมการที่ได้นั้นมีความเหมาะสม (Bhadra & Bandyopadhyay, 2021)

3. การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward Elimination)

เป็นวิธีที่ตรงข้ามกับการเลือกตัวแปรโดยวิธีเพิ่มตัวแปร โดยเริ่มจากการสร้างสมการถดถอยที่นำตัวแปรอิสระทุกตัวเข้าสู่สมการ ทำการคำนวณค่า p-value ของตัวแปรอิสระทุกตัวแปรในตัวแบบพยากรณ์ จากนั้นจึงพิจารณาตัวแปรที่มีค่า p-value มากที่สุด และทำการเปรียบเทียบค่า p-value กับระดับนัยสำคัญที่กำหนด เช่น Pr = 0.05 ถ้า p-value มากกว่าระดับนัยสำคัญที่กำหนดจะทำการตัดตัวแปรนั้นออกจากสมการ ทำซ้ำกระบวนการทั้งหมดกับตัวแปรที่เหลือจนกระทั่งไม่มีตัวแปรใดมีค่า p-value มากกว่าระดับนัยสำคัญที่กำหนดจึงหยุดการคัดเลือกตัวแปรอิสระ และถือว่าสมการที่ได้นั้นมีความเหมาะสม (Yoni Nazarathy, 2021)

การวัดประสิทธิภาพของแบบจำลอง (Method Evaluation)

Confusion Matrix สามารถใช้ในการวัดค่าความแมนยำสำหรับการจำแนกแบบไบนารี (Binary classification) หรือ ตัวแปรที่มีผลลัพธ์เพียง 2 กลุ่ม ในขณะที่ Regression จะใช้วิธีการวัดผล เช่น ค่า R-squared ค่า Mean Absolute Percentage Error (MAPE) หรือ RMSE

การเรียนรู้โดยมีผู้สอน (Supervised Learning) สามารถวัดค่าความแม่นยำ (Accuracy) ได้ เนื่องจากมี Target เป็นข้อมูลตั้งต้น ทำให้สามารถนำค่าพยากรณ์ที่ได้นำมาวัดค่าความคาดเคลื่อนจากข้อมูลตั้งต้นได้ ดังนั้นในกรณีของทั้ง Classification และ Regression จึงสามารถประสิทธิภาพของแบบจำลองได้จากค่าความแม่นยำ (Accuracy)

ตาราง 1 เมทริกซ์ความสับสน (Confusion matrix)

	True Positive	True Negative
Prediction Positive	a	b
Prediction Negative	c	d

1. ค่าความถูกต้อง (Accuracy)

คือค่าที่แสดงว่าตัวแบบพยากรณ์สามารถทำนายได้อย่างถูกต้องซึ่งแสดงค่าเป็นร้อยละ โดยสามารถคำนวณได้จากสมการที่ 2

$$\text{Accuracy} = \frac{a+b}{(a+b+c+d)} \quad (2)$$

2. ค่าความระลึก (Class Recall)

คือ ค่าความครบถ้วนที่แสดงให้เห็นถึง ความสามารถของตัวแบบพยากรณ์ในการเลือกข้อมูลหรือคำตอบที่เกี่ยวข้องได้จำนวนมากน้อยเพียงใดแบ่งเป็น 2 ชนิดคือค่าความไวในการตรวจพบค่าบวก (sensitivity) และค่าความจำเพาะในการตรวจพบค่าลบ (specificity) โดยสามารถคำนวณได้จากจากสมการที่ 3 และ 4

$$\text{Sensitivity} = \frac{a}{a+c} \quad (3)$$

$$\text{Specificity} = \frac{d}{b+d} \quad (4)$$

3. ค่าความแม่นยำ (Class Precision)

คือ การวัดความสามารถของตัวแบบพยากรณ์ในการขจัดคำตอบหรือข้อมูลที่ไม่เกี่ยวข้องออกไป ถ้าตัวแบบพยากรณ์สามารถขจัดคำตอบหรือข้อมูลที่ไม่เกี่ยวข้องออกไปได้มากแสดงถึงความแม่นยำของตัวแบบพยากรณ์ ประกอบด้วยค่าการทำนายผลบวก (positive predictive value; PPV) และค่าการทำนายผลลบ (negative predictive value: NPV) สามารถคำนวณได้จากสมการที่ 5 และ 6

$$\mathrm{PPV} = \frac{a}{a+b} \tag{5}$$

$$\mathrm{NPV} = \frac{d}{c+d} \tag{6}$$

4. ค่าพื้นที่ใต้กราฟ ROC (Area Under the Curve; AUC)

คือ ค่าแสดงความสามารถในการจำแนกกลุ่มของตัวแบบพยากรณ์ หากค่า AUC มีค่าเข้าใกล้1 มากเท่าใดแสดงวาตัวแบบพยากรณ์มีความสามารถในการจำแนกกลุ่มออกจากกันได้อย่างถูกต้อง

งานวิจัยที่เกี่ยวข้อง

Yuan (2017) นำเสนอการค้นหาความรู้สึกของผู้บริโภคเกี่ยวกับผลิตภัณฑ์จากบทวิจารณ์ตามหัวข้อบนแพลตฟอร์มอีคอมเมิร์ซด้วยการทำเหมืองข้อมูลโดยใช้เทคนิคโครงข่ายประสาทเทียมเพื่อช่วยสนับสนุนการตัดสินใจของผู้บริหารเพื่อเพิ่มประสิทธิภาพในการพยากรณ์การขาย จากกรณีศึกษา บริษัท TaoBao ซึ่งเป็นอีคอมเมิร์ซที่ใหญ่ที่สุดแห่งหนึ่งในประเทศจีน โดยมีวัตถุประสงค์เพื่อ (1) ออกแบบวิธีการใหม่เพื่อดึงความรู้สึกและหัวข้อจากบทวิจารณ์ผลิตภัณฑ์ออนไลน์จำนวนมาก และ (2) เพื่อพัฒนารูปแบบการพยากรณ์ใหม่ที่ใช้ประโยชน์จากคุณลักษณะความรู้สึกของหัวข้อเชิงคุณภาพ และคุณลักษณะเชิงปริมาณเพื่อพยากรณ์ยอดขายผลิตภัณฑ์ โดยการระบุป้ายกำกับความรู้สึกตามหัวข้อแบ่งออกเป็นสามป้าย ได้แก่ เชิงบวก เชิงลบ และเป็นกลาง โดยทำการรวบรวมข้อมูลได้แก่ ข้อมูลการขายในอดีต การเปลี่ยนแปลงราคา กิจกรรมส่งเสริมการขาย และความคิดเห็น การทดสอบเชิงประจักษ์แสดงให้เห็นว่าหากมีการนำความรู้สึกที่มีต่อหัวข้อต่างๆ จากบทวิจารณ์บน

แพลตฟอร์มอีคอมเมิร์ซร่วมกับคุณลักษณะเชิงปริมาณอื่นๆ จะสามารถพยากรณ์ปริมาณการขายได้แม่นยำกว่าเมื่อเปรียบเทียบกับการใช้คุณลักษณะเชิงปริมาณเพียงอย่างเดียว (Yuan, 2017)

Namin (2017) ได้ศึกษาปัจจัยที่ส่งผลกระทบต่อความพึงพอใจของผู้บริโภคร้านอาหารฟาสต์ฟู้ด โดยทำการศึกษาจากวรรณกรรมโดยใช้การวิเคราะห์ปัจจัยเชิงยืนยันการสร้างแบบจำลองสมการโครงสร้างและการวิเคราะห์เส้นทาง ผลการศึกษาชี้ให้เห็นว่าการปรับปรุงคุณภาพการบริการสำหรับร้านอาหารฟาสต์ฟู้ดไม่ส่งผลโดยตรงต่อการเพิ่มความตั้งใจเชิงพฤติกรรมของผู้บริโภค แต่ความตั้งใจในเชิงพฤติกรรมสามารถเพิ่มขึ้นได้โดยอาศัยความพึงพอใจของผู้บริโภค ความพึงพอใจของผู้บริโภคสามารถปรับปรุงได้ผ่านคุณภาพการบริการ คุณภาพอาหาร และอัตราส่วนของราคาต่อคุณค่า คุณภาพเป็นปัจจัยสำคัญที่ทำให้ผู้บริโภคพึงพอใจ (Namin, 2017)

Fernandes (2021) ได้ทำการวิจัยแนวทางในการลดความซับซ้อนของกระบวนการตัดสินใจของผู้จัดการร้านอาหารโดยการรวมความคิดเห็นของผู้บริโภคบนโซเชียลมีเดียและข้อมูลการขายในอดีตในรูปแบบการพยากรณ์การขาย โดยการตรวจสอบข้อมูลทั้งภายในและภายนอก เช่น บทวิจารณ์ออนไลน์และรายได้ที่รวบรวมจากร้านอาหารหกแห่ง ผลการวิจัยสรุปได้ว่าคุณภาพอาหารและการบริการมีความสำคัญและได้รับการกล่าวถึงอย่างสูงในบทวิจารณ์ที่เป็นบวกน้อยกว่า จำนวนผู้บริโภคต่อวันและรายได้มีอิทธิพลในทางลบต่อคะแนนรีวิว เนื่องจากการมีผู้บริโภคจำนวนมากทำให้เวลาในการรอรับบริการนานขึ้น และผู้บริโภคมีแนวโน้มที่จะพึงพอใจมากขึ้นเมื่อมีประสบการณ์รอที่สั้นกว่าก่อนที่จะได้รับบริการ (Fernandes, 2021)

Sohn (2020) ได้ทำการวิจัยความสัมพันธ์ที่หลากหลายระหว่างบทวิจารณ์ออนไลน์กับแบรนด์ในบริบทของร้านอาหาร โดยใช้การวิเคราะห์กำลังสองน้อยที่สุดแบบสองขั้นเพื่อวิเคราะห์ผลกระทบของปัจจัยทั้งสองที่มีต่อประสิทธิภาพทางการเงินของร้านอาหาร 1,210 แห่งในรัฐไอโอวาระหว่างปี 2010 ถึง 2017 ผลจากการวิจัยพบว่าผลกระทบทางการเงินจากบทวิจารณ์ในเชิงบวกลดลงเมื่อร้านอาหารมีแบรนด์ และการมีส่วนร่วมของผู้บริโภคในการรีวิวออนไลน์นั้นแตกต่างกันไปตามความแข็งแกร่งของแบรนด์ จากการค้นพบนี้ผู้วิจัยได้เสนอว่าผู้จัดการร้านอาหารควรเปลี่ยนจุดเน้นของกลยุทธ์การตลาดเมื่อความแข็งแกร่งของแบรนด์เติบโตขึ้น (Sohn, 2020)

Zhang (2020) ได้วิจัยการประยุกต์ใช้เทคโนโลยีการทำเหมืองข้อมูลกับอุตสาหกรรมร้านอาหาร โดยการใช้อัลกอริธึม linear regression (LR) และ random forest (RF) ในการคาดการณ์ปริมาณลูกค้าและความต้องการของลูกค้า เพื่อให้ร้านอาหารสามารถจัดเตรียมอาหารในปริมาณที่เหมาะสมกับความต้องการของลูกค้าตามที่คาดการณ์ไว้ ซึ่งจะเอื้อต่อการปรับปรุงประสบการณ์การใช้บริการของลูกค้าและปรับปรุงคุณภาพในการเตรียมอาหารและทำให้ร้านอาหารดำเนินการได้อย่างมีประสิทธิภาพมากยิ่งขึ้น ผลการวิจัยพบว่าการสร้างโมเดลโดยการผสมผสานกันระหว่างอัลกอริธึมทั้งสอง สามารถลดอัตราความผิดพลาดลงได้ประมาณ 5.503% เมื่อเทียบกับ

อัลกอริธึม linear regression และต่ำกว่าอัตราความผิดพลาดของอัลกอริธึม random forest ประมาณ 3.719% (Zhang, 2020)

Li (2020) ได้วิจัยผลกระทบของความเข้มข้นทางอารมณ์ต่อการรับรู้ถึงประโยชน์ โดยรวบรวมข้อมูลจากบทวิจารณ์ 600,686 รายการจากร้านอาหารยอดนิยม 300 แห่งในสหรัฐอเมริกาจากเว็บไซต์ Yelp โดยใช้วิธีเหมืองข้อความและการวิเคราะห์ทางเศรษฐมิติ ผลลัพธ์เชิงประจักษ์แสดงให้เห็นว่า ความเข้มข้นทางอารมณ์เชิงบวกมีผลกระทบในทางลบต่อการรับรู้ถึงประโยชน์ของการทบทวน ในขณะที่ความรุนแรงทางอารมณ์เชิงลบมีผลกระทบเชิงบวกต่อการรับรู้ถึงประโยชน์ของการทบทวนที่รับรู้ ในบรรดาอารมณ์เชิงลบที่ไม่ต่อเนื่องกันที่แพร่หลายมากที่สุดในบทวิจารณ์ออนไลน์ (เช่น ความโกรธและความวิตกกังวล) บทวิจารณ์ที่แสดงความโกรธนั้นมีประโยชน์มากกว่าการแสดงความวิตกกังวล และ ระยะเวลาในการทบทวนและความเชี่ยวชาญของผู้ตรวจทานสามารถกลั่นกรองผลกระทบของความเข้มข้นทางอารมณ์ต่อการรับรู้ถึงประโยชน์ของการทบทวนที่รับรู้ได้ (Li, 2020)

Jia (2020) ได้ทำการวิจัยเกี่ยวกับแรงจูงใจและความพึงพอใจของนักท่องเที่ยวชาวจีนและชาวอเมริกันในธุรกิจร้านอาหาร การขุดเหมืองข้อความ (Text Mining) ข้ามวัฒนธรรมของบทวิจารณ์ออนไลน์ เนื่องจากนักท่องเที่ยวที่มีภูมิหลังทางวัฒนธรรมต่างกัน คิดและปฏิบัติตนแตกต่างกัน การเข้าใจความแตกต่างทางวัฒนธรรมและการจับตลาดอย่างถูกต้องจะช่วยให้ผู้จัดการการท่องเที่ยวสร้างความพึงพอใจของลูกค้าได้มากขึ้นและเพิ่มรายได้ของธุรกิจ โดยได้ทำการเปรียบเทียบแรงจูงใจและความพึงพอใจของลูกค้าที่เป็นนักท่องเที่ยวร้านอาหารที่มาจากประเทศจีนและสหรัฐอเมริกาโดยการตรวจสอบการให้คะแนนและรีวิวออนไลน์ โดยมีการดึงข้อมูลการให้คะแนนของลูกค้าและคำวิจารณ์หาปริมาณ เหมืองข้อความ (Text Mining) เปรียบเทียบ และตีความโดยใช้สถิติ และการวิเคราะห์ความถี่ ผลการวิจัยชี้ว่า นักท่องเที่ยวชาวจีนมีแนวโน้มที่จะให้คะแนนร้านอาหารที่ต่ำกว่าเล็กน้อยและรู้สึกประทับใจกับอาหารที่ถูกนำเสนอ ในขณะที่นักท่องเที่ยวชาวอเมริกันมักจะสนุกสนานกับการแสวงหา และไม่สบายใจกับความแออัดยัดเยียด (Jia, 2020)

Chen (2016) ได้ทำการวิจัยเกี่ยวกับกรอบการทำเหมืองข้อมูลตามทฤษฎี Rough Set เพื่อปรับปรุงการตัดสินใจในการเลือกสถานที่ กรณีศึกษาเครือธุรกิจร้านอาหาร โดยได้พัฒนากรอบงานการทำเหมืองข้อมูลโดยใช้ทฤษฎี Rough Set (RST) เพื่อสนับสนุนการตัดสินใจเลือกสถานที่ กรอบงานที่เสนอประกอบด้วย 4 ขั้นตอนได้แก่ 1) คำจำกัดความของปัญหาและการรวบรวมข้อมูล 2)การวิเคราะห์ RST 3) การตรวจสอบกฎ และ 4) การสกัดและการใช้ความรู้ การศึกษาเชิงประจักษ์มุ่งเน้นไปที่เครือร้านอาหารเพื่อแสดงความถูกต้องของแนวทางที่นำเสนอ มีการตรวจสอบตัวแปรตำแหน่ง 20 ตัวแปรที่เกี่ยวข้องกับ 5 ด้านของสถานที่ และผลการวิจัยพบว่าสามารถระบุความรู้ที่แฝงอยู่ได้เพื่อสนับสนุนการตัดสินใจเลือกสถานที่ จากผลการวิเคราะห์ ปัจจัยที่ตั้งที่สำคัญที่สุดที่ส่งผลต่อ

ประสิทธิภาพของร้านค้าคือ ขนาดร้าน ความพร้อมใช้งานของพื้นที่จอดรถ การมองเห็นร้าน และอัตราการเติบโตของประชากรในบริเวณใกล้เคียง (Chen, 2016)

Chang (2019) ได้ทำการวิจัยเรื่องการคาดการณ์ประสิทธิภาพของธุรกิจในโซเชียลคอมเมิร์ซตามสถานที่ตั้ง โดยใช้เทคนิคเหมืองข้อมูลจากแพลตฟอร์มโซเชียลคอมเมิร์ซ เพื่อค้นหาข้อมูลเชิงลึกสำหรับธุรกิจที่อยู่ร่วมกันและแข่งขันกันในละแวกใกล้เคียงทางภูมิศาสตร์ เพื่อสนับสนุนการตัดสินใจทางธุรกิจเชิงกลยุทธ์และการดำเนินธุรกิจได้ดียิ่งขึ้น โดยพัฒนากรอบการคำนวณแบบรวมศูนย์ที่รวบรวมสถานที่ตั้งของธุรกิจและข้อมูลตามบริบทในการรวมตัวกันทางภูมิศาสตร์สำหรับการคาดการณ์ประสิทธิภาพ โมเดลได้รับการประเมินโดยใช้ทั้งชุดข้อมูลสังเคราะห์และชุดข้อมูลจริง โดยการสร้างแบบจำลองได้แก่ 1) แบบจำลองความสัมพันธ์ของคุณลักษณะ 2) แบบจำลองความสัมพันธ์ทางภูมิศาสตร์ 3) แบบจำลองความสัมพันธ์ตามบริบท 4) แบบจำลองไฮบริด โดยใช้ข้อมูลจาก yelp.com ซึ่งให้บริการโซเชียลคอมเมิร์ซ เพื่อช่วยให้ผู้ใช้สามารถค้นหาธุรกิจโดยพิจารณาจากบทวิจารณ์และคำแนะนำของสมาชิกในปี 2013 จากธุรกิจร้านอาหารทั้งหมด 10,618 แห่ง ผลการวิจัยแสดงให้เห็นว่าแบบจำลองไฮบริดมีประสิทธิภาพในการสนับสนุนที่ดีขึ้นแก่ธุรกิจเกี่ยวกับการตัดสินใจเชิงกลยุทธ์ เช่น การเลือกไซต์ และการตัดสินใจในการดำเนินงาน เช่น ความแตกต่างของผลิตภัณฑ์/บริการ (Chang, 2019)

Sohn (2020) ได้ทำการศึกษาปัจจัยที่มีอิทธิพลต่อความตั้งใจในการซื้อบนโซเชียลคอมเมิร์ซ ผลการวิจัยพบว่า การวิเคราะห์ปัจจัยเผยให้เห็นคุณลักษณะ 5 ประการที่สามารถใช้ในการจำแนกโซเชียลคอมเมิร์ซ ได้แก่ เศรษฐกิจ ความจำเป็น ความน่าเชื่อถือ ปฏิสัมพันธ์ และการส่งเสริมการขาย ประการที่สอง จากการวิเคราะห์การถดถอยแบบพหุคูณ เศรษฐกิจ ความจำเป็น ความน่าเชื่อถือ และการส่งเสริมการขายจะส่งผลต่อความตั้งใจในการซื้อ ผลการวิจัยสามารถช่วยให้ผู้ประกอบการในธุรกิจการค้าเพื่อสังคมสามารถดึงดูดลูกค้าได้มากขึ้นด้วยการกำหนดเหตุผลที่ผลักดันการซื้อและความต้องการ นอกจากนี้ยังสามารถช่วยในการพัฒนากลยุทธ์ในการจัดการการค้าเพื่อสังคมได้อย่างมีประสิทธิภาพมากขึ้น (Sohn, 2020)

Najwa AlGhamdi (2021) ได้ทำการศึกษาเปรียบเทียบประสิทธิภาพการทำงานของ Ensemble Method 3 วิธี ได้แก่ Bagging, AdaBoost และ Stacking โดยใช้การเรียนรู้ของเครื่องได้แก่ Naive Bayes, Linear Regression, Decision Tree, K-Nearest Neighbor และ Support Vector Machine เพื่อทำนายการจำแนกประเภทความรู้สึกรีวิวออนไลน์ใน 3 โดเมนการที่แตกต่างกัน ได้แก่ โทรศัพท์มือถือ ร้านอาหาร และภาพยนตร์ จากผลการทดลองพบว่าวิธี Stacking ได้ผลลัพธ์ที่ดีที่สุด โดยมีค่าความถูกต้อง 83% (Najwa AlGhamdi, 2021)

ปราโมชย์ นามวงศ์ (2022) ได้ทำการวัดประสิทธิภาพการจำแนกความคิดเห็นบนเครือข่ายสังคมออนไลน์เฟสบุ๊คของผู้ใช้บริการร้านอาหารในจังหวัดอุบลราชธานี โดยใช้เทคนิคเหมืองข้อความ

ในการจำแนกความคิดเห็นออกเป็น เชิงบวก (Positive) และเชิงลบ (Negative) เพื่อการเปรียบเทียบประสิทธิภาพในการจำแนกข้อความ งานวิจัยนี้ได้เลือก 5 อัลกอริทึม ที่ได้รับความนิยมในการจำแนกประเภทข้อความ ได้แก่ Support Vector Machine (SVM) Naïve-Bayes (NB) Decision Tree (DT) K-Nearest Neighbor (KNN) และ Long Short-term Memory (LSTM) ซึ่งจากผลการทดลองพบว่าอัลกอริทึมที่ให้ประสิทธิภาพในการจำแนกสูงที่สุดคือ DT ได้ค่าความถูกต้องเท่ากับ 89.00 เปอร์เซ็นต์ และเมื่อพิจารณาประสิทธิภาพในการจำแนกข้อความโดยรวมด้วยค่าประสิทธิภาพโดยรวม (F-measure) อัลกอริทึมที่ให้ประสิทธิภาพสูงที่สุดคือ SVM ได้ค่าประสิทธิภาพการจำแนกโดยรวมเท่ากับ 90.46 เปอร์เซ็นต์ (ปราโมชย์ นามวงศ์, 2022)

ชาลิสา จิตบุญญาพินิจ (2022) ได้นำเสนอการพัฒนาแบบจำลองการวิเคราะห์ความรู้สึกบนสื่อสังคมออนไลน์ไทยโดยใช้เทคนิคการเรียนรู้เชิงลึกประกอบด้วย โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network) หน่วยความจำระยะยาว-ระยะสั้น (Long Short-Term Memory) และ หน่วยเวียนกลับแบบมีประตู (Gated Recurrent Unit) โดยการทดสอบกับชุดข้อมูลความคิดเห็นต่อสินค้าและบริการของเว็บวงใน (Wongnai) เมื่อวัดประสิทธิภาพจากค่าความถูกต้อง พบว่าแบบจำลองที่ดีที่สุดคือหน่วยความจำระยะยาว-ระยะสั้นมีความถูกต้องในการจำแนกสูงถึง 83.7% รองลงมาเป็น โครงข่ายประสาทแบบคอนโวลูชันมีประสิทธิภาพความถูกต้องในการจำแนก 77.0% ตามมาด้วยหน่วยเวียนกลับแบบมีประตู มีความถูกต้องในการจำแนก 65.4% ตามลำดับ ดังนั้นจึงสรุปได้ว่าแบบจำลองหน่วยความจำระยะยาว-ระยะสั้นมีความเหมาะสมและมีประสิทธิภาพที่จะนำไปสร้างระบบวิเคราะห์ความคิดเห็นอัตโนมัติอันเนื่องจากอัลกอริทึมชนิดนี้มีการพิจารณาบริบทของคำในภาษาไทย และถูกออกแบบมาสำหรับการประมวลผลแบบลำดับตลอดจนมีหน่วยความจำภายในตัวเอง (ชาลิสา จิตบุญญาพินิจ, 2022)

พิศิษฐ์ บวรเลิศสุธี (2022) ได้นำเสนอตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์โดยใช้เทคนิคส่วนการเรียนรู้ของเครื่อง (Machine Learning) ร่วมกับเทคนิคการตัดคำ (Word Tokenization) และการสร้างคลังคำศัพท์ (Bag of Words) จากนั้นนำเข้ากระบวนการจำแนกประเภทผลการวิเคราะห์ 4 เทคนิค ได้แก่ LSTM, SGD, Logistic Regressionและ Support Vector การประเมินตัวแบบทำการทดลองโดยใช้ข้อมูลตัวอย่างการแสดงความคิดเห็นต่อสินค้าและบริการออนไลน์ภาษาไทย จำนวน 12,900 ข้อมูล ใช้เป็นข้อมูลในการสร้างตัวแบบ เพื่อช่วยให้ผู้บริโภคใช้ในการตัดสินใจซื้อสินค้าและบริการ และช่วยให้ผู้ประกอบการมีข้อมูลสำหรับการพัฒนาสินค้าและบริการต่อไปในอนาคต จากผลการทดลองสร้างตัวแบบการวิเคราะห์ระดับความรู้สึกทางอารมณ์ 3 ระดับ คือ ความรู้สึกเชิงบวก (Positive) เป็นกลาง (Neutral) และเชิงลบ (Negative) โดยแต่ละตัวแบบให้ค่าความถูกต้องในการทำนายผล ดังนี้ LSTM 81.27%, Logistic Regression 69%, SGD 66%, และ Support Vector Machines 65% สรุปได้

ว่าการสร้างตัวแบบโดยใช้เทคนิคการจำแนกประเภทแบบ LSTM ให้คะแนนความถูกต้องสูงที่สุดในการทำนายผลด้วยการเรียนรู้เชิงลึกจึงเหมาะสมแก่การนำไปใช้สร้างตัวแบบในการวิเคราะห์ระดับความรู้สึกกับข้อความที่เป็นภาษาไทยโดยพิจารณาจากค่า F1 (พิศิษฐ์ บวรเลิศสุธี, 2022)

กม จารุเศรนี (2022) ได้ทำการวิจัยเพื่อศึกษาปัจจัยที่มีอิทธิพลต่อความสำเร็จของธุรกิจร้านอาหารขนาดกลางในเขตกรุงเทพมหานคร เพื่อศึกษารูปแบบความสัมพันธ์เชิงสาเหตุของปัจจัยที่มีอิทธิพลต่อความสำเร็จของธุรกิจร้านอาหารขนาดกลางในเขตกรุงเทพมหานครการวิจัยนี้ใช้ระเบียบวิธีการวิจัยแบบผสมผสาน โดยการวิจัยเชิงปริมาณใช้วิธีเก็บรวบรวมข้อมูลด้วยแบบสอบถามจากผู้ประกอบการธุรกิจร้านอาหารขนาดกลางในเขตกรุงเทพมหานครจำนวน 300 คน การวิจัยเชิงคุณภาพใช้วิธีการสัมภาษณ์เชิงลึกผู้ให้ข้อมูลสำคัญตัวแทนธุรกิจร้านอาหาร ผู้แทนหน่วยงานของกรุงเทพมหานคร รวม 30 คน การวิเคราะห์ข้อมูลเพื่อทดสอบสมมติฐานวิจัยใช้การวิเคราะห์สมการเชิงโครงสร้างโดยโปรแกรม AMOS ผลการวิจัยพบว่าปัจจัยที่มีอิทธิพลต่อความสำเร็จของธุรกิจร้านอาหารขนาดกลางมากที่สุดคือการตอบสนองต่อตลาด รองลงมาคือกระบวนการจัดการและการให้ความสำคัญกับตลาด ส่วนประสิทธิผลของการจัดการไม่มีอิทธิพลทางตรงต่อความสำเร็จของธุรกิจร้านอาหารขนาดกลาง โดยปัจจัยทั้งสามสามารถอธิบายความแปรปรวนของความสำเร็จของธุรกิจร้านอาหารขนาดกลาง (R2) ได้เท่ากับ 96% (กม จารุเศรนี, 2022)

จิราพัชร จุรุเทียบ (2022) ได้ทำการศึกษาระดับและความสัมพันธ์ของช่องทางโซเชียลมีเดียและกลยุทธ์การตลาดที่มีผลต่อการตั้งใจเข้าร้านอาหารในจังหวัดอุบลราชธานี โดยใช้แบบสอบถามจำนวน 16 ข้อเป็นเครื่องมือในการเก็บรวบรวมความเห็นของกลุ่มตัวอย่าง 400 คน สถิติที่ใช้วิเคราะห์ได้แก่ ความถี่ ร้อยละ ค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ความคลาดเคลื่อน One-Way ANOVA, MANOVA และ Correlation ผลวิจัยพบว่า ระดับค่าเฉลี่ยของกลยุทธ์การตลาดที่มีผลต่อการตั้งใจเข้าร้านอาหาร พบว่า โดยรวมมีค่าเฉลี่ยเท่ากับ 5.97 ซึ่งมีความคิดเห็นอยู่ในระดับมาก โดยปัจจัยกลยุทธ์ด้านสถานที่มีค่าเฉลี่ยความคิดเห็นสูงที่สุดเท่ากับ 4.07 นอกจากนี้ ปัจจัยที่มีผลต่อการตั้งใจเข้าร้านอาหาร ด้านกลยุทธ์การตลาดประกอบด้วย กลยุทธ์ด้านการสนับสนุนการขาย และกลยุทธ์ด้านสถานที่ พบว่าปัจจัยทั้ง 2 มีผลต่อการตั้งใจเข้าร้านอาหารอย่างมีนัยสำคัญทางสถิติที่ 0.05 และปัจจัยดังกล่าวสามารถร่วมพยากรณ์ปัจจัยที่มีผลต่อการตั้งใจเข้าร้านอาหารได้ร้อยละ (R = 0.663) (จิราพัชร จุรุเทียบ, 2022)

จักรรัตน์ กนะกาศัย (2022) ได้ทำการศึกษาปัจจัยเชิงสาเหตุด้านส่วนประสมการตลาดบริการ ด้านเทคโนโลยีสารสนเทศ และด้านการตลาดบนโซเชียล ที่ส่งผลต่อการรักษาฐานลูกค้า เพื่อศึกษาปัจจัยเชิงสาเหตุ ด้านส่วนประสมการตลาดบริการ ด้านเทคโนโลยีสารสนเทศ และด้านการตลาดบนโซเชียล ที่ส่งผลต่อการดำเนินงานธุรกิจที่ทันสมัยของธุรกิจร้านอาหารแบบดั้งเดิมเพื่อผลสำเร็จขององค์กร และเพื่อสร้างแบบจำลองระบบการรักษาฐานของลูกค้าของธุรกิจร้านอาหาร

แบบดั้งเดิมเพื่อผลสำเร็จขององค์กร กลุ่มตัวอย่างได้แก่ ผู้ประกอบการธุรกิจร้านอาหารในประเทศไทย จำนวน 544 ราย สุ่มตัวอย่างแบบแบ่งชั้น เครื่องมือที่ใช้ในการวิจัย คือ แบบสอบถาม วิเคราะห์ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน และทดสอบสมมติฐานด้วยการวิเคราะห์สมการโครงสร้างด้วยเทคนิควิธี Partial Least Square ผลการวิจัยพบว่า ปัจจัยส่วนประสมการตลาดบริการ ปัจจัยเทคโนโลยีสารสนเทศ ปัจจัยการตลาดโซเชียลมีเดีย ส่งผลต่อการรักษาฐานลูกค้า ในภาพรวมอยู่ในระดับมาก สำหรับการศึกษาการรักษาฐานลูกค้า และการดำเนินงานของธุรกิจที่ทันสมัย ในภาพรวมอยู่ในระดับมาก และแบบจำลองการสร้างระบบการรักษาฐานของลูกค้าของธุรกิจร้านอาหารแบบดั้งเดิมเพื่อผลสำเร็จขององค์กรพบว่าโมเดลแบบจำลองสมการเชิงโครงสร้างมีความเหมาะสมกลมกลืนกับข้อมูลเชิงประจักษ์ผลวิจัยชี้ให้เห็นถึงความสำคัญในการทำความเข้าใจถึงอิทธิพลและรูปแบบความสัมพันธ์เชิงสาเหตุของการสร้างระบบการรักษาฐานของลูกค้าของธุรกิจร้านอาหารแบบดั้งเดิมเพื่อผลสำเร็จขององค์กร (จักรรัตน์ กนะกาศัย, 2022)

Takashi Tanizaki (2019) ได้นำเสนอการพยากรณ์ความต้องการในร้านอาหารโดยใช้เทคนิคแมชชีนเลิร์นนิ่ง เช่น Bayesian Linear Regression, Boosted Decision Tree Regression, และ Decision Forest Regression โดยใช้ข้อมูลจากระบบขายหน้าร้าน (Point of sale system) แบบจำลองการพยากรณ์คำนึงถึงปัจจัยต่างๆ เช่น ที่ตั้งร้าน สภาพอากาศ เหตุการณ์ต่างๆ เป็นต้น ผลการวิจัยพบว่า ไม่มีความแตกต่างกันมากนักระหว่างแต่ละเทคนิคแมชชีนเลิร์นนิ่ง โดยเทคนิค Boosted Decision Tree Regression ให้ค่าความถูกต้องต่ำที่สุด โดยรวมแล้วทุกเทคนิคให้ค่าความถูกต้องประมาณ 85% (Takashi Tanizaki, 2019)

Leonard Gunawan (2023) ได้ทำการศึกษาความพึงพอใจของผู้มาใช้บริการร้านอาหารโดยการวิเคราะห์ความรู้สึกจากบทวิจารณ์ออนไลน์จาก TripAdvisor ว่าเป็นไปในเชิงบวก เชิงลบ หรือเป็นกลาง โดยใช้เทคนิค SVM (Support Vector Machine) เปรียบเทียบกับ Naive Bayes (NB) ผลการวิจัยบ่งชี้ว่า SVM มีความแม่นยำมากกว่าโดยมีค่าความถูกต้องที่ 79% (Leonard Gunawan, 2023)

Yi Luo (2021) ได้ทำการวิเคราะห์ความรู้สึกจากบทวิจารณ์ออนไลน์เพื่อทำความเข้าใจความคิดเห็นที่ผู้มารับประทานอาหารนำเสนอจากร้านอาหารจำนวน 112,412 ในช่วงการระบาดของ COVID-19 ระหว่างเดือนมกราคม-มิถุนายน 2020 บน Yelp.com โดยการเปรียบเทียบอัลกอริทึมต่าง ๆ ที่เกี่ยวข้องกับการเรียนรู้เชิงลึก (Deep learning) ได้แก่ Random Forest, GBDT, Simple Embedding + Average Pooling, Bidirectional LSTM ผลการวิจัยพบว่า Bidirectional LSTM ซึ่งพัฒนามาจากเทคนิคโครงข่ายประสาทเทียม (Neural Network) มีประสิทธิภาพดีที่สุดโดยมีค่าความถูกต้องโดยรวมที่ 92% (Yi Luo, 2021)

จากข้อมูลงานวิจัยที่เกี่ยวข้องข้างต้นสามารถนำมาสรุปเป็นตารางเปรียบเทียบแหล่งที่มาของข้อมูลเทคนิคที่ใช้และผลการวิจัยได้ตามตารางที่ 2

ตาราง 2 งานวิจัยที่เกี่ยวข้องด้วยเทคโนโลยีการเรียนรู้ของเครื่อง

ผู้วิจัย	แหล่งที่มาของข้อมูล	อัลกอริทึม	ค่าความถูกต้อง
(ณัฐวัฒน์ เอกธรรมนิตย์, 2021)	ชุดข้อมูลจำลอง (Synthetic Data) ที่อาศัยหลักการจำลองข้อมูลแบบแจกแจงปกติ (Normal Distribution)	Reinforcement Learning	82 %
ปราโมชย์ นามวงศ์ (2022)	เว็บไซต์เครือข่ายสังคมออนไลน์เฟสบุ๊ค (Facebook.com) จำนวน 56 ร้าน	Decision Tree	89 %
ชาลิสา จิตบุญญาพินิจ (2022)	ชุดข้อมูลการแสดงความคิดเห็นต่อสินค้าและบริการของร้านอาหารของเว็บไซต์ Wongnai.com	Neural Network	83 %
Mehrbakhsh Nilashi (2021)	ชุดข้อมูลการแสดงความคิดเห็นต่อสินค้าและบริการของร้านอาหารของเว็บไซต์ TripAdvisor	Decision Tree	93%
พิศิษฐ์ บวรเลิศสุธี (2022)	ชุดข้อมูลที่ให้เข้าถึงสาธารณะจากเว็บไซต์ https://pythainlp.github.io/ (Wisesight Sentiment Corpus)	Neural Network	81.27%
Xiaohui Chang (2019)	ชุดข้อมูลจากเว็บไซต์ค้นหาธุรกิจในท้องถิ่น yelp.com	Support Vector Machines	77%
Najwa AlGhamdi (2021)	ชุดข้อมูลรีวิวออนไลน์ ร้านอาหาร จากเว็บไซต์https://www.kaggle.com/hj5992/restaurantreviews?select=Restaurant_Reviews.tsv	Naive Bayes, Linear Regression, Decision Tree, K-Nearest Neighbor,	83%

ผู้วิจัย	แหล่งที่มาของข้อมูล	อัลกอริทึม	ค่าความถูกต้อง
		Support Vector Machine	
Takashi Tanizaki (2019)	ข้อมูลจากระบบขายหน้าร้าน (Point of sale system)	Bayesian Linear Regression, Boosted Decision Tree Regression, Decision Forest Regression	85%
Yi Luo (2021)	บทวิจารณ์ออนไลน์จากร้านอาหารจำนวน 112,412 ในช่วงการระบาดของ COVID-19 ระหว่างเดือนมกราคม-มิถุนายน 2020 บน Yelp.com	Random Forest, GBDT, Simple Embedding + Average Pooling, Bidirectional LSTM	92%
Leonard Gunawan (2023)	ชุดข้อมูลการแสดงความคิดเห็นต่อสินค้าและบริการของร้านอาหารจากเว็บไซต์ TripAdvisor	Support Vector Machine, Naive Bayes	79%

จากงานวิจัยที่เกี่ยวข้องข้างต้น Decision Tree, Linear Regression และ Neural Network เป็นเครื่องมือที่มีประสิทธิภาพและได้รับความนิยมอย่างกว้างขวางในงานวิจัยที่เกี่ยวข้องกับธุรกิจร้านอาหาร เช่น การค้นหาปัจจัยที่เกี่ยวข้องกับประสิทธิภาพ การพยากรณ์ความต้องการในร้านอาหาร รวมถึงการวิเคราะห์ความรู้สึกจากบทวิจารณ์ออนไลน์ผ่านสื่อโซเชียลคอมเมิร์ซต่าง ๆ จึงมีความเหมาะสมที่จะนำมาใช้ในงานวิจัยนี้ และจากงานวิจัยที่ผ่านมายังเป็นการวิจัยโดยใช้แหล่งข้อมูลจากแพลตฟอร์มใดแพลตฟอร์มหนึ่งเช่น Facebook.com TripAdvisor.com Yelp.com หรือ Wongnai.com แต่ยังไม่มีการเชื่อมโยงว่าธุรกิจร้านอาหารที่มีข้อมูลหรือช่องทางการประชาสัมพันธ์ในหลายแพลตฟอร์มจะสามารถส่งผลอย่างไรต่อประสิทธิภาพของธุรกิจร้านอาหาร จึงเป็นช่องว่างและโอกาสที่ผู้วิจัยนำมาศึกษาในวิทยานิพนธ์ฉบับนี้

บทที่ 3
วิธีดำเนินการวิจัย

งานวิจัยฉบับนี้เป็นงานวิจัยเชิงปริมาณ (Quantitative Research) เพื่อวิเคราะห์ข้อมูลด้วยเทคนิคแมชชีนเลิร์นนิ่ง โดยผู้วิจัยได้ทำการศึกษาวิจัยตามลำดับขั้นตอน ดังนี้

1. ประชากรและกลุ่มตัวอย่าง
2. เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล
3. การเก็บรวบรวมข้อมูล
4. การจัดกระทำกับข้อมูลและวิเคราะห์ข้อมูล
5. การทำความสะอาดข้อมูล
6. การแปลงข้อมูล
7. การวิเคราะห์ความคิดเห็น
8. การลดมิติข้อมูล
9. การแบ่งกลุ่มข้อมูล
10. การวิเคราะห์ข้อมูล
11. การวัดประสิทธิภาพของแบบจำลอง

ประชากรและกลุ่มตัวอย่าง

ร้านอาหารในเขตกรุงเทพมหานครซึ่งมีข้อมูลอยู่บนเว็บไซต์ wongnai.com และมี facebook fanpage จำนวน 1,750 ร้านค้า จากนั้นทำการตรวจสอบข้อมูลผิดพลาดและสูญหาย จำนวน 113 รายการ คงเหลือจำนวน 1,637 ร้านค้า และข้อมูลรีวิวจากผู้ใช้บริการร้านอาหารจากเว็บไซต์ wongnai.com จำนวน 45,119 รายการ

เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

1. รวบรวมข้อมูลจากแหล่งต่างๆ ลงในฐานข้อมูล
2. นำเข้าและวิเคราะห์ข้อมูลโดยใช้โปรแกรม RapidMiner

การเก็บรวบรวมข้อมูล

ผู้วิจัยได้ทำการศึกษาและเลือกใช้ข้อมูลทุติยภูมิจากแหล่งต่างๆ นำมาเป็นปัจจัยนำเข้าเพื่อใช้ในการวิเคราะห์ข้อมูล ดังนี้

ข้อมูลร้านอาหารที่รวบรวมจากเว็บไซต์ wongnai.com ที่อยู่ในเขตกรุงเทพมหานคร ได้แก่ ประเภทร้านอาหาร ช่วงราคา ที่จอดรถ จำนวนบทวิจารณ์ การรับบัตรเครดิต บริการ Wi-Fi มีโซเชียลมีเดียอื่นๆ เหมาะสำหรับมาเป็นกลุ่ม การมีบริการเครื่องดื่มแอลกอฮอล์ มีเว็บไซต์(Website) บริการดิลิเวอรี และคะแนนรีวิว

ข้อมูลจาก facebook fanpage ของร้านที่มีข้อมูลบนเว็บไซต์ wongnai.com ที่อยู่ในเขตกรุงเทพมหานคร ได้แก่ จำนวนเช็คอิน จำนวนผู้ถูกใจเพจ จำนวนผู้ติดตามเพจ

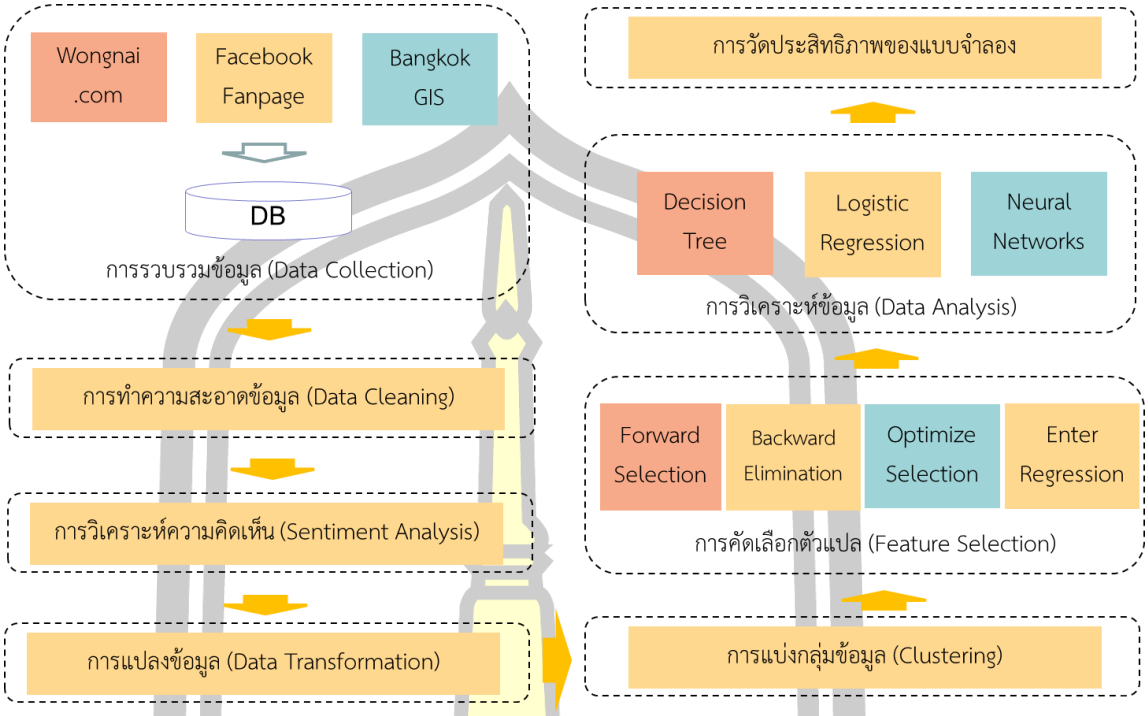
ข้อมูลความหนาแน่นของประชากรในพื้นที่จากศูนย์เทคโนโลยีสารสนเทศภูมิศาสตร์กรุงเทพมหานคร ซึ่งเป็นหน่วยงานที่มีหน้าที่ดูแลรับผิดชอบ วางแผน จัดการกิจกรรมต่าง ๆ เกี่ยวข้องกับการพัฒนาและส่งเสริมการปฏิบัติงานด้านเทคโนโลยีสารสนเทศภูมิศาสตร์ของกรุงเทพมหานคร จากนั้นทำการบันทึกข้อมูลลงในฐานข้อมูลเพื่อนำไปใช้ในขั้นตอนการเตรียมข้อมูลต่อไป

ระดับความคิดเห็น โดยใช้ข้อมูลรีวิวจากลูกค้าผู้ใช้บริการซึ่งจะถูกนำไปหาค่าระดับความคิดเห็นโดยใช้ Natural Language AI ซึ่งให้บริการโดย Google สำหรับวิเคราะห์ความรู้สึกด้วยภาษาธรรมชาติ (Natural Language) โดยใช้ AI

การจัดกระทำกับข้อมูลและการวิเคราะห์ข้อมูล

การจัดกระทำกับข้อมูลในงานวิจัยฉบับนี้เป็นการวิเคราะห์ข้อมูลด้วยเทคนิคแมชชีนเลิร์นนิ่ง โดยสามารถแสดงขั้นตอนการทำงานดังภาพประกอบ 8

พหุน ปณฺฑิโต ชีเว



ภาพประกอบ 8 ขั้นตอนการจัดกระทำกับข้อมูลและการวิเคราะห์ข้อมูล

การทำความสะอาดข้อมูล (Data Cleaning)

โดยการนำข้อมูลที่ได้ทำการตรวจสอบและขจัดหรือแทนค่าข้อมูลที่มีความผิดปกติ (Outliers) เช่น ค่าสูญหาย (missing value) ซึ่งเกิดจากข้อมูลร้านอาหารบางร้านที่รวบรวมจากเว็บไซต์แสดงข้อมูลไม่ครบถ้วนในทุกปัจจัย ในกรณีนี้จะแทนค่าปัจจัยที่สูญหายด้วย 0

การแปลงข้อมูล (Data Transformation)

ผู้วิจัยได้ทำการแปลงข้อมูลให้อยู่ในรูปแบบที่สามารถนำมาวิเคราะห์ข้อมูลด้วยเทคนิคแมชชีนเลิร์นนิ่งได้ ดังแสดงในตาราง 3

ตาราง 3 เกณฑ์การแปลงข้อมูล

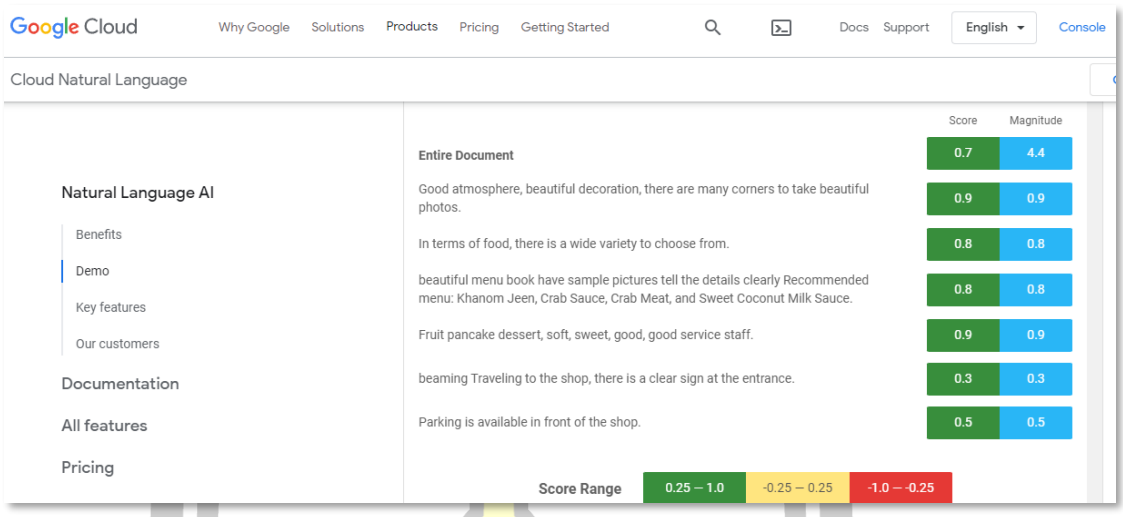
Attribute	Description
RestaurantType	ประเภทร้านอาหาร
Price	ช่วงราคา ต่ำกว่า 100 = 1; 101-250 = 2; 251-500 = 3; 501-1000 = 4; มากกว่า 1000 = 5
Seats	จำนวนที่นั่ง 10 = 1; 40 = 2; 80 = 3; 150 = 4
Parking	ที่จอดรถ ไม่มี = 1; มี = 2
Delivery	บริการดิลิเวอรี ไม่มี = 1; มี = 2
CreditCard	รับบัตรเครดิต ไม่มี = 1; มี = 2
Wi-Fi	มีบริการ Wi-Fi ไม่มี = 1; มี = 2
Animal	สัตว์เลี้ยงเข้าได้ ไม่ได้ = 1; ได้ = 2
Group	เหมาะสำหรับมาเป็นกลุ่ม ไม่ใช่ = 1; ใช่ = 2
Alcohol	มีบริการเครื่องดื่มแอลกอฮอล์ ไม่มี = 1; มี = 2
SocialMedia	มีโซเชียลมีเดียอื่นๆ ไม่มี = 1; มี = 2
PopulationDensity	ความหนาแน่นของประชากรในพื้นที่ น้อยกว่า 50,000 = 1; 50,000–100,000 = 2; 100,001– 150,000 = 3; 150,001–200,000 = 4; มากกว่า 200,000 = 5
Reviews	จำนวนบทวิจารณ์ ต่ำกว่า 50 = 1; 51-100 = 2; 101-150 = 3; 151-200 = 4;

Attribute	Description
	มากกว่า 200 = 5
Sentiment	ความคิดเห็น -1.0 – -0.25 = 1; -0.26- 0.24 = 2; 0.25 – 1.0 = 3
Check-in	จำนวนเช็คอิน น้อยกว่า 5,000 = 1; 5,000–10,000 = 2; 10,001–15,000 = 3; 15,001–20,000 = 4; มากกว่า 20,000 = 5
Like	จำนวนผู้ถูกใจเพจ น้อยกว่า 25,000 = 1; 25,000–50,000 = 2; 50,001–75,000 = 3; 75,001–100,000 = 4; มากกว่า 100,000 = 5
Followers	จำนวนผู้ติดตามเพจ น้อยกว่า 25,000 = 1; 25,000–50,000 = 2; 50,001–75,000 = 3; 75,001–100,000 = 4; มากกว่า 100,000 = 5
RatingScore	คะแนนรีวิว 0 – 3.9 = (ไม่ดี); มากกว่า 3.9 = 1 (ดี)

การวิเคราะห์ความคิดเห็น (sentiment analysis)

เป็นการวิเคราะห์ระดับความคิดเห็นและนำข้อมูลดังกล่าวมาคำนวณเพื่อให้ได้ค่าร้อยละของจำนวนความคิดเห็นทั้งหมดจากร้านเดียวกันได้แก่ ร้อยละของจำนวนบทวิจารณ์เชิงบวก ร้อยละของจำนวนบทวิจารณ์เชิงลบ ร้อยละของจำนวนบทวิจารณ์ที่เป็นกลาง ซึ่งจำเป็นต้องอาศัยเครื่องมือในการสังเคราะห์ข้อมูล ดังภาพประกอบ 9

พหุน ปณ ฑิโต ชีเว



ที่มา : https://cloud.google.com/natural-language

ภาพประกอบ 9 เว็บไซต์ Cloud Natural Language ของบริษัท Google

จากภาพประกอบ 9 เป็นการแสดงกระบวนการวิเคราะห์ความคิดเห็นจากบทวิจารณ์ออนไลน์ซึ่งเป็นการนำบทวิจารณ์ออนไลน์ ซึ่งลูกค้าที่มาใช้บริการร้านอาหารเป็นผู้เขียนแสดงความคิดเห็นลงบนเว็บไซต์ wongnai ของแต่ละร้าน นำมาวิเคราะห์ความคิดเห็นโดยใช้ Natural Language AI ซึ่งเป็นบริการคลาวด์ของบริษัท Google ที่สามารถวิเคราะห์ข้อความที่ไม่มีโครงสร้างโดยการทำความเข้าใจภาษาธรรมชาติ (NUL) โดยใช้แมชชีนเลิร์นนิง ซึ่งมักจะใช้ในการวิเคราะห์ความรู้สึกเพื่อทำความเข้าใจความคิดเห็นของลูกค้าเพื่อให้ได้ข้อมูลเชิงลึกที่สามารถนำไปใช้ในการพัฒนาผลิตภัณฑ์และบริการของธุรกิจ โดยจะได้ผลลัพธ์เป็นระดับคะแนนซึ่งสามารถแปลความหมายได้ ดังนี้ 1) -1.0 - -0.25 คะแนน เป็นความคิดเห็นเชิงลบ 2) -0.25 - 0.25 เป็นความคิดเห็นที่เป็นกลาง 3) 0.25 - 1.0 คะแนน เป็นความคิดเห็นเชิงบวก

การลดมิติข้อมูล (Dimensionality Reduction)

เพื่อลดขนาดข้อมูลโดยยังรักษาคุณภาพข้อมูลที่เหมาะสมที่สุด และเพื่อเป็นการระบุปัจจัยที่ส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซมากที่สุด เพื่อที่จะนำมาใช้ในการวิเคราะห์ด้วยเทคนิค Evolutionary Selection ซึ่งมีกระบวนการทำงานโดยการคำนวณค่าน้ำหนักของปัจจัยที่กำหนด โดยเมื่อน้ำหนักของปัจจัยที่นำมาคำนวณสูงเท่าใด ก็จะถือว่ามีความเกี่ยวข้องมากขึ้นเท่านั้น หากปัจจัยที่นำมาคำนวณส่งผลให้นำหนักสูงขึ้นจะเก็บไว้ แต่หากปัจจัยที่นำมาคำนวณ

ส่งผลให้ค่าน้ำหนักต่ำลงทำการจะตัดปัจจัยนั้นออกจากกระบวนการวิเคราะห์ข้อมูล ซึ่งวิธีการที่นำมาใช้ในการวิจัยนี้ได้แก่ การเลือกตัวแปรโดยวิธีเพิ่มตัวแปร (Forward Selection) การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward Elimination) และการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

การแบ่งกลุ่มข้อมูล (Clustering)

ธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซมีหลากหลายประเภท จากข้อมูลตัวอย่างที่นำมาใช้ในงานวิจัยพบว่ามีประเภทร้านอาหารทั้งหมด 23 ประเภท ผู้วิจัยจึงทำการแบ่งกลุ่มประเภทร้านอาหารออกเป็น 3 กลุ่ม โดยใช้ราคาเฉลี่ยซึ่งเป็นคุณลักษณะที่เด่นชัดของร้านอาหารแต่ละประเภท ตัวอย่างเช่นเช่น อาหารตามสั่ง และสตรีทฟู้ด มักจะมีราคาเฉลี่ยไม่เกิน 100 บาท ส่วนร้านอาหารญี่ปุ่นจะมีราคาเฉลี่ยมากกว่า 1,000 บาท เป็นต้น เพื่อให้ทราบว่าแต่ละกลุ่มประเภทมีปัจจัยที่ส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารที่แตกต่างกันอย่างไร โดยผู้วิจัยได้เลือกใช้เทคนิค K-means Clustering ในการแบ่งกลุ่ม

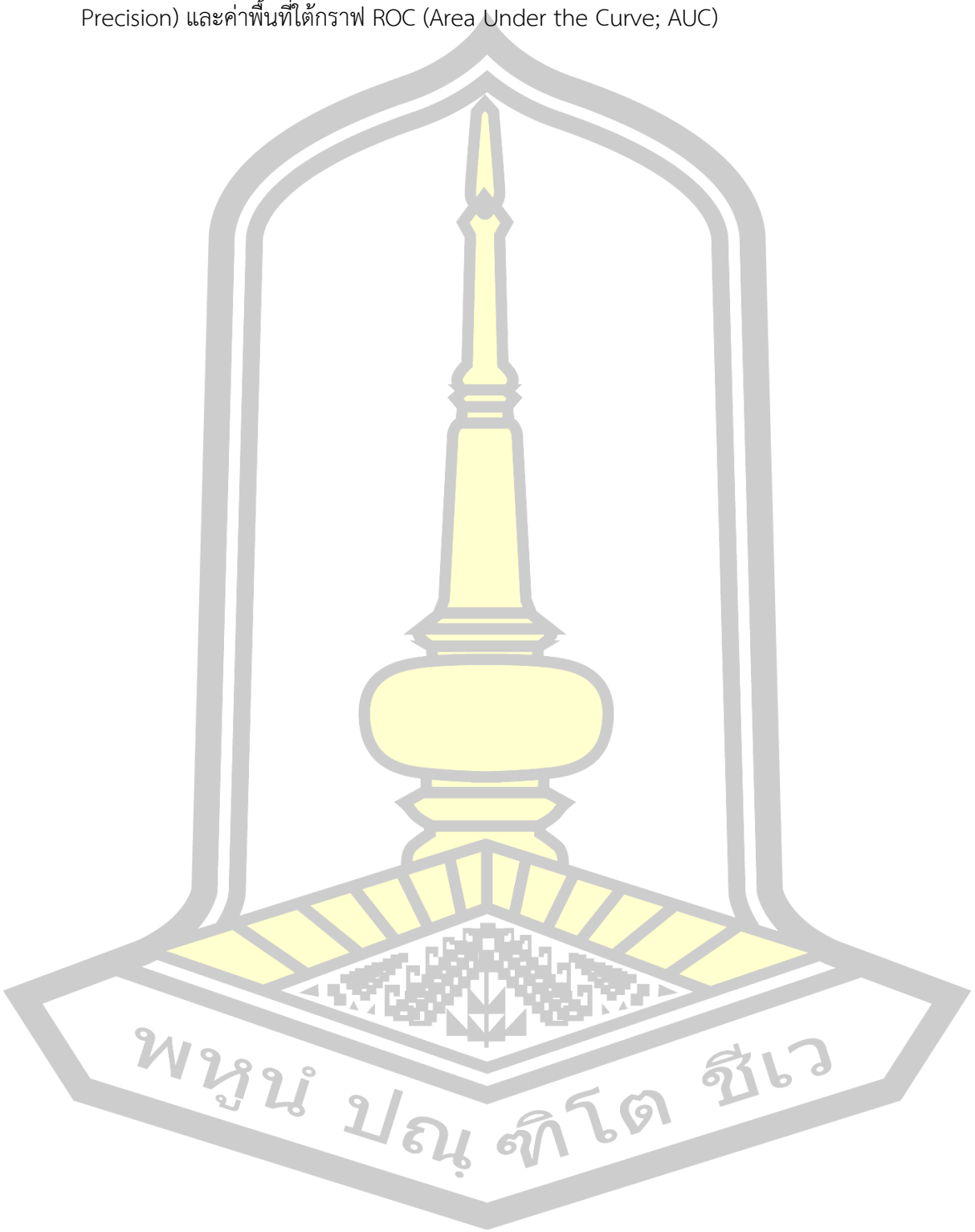
การวิเคราะห์ข้อมูล (Data Analysis)

งานวิจัยนี้มุ่งเน้นการเปรียบเทียบประสิทธิภาพความแม่นยำของตัวแบบการพยากรณ์ (Prediction Model) ที่สร้างขึ้นโดยใช้อัลกอริธึมแมชชีนเลิร์นนิ่งด้วยโปรแกรม RapidMiner ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) โครงข่ายประสาทเทียม (Neural Networks) และการวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis) ซึ่งจากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าเป็นอัลกอริธึมที่มีความแม่นยำสำหรับการสร้างตัวแบบพยากรณ์ที่มีประสิทธิภาพสูง โดยทำการแบ่งข้อมูลเพื่อใช้ในการทดสอบออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลเรียนรู้ (Training Data set) และทำการแบ่งข้อมูลที่เหลือสำหรับใช้เป็นข้อมูลในการทดสอบ (Testing Data set) โดยใช้วิธี K-fold Cross validation

การวัดประสิทธิภาพของแบบจำลอง (Method Evaluation)

ทำการวัดผลและเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์โดยการพิจารณาค่าความถูกต้อง (Accuracy) การวัดประสิทธิภาพโดยรวม (F-measure) ค่าความไวในการตรวจพบค่าบวก

(Sensitivity) และค่าความจำเพาะในการตรวจพบค่าลบ (Specificity) ค่าความแม่นยำ (Class Precision) และค่าพื้นที่ใต้กราฟ ROC (Area Under the Curve; AUC)



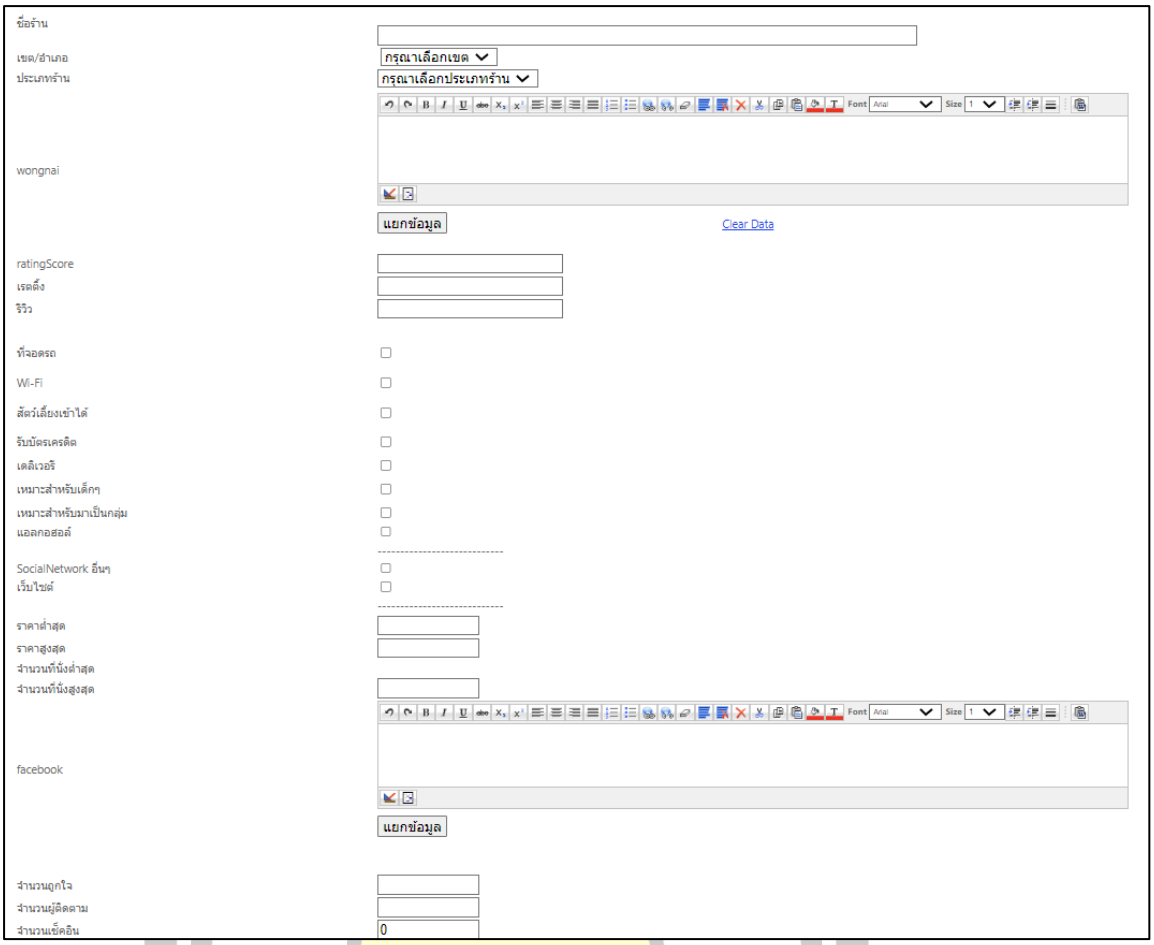
บทที่ 4
ผลการวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลงานวิจัย เรื่อง การพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซโดยมีจุดมุ่งหมายเพื่อศึกษาปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ เปรียบเทียบประสิทธิภาพอัลกอริธึมในการพยากรณ์ของการทำเหมืองข้อมูลสำหรับธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ เพื่อพัฒนาแบบจำลองสำหรับการวิเคราะห์ปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ โดยผู้วิจัยได้ทำการนำเข้าข้อมูลร้านอาหารบนเว็บไซต์ wongnai ที่อยู่ในเขตกรุงเทพมหานคร และ facebook fanpage ของแต่ละร้านจำนวนทั้งสิ้น 1,750 ร้านค้า จากนั้นทำการตรวจสอบข้อมูลผิดพลาดและสูญหายจำนวน 113 รายการ คงเหลือจำนวน 1,637 ร้านค้า

การนำเข้าข้อมูล

ผู้วิจัยได้พัฒนาโปรแกรมสำหรับรวบรวมข้อมูลร้านอาหารบนเว็บไซต์ wongnai ที่อยู่ในเขตกรุงเทพมหานคร รวมถึงข้อมูลจาก facebook fanpage ของร้านที่มีข้อมูลบนเว็บไซต์ wongnai โดยใช้ภาษา C# ในการพัฒนา โดยโปรแกรมจะทำการแยกข้อมูลจากเว็บไซต์ออกตามปัจจัยต่างๆ ตามข้อมูลที่ปรากฏในหน้าเว็บไซต์ จากนั้นทำการบันทึกข้อมูลลงในฐานข้อมูล MS SQLServer ตัวอย่างหน้าจอโปรแกรมสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai สามารถแสดงได้ดังภาพประกอบ 10

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 10 ตัวอย่างหน้าจอโปรแกรมสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai

ตัวอย่างโค๊ดโปรแกรมภาษา C# ที่ใช้ในการพัฒนาระบบสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai สามารถแสดงได้ดังภาพประกอบ 11

พหุน ปณ ฑิโต ชีเว

```
    using(DataClassesDataContext dc = new DataClassesDataContext())
    {
        if (DropDownList1.SelectedItem.Text != "กรุณาเลือกเขต" && ddlResType.SelectedItem.Text != "กรุณาเลือกประเภทร้าน"){
            Info i = new Info();
            if (Request.QueryString["edit"] != null){
                var iList = from t in dc.Infos where t.InfoID == Convert.ToInt32(Request.QueryString["edit"].ToString()) select t;
                foreach (var _i in iList){
                    i = _i;
                }
            }
            i.name = txtName.Text.Trim();
            i.Location = DropDownList1.SelectedItem.Text;
            i.ratingScore = txtRatingScore.Text.Trim();
            i.ratingCount = txtRatingCount.Text.Trim();
            i.reviewCount = int.Parse(txtReviewCount.Text.Trim());
            i.restaurantType = ddlResType.SelectedValue;
            i.parking = (ckbParking.Checked) ? "Yes" : "No";
            if(Request.QueryString["edit"]==null){
                dc.Infos.InsertOnSubmit(i);
    hasWebSite(txtContentText.Text);
    GetFactor(txtContentText.Text);
}

private void GetResName(string strInfo)
{
    using(DataClassesDataContext dc = new DataClassesDataContext())
    {
        int fp = strInfo.IndexOf(">") + 1;
        while (strInfo[fp] != '>')
        {
            fp++;
        }
        fp++;

        int lp = strInfo.IndexOf("</span>");
        txtName.Text = strInfo.Substring(fp,lp-fp);
        var l = from t in dc.Infos where t.name == txtName.Text select t;
        foreach(var i in l)
        {
```

ภาพประกอบ 11 ตัวอย่างโค๊ดโปรแกรมสำหรับบันทึกข้อมูลร้านอาหารบนเว็บไซต์ wongnai

ตัวอย่างหน้าจอข้อมูลร้านอาหารที่ทำการบันทึกไว้ในฐานข้อมูลสามารถแสดงได้ดังภาพประกอบ 12

Location	minPrice	maxPrice	rating Score	rating Count	review Count	restaurant Type	parking	WiFi	animal	credit Card	delivery	forChildren	forGroup	alcohol	maxSeats	SocialNetwork	hasWebSite	NumberofLike	NumberofFollowers	NumberofCheckin
วัฒนา	1,000		3.4	35	22	ok	0	1	0	1	1	0	0	1	40	1	0	4500	4700	0
บางกอกน้อย	1,000		4.4	1101	714	ok	1	0	0	1	1	1	1	1	150	1	0	377,083	395,426	170,037
ทวีวัฒนา	251	500	4.2	657	486	ok	1	1	1	1	1	1	1	1	150	0	1	84,106	85,170	308,051
ปทุมวัน	1,000		3.9	196	153	ok	1	1	0	1	1	1	1	1	80	1	1	4,166	4,300	0
บางรัก	1,000		4.4	129	68	ok	1	1	0	1	1	0	1	1	80	1	1	0	0	0
คลองเตย	501	1,000	4.3	166	116	ok	1	1	0	1	1	0	1	1	150	0	1	23,929	25,329	35,837
คลองเตย	1,000		4.1	219	168	ok	1	0	1	1	1	1	1	1	150	0	1	64,262	65,466	33,892
ปทุมวัน	1,000		4.3	160	117	ok	1	1	0	1	0	1	1	1	80	0	1	50,349	52,815	589,170
บางรัก	251	500	4.4	202	154	ok	0	0	0	0	1	0	0	0	40	0	0	24,288	25,420	13,399
ปทุมวัน	251	500	4.1	341	280	ok	1	1	0	0	1	1	0	0	10	1	0	14,717	14,840	399
จตุจักร	501	1,000	4.2	404	270	ok	1	1	1	1	1	1	1	1	150	1	0	3,204	3,320	4,155
ป้อมปราบฯ	251	500	3.9	546	413	ok	0	0	0	1	1	0	1	0	150	0	0	3,065	3,274	6,787
พระนคร	101	250	3.7	839	653	ok	0	0	0	0	1	0	0	0	40	1	0	60,000	62,000	0
สาทร	1,000		4.4	455	270	ok	1	1	0	1	1	1	1	1	150	1	1	23,840	24,558	52,343
จตุจักร	251	500	3.9	409	273	ok	1	1	1	1	1	1	1	1	150	0	0	21,516	22,268	21,838
พระนคร	101	250	3.9	326	248	ok	0	0	0	0	1	0	0	1	40	0	0	797	842	2,183
ป้อมปราบฯ	501	1,000	4.0	309	201	ok	0	0	0	0	0	0	0	0	40	0	0			
ลาดพร้าว	101	250	4.3	146	100	ok	0	1	0	1	1	0	0	1	40	0	0	1,007	1,071	2,295
ปทุมวัน	501	1,000	4.2	115	85	ok	1	1	0	1	1	0	0	1	40	0	0	18,611	18,917	21,037
ดุสิต	101	250	4.0	307	235	ok	0	0	0	1	1	0	0	0	40	1	0			
ดุสิต	251	500	4.1	721	553	ok	1	0	0	1	1	0	0	1	150	1	1	34,604	35,002	2,583
อ.เมืองปทุมธานี	251	500	3.6	211	153	ok	1	1	0	1	0	1	1	1	150	0	0	6,800	7,000	0
บางรัก	251	500	4.0	273	218	ok	1	0	0	1	1	0	0	1	40	0	1	7,100	7,500	0

ภาพประกอบ 12 ตัวอย่างหน้าจอข้อมูลร้านอาหาร

ข้อมูลรีวิวจากลูกค้าจะถูกนำไปหาค่าระดับความคิดเห็นโดยใช้ Natural Language AI ซึ่งให้บริการโดย Google สำหรับวิเคราะห์ความรู้สึกด้วยภาษาธรรมชาติ (Natural Language) โดยใช้ AI และนำข้อมูลที่ได้บันทึกลงในฐานข้อมูลเพื่อหาค่าเฉลี่ยของระดับความคิดเห็นทั้งหมดแบ่งตามร้านอาหาร หน้าจอ Natural Language AI สามารถแสดงได้ดังภาพประกอบ 13

พหุน ปณ ฑิโต ชีเว

ที่มา : https://cloud.google.com/natural-language

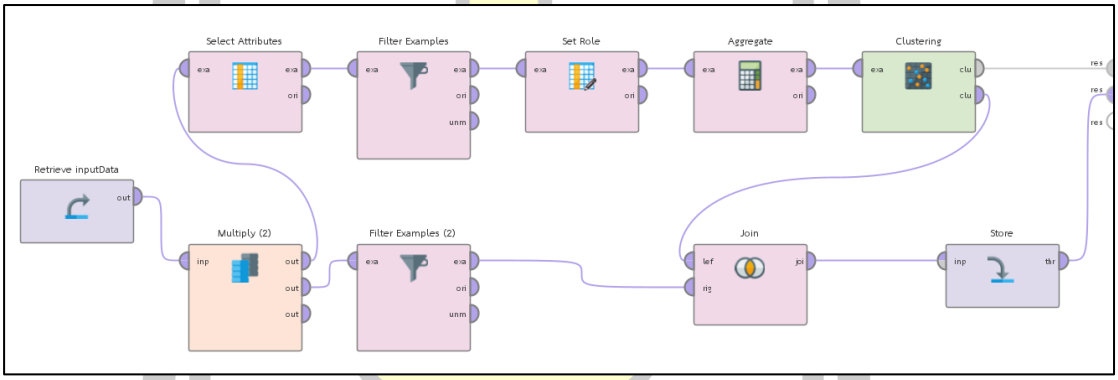
ภาพประกอบ 13 หน้าจอ Natural Language AI

ตัวอย่างหน้าจอข้อมูลคะแนนรีวิวที่ผ่านขั้นตอนการวิเคราะห์ความรู้สึกด้วยภาษาธรรมชาติ (Natural Language) โดยใช้ AI สามารถแสดงได้ดังภาพประกอบ 14

ภาพประกอบ 14 ตัวอย่างหน้าจอข้อมูลคะแนนรีวิว

การแบ่งกลุ่มข้อมูล (Clustering)

ธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซมีหลากหลายประเภท จากข้อมูลตัวอย่างที่นำมาใช้ในงานวิจัยพบว่ามีประเภทร้านอาหารทั้งหมด 23 ประเภท ผู้วิจัยจึงทำการแบ่งกลุ่มประเภทร้านอาหารออกเป็น 3 กลุ่ม โดยใช้ราคาเฉลี่ยซึ่งเป็นคุณลักษณะที่เด่นชัดของร้านอาหารแต่ละประเภท ตัวอย่างเช่นเช่น อาหารตามสั่ง และสตรีทฟู้ด มักจะมีราคาเฉลี่ยไม่เกิน 100 บาท ส่วนร้านอาหารญี่ปุ่นจะมีราคาเฉลี่ยมากกว่า 1,000 บาท เป็นต้น เพื่อให้ทราบว่าแต่ละกลุ่มประเภทมีปัจจัยที่ส่งผลต่อประสิทธิภาพของธุรกิจร้านอาหารที่แตกต่างกันอย่างไร โดยผู้วิจัยได้เลือกใช้เทคนิค K-means Clustering ในการแบ่งกลุ่ม ขั้นตอนการแบ่งกลุ่มประเภทร้านอาหารตามค่าเฉลี่ยของราคา สามารถแสดงได้ดังภาพประกอบ 15



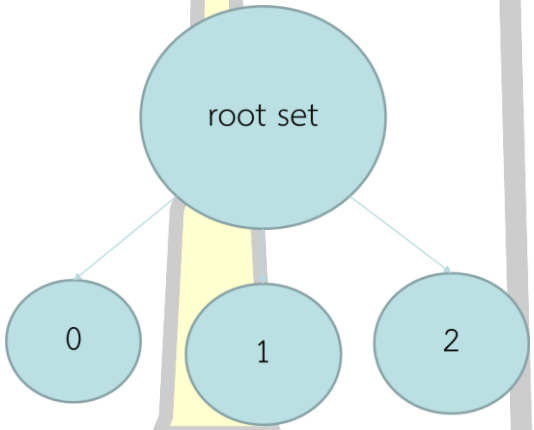
ภาพประกอบ 15 ขั้นตอนการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering

1. Retrieve Data : เชื่อมต่อกับชุดข้อมูลนำเข้า
2. Multiply : ทำการแยกข้อมูลเพื่อเข้าสู่กระบวนการแบ่งกลุ่มข้อมูลและการรวมข้อมูลหลังจากการแบ่งข้อมูลเสร็จสิ้น
3. Filter Examples : ทำการเลือกเฉพาะข้อมูลที่ไม่มีค่าสูญหาย (Missing Value)
4. Set Role : กำหนดให้ประเภทร้านอาหาร (RestaurantTypeData) เป็นคีย์หลัก (id)
5. Aggregate : หาค่าเฉลี่ยของช่วงราคาเพื่อนำไปใช้ในกระบวนการแบ่งกลุ่มด้วย K-means clustering
6. Clustering : ทำการแบ่งกลุ่มข้อมูลโดยใช้เทคนิค K-means กำหนดค่าพารามิเตอร์ k = 3

7. Join : ทำการรวมข้อมูลที่ได้จากกระบวนการ Clustering เข้ากับข้อมูลนำเข้าโดยกำหนดให้ RestaurantTypeData เป็นคีย์แอตตริบิวต์ที่ใช้ในการรวมข้อมูล

8. Store : ทำการบันทึกข้อมูลที่ได้ในชื่อ ClusterData เพื่อใช้เป็นข้อมูลนำเข้าสำหรับกระบวนการสร้างตัวแบบพยากรณ์ ด้วยเทคนิคการเรียนรู้ของเครื่องในขั้นตอนต่อไป

ผลการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering สามารถแสดงได้ดังภาพประกอบ 16



ภาพประกอบ 16 ผลการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering

จากภาพประกอบ 16 พบว่า Cluster_1 มีช่วงราคาสูงที่สุด โดยมีจำนวนประเภทร้านอาหาร 8 ประเภท คิดเป็น 34.78% รองลงมาคือ Cluster_2 ซึ่งมีช่วงราคาอยู่ในระดับปานกลาง โดยมีจำนวนประเภทร้านอาหาร 8 ประเภท คิดเป็น 34.78% และ Cluster_0 มีช่วงราคาต่ำที่สุด โดยมีจำนวนประเภทร้านอาหาร 7 ประเภท คิดเป็น 30.43% โดยมีรายละเอียดข้อมูลดังตาราง 4

ตาราง 4 ข้อมูลการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering

Cluster	ประเภทร้านอาหาร	ค่าเฉลี่ยช่วงราคา	ระดับราคา
Cluster_0	ก๋วยเตี๋ยว ร้านกาแฟ/ชา ร้านข้าวต้ม สตรีทฟู้ด/รถเข็น อาหารกล่อง/ข้าวกล่อง อาหารจานเดียว อาหารตามสั่ง	1.557	ต่ำ

Cluster	ประเภทร้านอาหาร	ค่าเฉลี่ยช่วงราคา	ระดับราคา
Cluster_1	ชาบู/สุกี้ยากี้/หม้อไฟ ซูชิ บุฟเฟ่ต์ ปิ้งย่าง อาหารญี่ปุ่น อาหารทะเล อาหารฟิวชั่น อิซากายะ	3.527	สูง
Cluster_2	ราเมง สเต๊ก อาหารจีน อาหารไทย อาหารอีสาน อาหารนานาชาติ เบเกอรี่/เค้ก คาเฟ่	2.55	ปานกลาง

ตัวอย่างข้อมูลที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูลด้วยเทคนิค K-means Clustering สามารถแสดงได้ดังภาพประกอบ 17

RestaurantTypeDa...	RatingScore	cluster	average(Price)	Price	Se...	Reviews	Check-in	Like	PopulationD...	Sentiment	delivery	Group	alcohol	parking	creditCard	Website	SocialMedia	WiFi
ก๋วยเตี๋ยว	1	cluster_0	1.557	2	4	4	1	1	1	1	2	2	2	2	2	1	1	2
ก๋วยเตี๋ยว	1	cluster_0	1.557	1	4	1	1	1	1	1	2	2	1	2	1	1	2	2
ก๋วยเตี๋ยว	1	cluster_0	1.557	1	4	1	4	1	1	1	2	2	1	2	1	1	1	1
ก๋วยเตี๋ยว	1	cluster_0	1.557	1	4	3	1	1	5	1	2	1	1	2	2	1	1	2
คาเฟ่	2	cluster_2	2.321	3	4	2	1	1	1	1	2	1	2	2	2	2	2	2
คาเฟ่	2	cluster_2	2.321	2	4	1	4	1	4	3	2	2	2	2	2	1	2	2
ชาบู/สุกี้ยากี้/หม้อไฟ	2	cluster_1	3.510	5	4	1	1	2	3	1	2	2	2	2	2	1	2	2
ชาบู/สุกี้ยากี้/หม้อไฟ	1	cluster_1	3.510	3	4	5	1	2	3	1	2	2	2	2	2	1	2	2
ชาบู/สุกี้ยากี้/หม้อไฟ	1	cluster_1	3.510	3	4	1	1	1	1	1	2	2	2	2	1	1	1	2
ชาบู/สุกี้ยากี้/หม้อไฟ	1	cluster_1	3.510	3	4	2	1	1	1	1	2	1	2	2	2	1	1	1
ชาบู/สุกี้ยากี้/หม้อไฟ	1	cluster_1	3.510	3	4	2	1	2	3	1	2	2	2	2	2	1	2	2
ชาบู/สุกี้ยากี้/หม้อไฟ	2	cluster_1	3.510	3	4	1	3	1	3	3	2	2	2	1	1	2	2	2
ชาบู/สุกี้ยากี้/หม้อไฟ	2	cluster_1	3.510	2	4	2	1	1	3	3	2	1	2	1	1	1	1	1

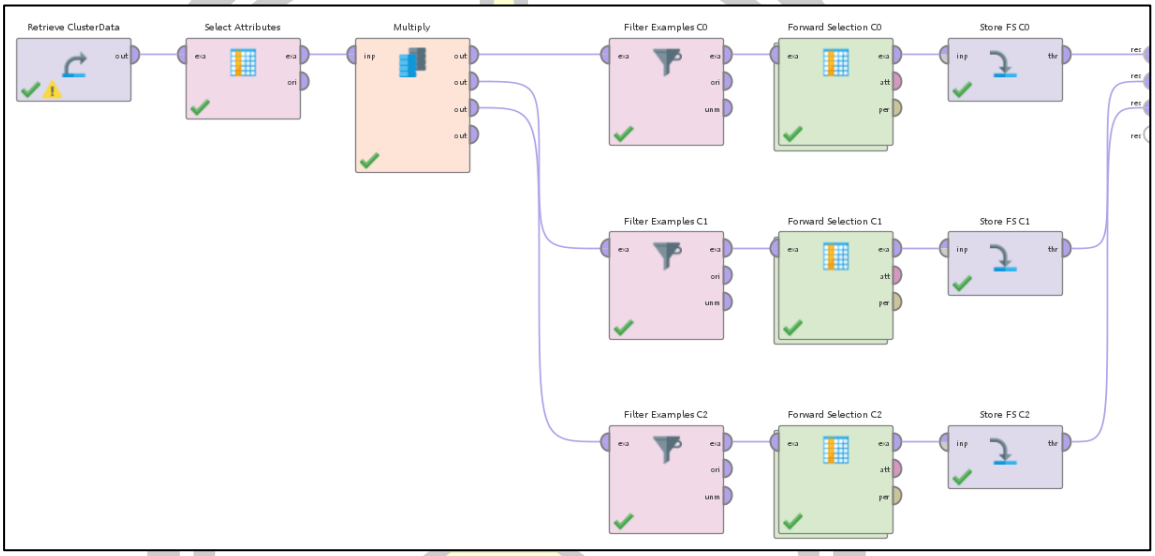
ภาพประกอบ 17 ตัวอย่างข้อมูลที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูลด้วยเทคนิค K-means Clustering

การคัดเลือกตัวแปร (Feature Selection)

งานวิจัยนี้ใช้วิธีการคัดเลือกตัวแปรเพื่อใช้ในการสร้างตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารทั้ง 3 กลุ่มซึ่งได้จากขั้นตอนการแบ่งกลุ่มประเภทร้านอาหารด้วยเทคนิค K-means Clustering ที่ผ่านมา 3 วิธี ได้แก่ การคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) การคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (backward elimination) และการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)

1. ผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection)

ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) สามารถแสดงได้ดังภาพประกอบ 18



ภาพประกอบ 18 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection)

1. Retrieve ClusterData : เชื่อมต่อกับชุดข้อมูล ClusterData ที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูล (Clustering)

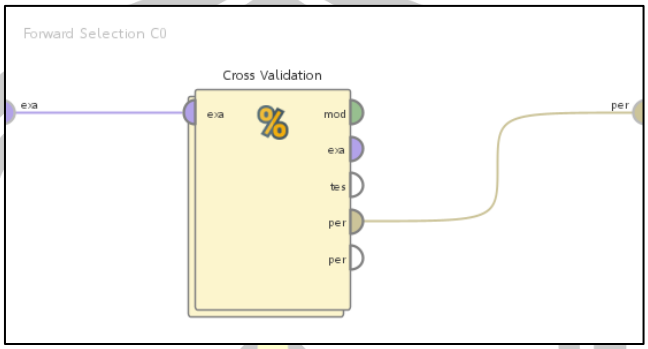
2. Select Attributes : ทำการนำตัวแปร average(Price) ซึ่งได้จากขั้นตอนการแบ่งกลุ่มข้อมูลออกก่อนนำเข้ากระบวนการเลือกตัวแปร

3. Multiply : ทำการแยกข้อมูลเพื่อเข้าสู่กระบวนการคัดเลือกตัวแปรตามแต่ละ Cluster

4. Filter Examples : ทำการเลือกเฉพาะข้อมูลที่ตรงกับ Cluster ที่กำหนด เช่น Filter Examples C0 เป็นการเลือกเฉพาะข้อมูลที่อยู่ใน Cluster_0 โดยการเพิ่ม filter จากนั้นกำหนดค่าเป็น cluster contain cluster_0 เป็นต้น

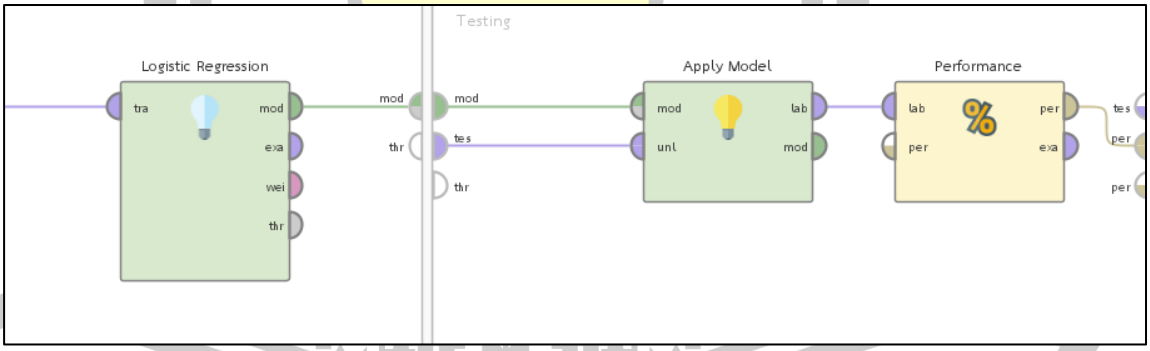
5. Forward Selection : ทำการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรเข้ามาทีละตัวแปร โดยในแต่ละรอบจะทำการทดสอบประสิทธิภาพ เมื่อตัวแปรที่เพิ่มเข้ามาทำให้ประสิทธิภาพเพิ่มขึ้นจะทำการเก็บตัวแปรนั้นไว้ และเพิ่มตัวแปรอื่นๆ เข้ามาจนกระทั่งประสิทธิภาพไม่สามารถเพิ่มขึ้นอีก

ขั้นตอนการทดสอบประสิทธิภาพภายใน Forward Selection โดยใช้ Cross Validation สามารถแสดงได้ดังภาพประกอบ 19



ภาพประกอบ 19 ขั้นตอนการทดสอบประสิทธิภาพภายใน Forward Selection โดยใช้ Cross Validation

5.1 Cross Validation : ภายใน Forward Selection ทำการประเมินประสิทธิภาพด้วยวิธี K-Fold Cross Validation โดยกำหนดพารามิเตอร์ number of folds = 10 ดังภาพประกอบ 20



ภาพประกอบ 20 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือกตัวแปรด้วยวิธี Forward Selection

5.1.1 Logistic Regression : เป็นการสร้างตัวแบบพยากรณ์เพื่อทดสอบประสิทธิภาพโดยใช้การถดถอยโลจิสติกส์

5.1.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing dataset)

5.1.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบที่ได้สร้างขึ้น

6. Store : บันทึกข้อมูลที่ได้เพื่อนำไปใช้ในขั้นตอนการสร้างตัวแบบพยากรณ์ชื่อ FS_Cluster_0_Data FS_Cluster_1_Data และ FS_Cluster_2_Data ตามลำดับ

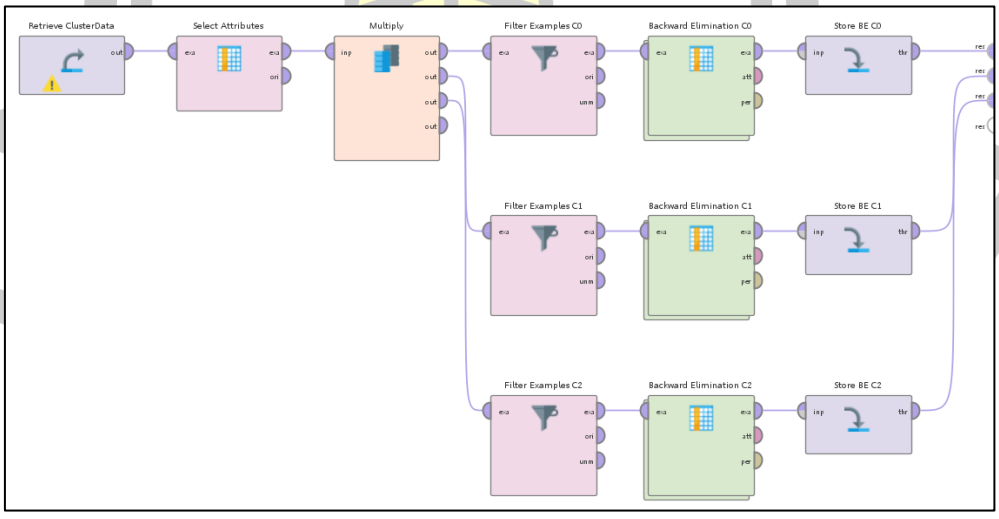
จากผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 3 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) จำนวนที่นั่ง(seats) และบริการดิลิเวอรี่(delivery)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 5 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) ช่วงราคา(price) มีบริการเครื่องดื่มแอลกอฮอล์(alcohol) จำนวนบทวิจารณ์(reviews) และเหมาะสำหรับมาเป็นกลุ่ม (group)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 1 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment)

2. ผลการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination)

ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) สามารถแสดงได้ดังภาพประกอบ 21



ภาพประกอบ 21 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination)

1. Retrieve ClusterData : เชื่อมต่อกับชุดข้อมูล ClusterData ที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูล (Clustering)

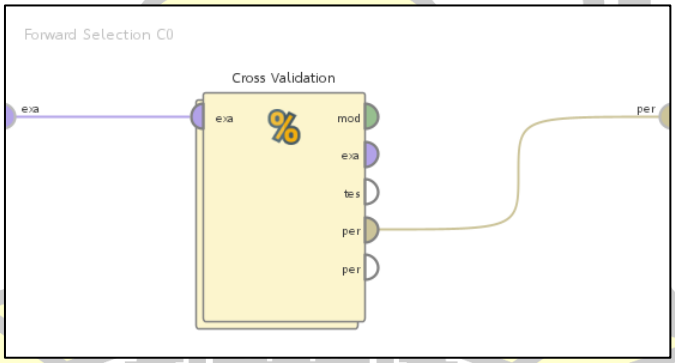
2. Select Attributes : ทำการนำตัวแปร average(Price) ซึ่งได้จากขั้นตอนการแบ่งกลุ่มข้อมูลออกก่อนนำเข้ากระบวนการเลือกตัวแปร

3. Multiply : ทำการแยกข้อมูลเพื่อเข้าสู่กระบวนการคัดเลือกตัวแปรตามแต่ละ Cluster

4. Filter Examples : ทำการเลือกเฉพาะข้อมูลที่ตรงกับ Cluster ที่กำหนด เช่น Filter Examples C0 เป็นการเลือกเฉพาะข้อมูลที่อยู่ใน Cluster_0 โดยการเพิ่ม filter จากนั้นกำหนดค่าเป็น cluster contain cluster_0 เป็นต้น

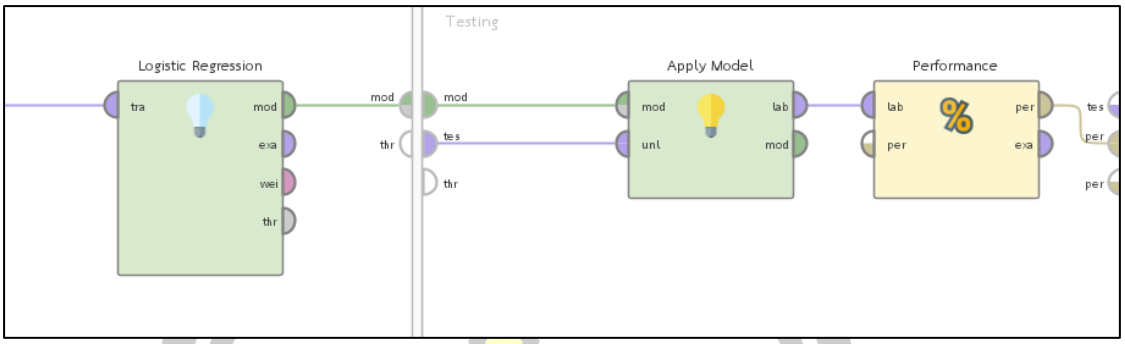
5. Backward Elimination : ทำการคัดเลือกตัวแปรด้วยวิธีนำเข้าตัวแปรทั้งหมด จากนั้นจึงนำตัวแปรออกทีละตัวแปร โดยในแต่ละรอบจะทำการทดสอบประสิทธิภาพ เมื่อตัวแปรที่นำออกทำให้ประสิทธิภาพเพิ่มขึ้นจะทำการนำตัวแปรอื่นๆ ออกไปจนกระทั่งประสิทธิภาพไม่สามารถเพิ่มขึ้นอีก

ขั้นตอนการทดสอบประสิทธิภาพภายใน Backward Elimination โดยใช้ Cross Validation สามารถแสดงได้ดังภาพประกอบ 22



ภาพประกอบ 22 ขั้นตอนการทดสอบประสิทธิภาพภายใน Backward Elimination โดยใช้ Cross Validation

5.1 Cross Validation : ภายใน Backward Elimination ทำการประเมินประสิทธิภาพด้วยวิธี K-Fold Cross Validation โดยกำหนดพารามิเตอร์ number of folds = 10 ดังภาพประกอบ 23



ภาพประกอบ 23 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือดตัวแปรด้วยวิธี Backward Elimination

5.1.1 Logistic Regression : เป็นการสร้างตัวแบบพยากรณ์เพื่อทดสอบประสิทธิภาพโดยใช้การถดถอยโลจิสติกส์

5.1.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing dataset)

5.1.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบที่ได้สร้างขึ้น

6. Store : บันทึกข้อมูลที่ได้เพื่อนำไปใช้ในขั้นตอนการสร้างตัวแบบพยากรณ์ชื่อ BE_Cluster_0_Data BE_Cluster_1_Data และ BE_Cluster_2_Data ตามลำดับ

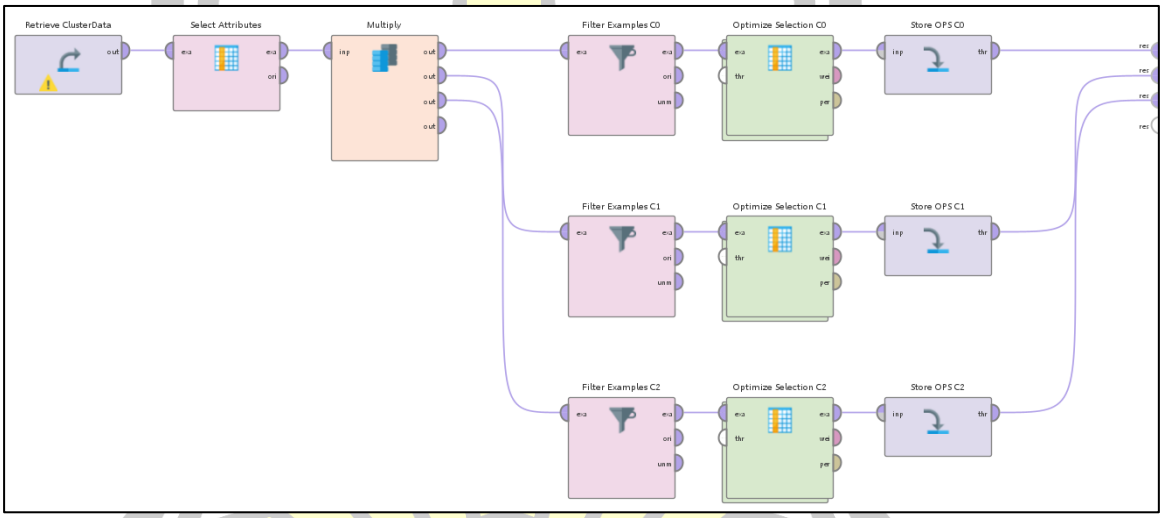
จากผลการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 12ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) ช่วงราคา(price) จำนวนที่นั่ง(seats) จำนวนบทวิจารณ์(reviews) จำนวนเช็คอิน(check-in) จำนวนผู้ถูกใจเพจ(like) ความหนาแน่นของประชากรในพื้นที่(populationDensity) บริการดิลิเวอรี(delivery) เหมาะสำหรับมาเป็นกลุ่ม(group) มีที่จอดรถ(parking) มีเว็บไซต์(Website) มีบริการ Wi-Fi(Wi-Fi)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 14 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) ช่วงราคา(price) จำนวนบทวิจารณ์(reviews) จำนวนเช็คอิน(check-in) จำนวนผู้ถูกใจเพจ(like) บริการดิลิเวอรี (delivery) เหมาะสำหรับมาเป็นกลุ่ม(group) มีบริการเครื่องดื่มแอลกอฮอล์(alcohol) มีที่จอดรถ (parking) รับบัตรเครดิต(creditCard) มีเว็บไซต์(Website) มีโซเชียลมีเดียอื่นๆ(SocialMedia) มีบริการ Wi-Fi(Wi-Fi)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 12 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) ช่วงราคา(price) จำนวนที่นั่ง(seats) จำนวนบทวิจารณ์(reviews) จำนวนเช็คอิน(check-in) ความหนาแน่นของประชากรในพื้นที่(populationDensity) บริการดิลิเวอรี่(delivery) มีบริการเครื่องดื่มแอลกอฮอล์(alcohol) มีที่จอดรถ(parking) รับบัตรเครดิต(creditCard) มีเว็บไซต์ (Website) มีโซเชียลมีเดียอื่นๆ (SocialMedia)

3. ผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) สามารถแสดงได้ดังภาพประกอบ 24



ภาพประกอบ 24 ขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

1. Retrieve ClusterData : เชื่อมต่อกับชุดข้อมูล ClusterData ที่ได้จากขั้นตอนการแบ่งกลุ่มข้อมูล (Clustering)

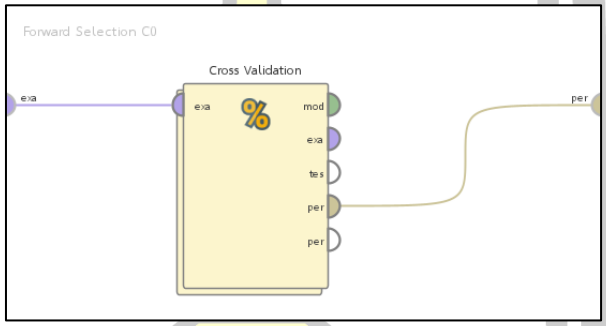
2. Select Attributes : ทำการนำตัวแปร average(Price) ซึ่งได้จากขั้นตอนการแบ่งกลุ่มข้อมูลออกก่อนนำเข้ากระบวนการเลือกตัวแปร

3. Multiply : ทำการแยกข้อมูลเพื่อเข้าสู่กระบวนการคัดเลือกตัวแปรตามแต่ละ Cluster

4. Filter Examples : ทำการเลือกเฉพาะข้อมูลที่ตรงกับ Cluster ที่กำหนด เช่น Filter Examples C0 เป็นการเลือกเฉพาะข้อมูลที่อยู่ใน Cluster_0 โดยการเพิ่ม filter จากนั้นกำหนดค่าเป็น cluster contain cluster_0 เป็นต้น

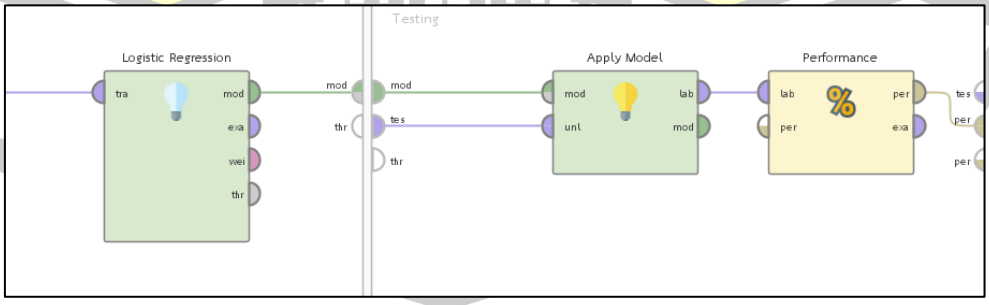
5. Optimize Selection : ทำการคัดเลือกตัวแปรด้วยวิธีสุ่มตัวแปรนำเข้าทีละตัวแปร โดยในแต่ละรอบจะทำการทดสอบประสิทธิภาพ เมื่อตัวแปรที่นำเข้าทำให้ประสิทธิภาพเพิ่มขึ้นจะทำการสุ่มตัวแปรอื่นๆ เข้ามาจนกระทั่งประสิทธิภาพไม่สามารถเพิ่มขึ้นอีก

ขั้นตอนการทดสอบประสิทธิภาพภายใน Optimize Selection โดยใช้ Cross Validation สามารถแสดงได้ดังภาพประกอบ 25



ภาพประกอบ 25 ขั้นตอนการทดสอบประสิทธิภาพภายใน Optimize Selection โดยใช้ Cross Validation

5.1 Cross Validation : ภายใน Optimize Selection ทำการประเมินประสิทธิภาพด้วยวิธี K-Fold Cross Validation โดยกำหนดพารามิเตอร์ number of folds = 10 ดังภาพประกอบ 26



ภาพประกอบ 26 ขั้นตอนการสร้างตัวแบบเพื่อทดสอบประสิทธิภาพภายใน Cross Validation เพื่อคัดเลือดตัวแปรด้วยวิธี Optimize Selection

5.1.1 Logistic Regression : เป็นการสร้างตัวแบบพยากรณ์เพื่อทดสอบประสิทธิภาพโดยใช้การถดถอยโลจิสติกส์

5.1.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing dataset)

5.1.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบที่ได้สร้างขึ้น

6. Store : บันทึกข้อมูลที่ได้เพื่อนำไปใช้ในขั้นตอนการสร้างตัวแบบพยากรณ์ชื่อ OPS_Cluster_0_Data OPS_Cluster_1_Data และ OPS_Cluster_2_Data ตามลำดับ

จากผลการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 3 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) จำนวนที่นั่ง(seats) และบริการดิลิเวอรี(delivery)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 5 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment) ช่วงราคา(price) มีบริการเครื่องดื่มแอลกอฮอล์(alcohol) จำนวนบทวิจารณ์(reviews) และเหมาะสำหรับมาเป็นกลุ่ม (group)

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่า มีตัวแปรที่ผ่านเกณฑ์คัดเลือกจำนวน 1 ตัวแปร ได้แก่ ระดับความคิดเห็น(sentiment)

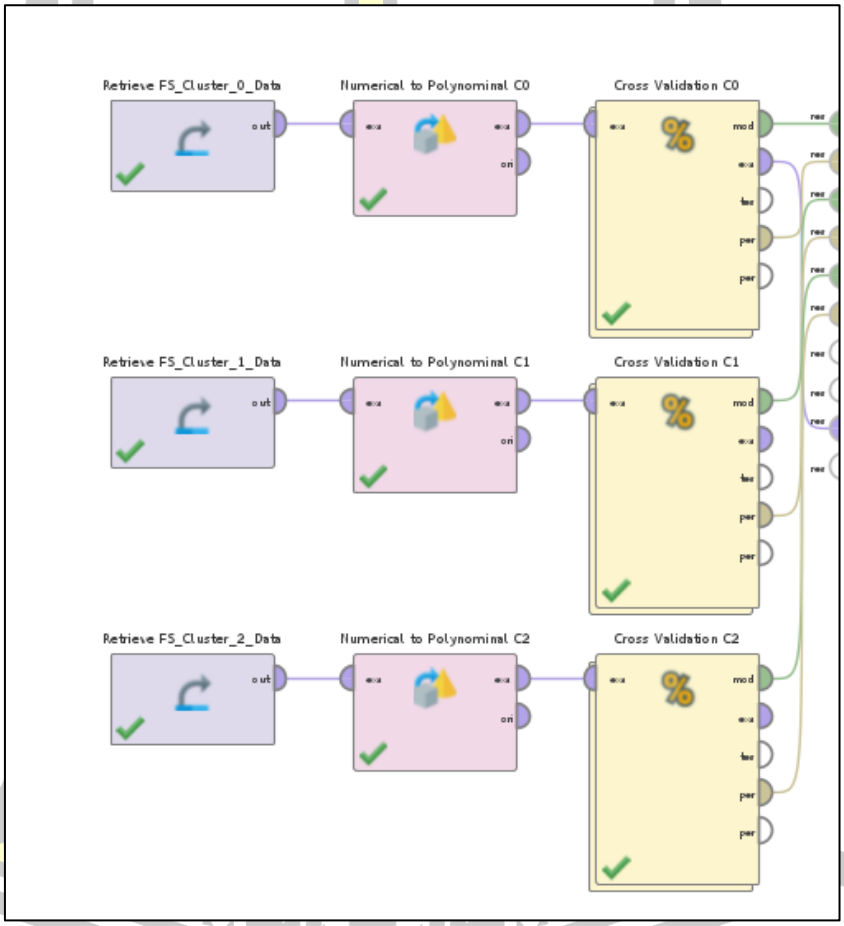
การสร้างตัวแบบพยากรณ์

งานวิจัยนี้เป็นการสร้างตัวแบบพยากรณ์ 3 ชนิดด้วยโปรแกรม RapidMiner โดยแต่ละชนิดจะทำการวัดประสิทธิภาพด้วยวิธี Cross-validation โดยการแบ่งข้อมูลออกเป็น 10 ส่วน โดยที่แต่ละส่วนมีจำนวนข้อมูลเท่ากัน (10-fold cross-validation) คัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด วิธีเพิ่มตัวแปร (forward selection) วิธีลดตัวแปร (Backward Elimination) และวิธีเพิ่มและลดตัวแปร (Optimize Selection) จากนั้นทำการเปรียบเทียบประสิทธิภาพระหว่างตัวแบบพยากรณ์ที่ดีที่สุดทั้ง 3 ชนิดซึ่งมีผลการวิจัยดังนี้

1. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree)

1.1 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) แสดงดังภาพประกอบ 27



ภาพประกอบ 27 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection)

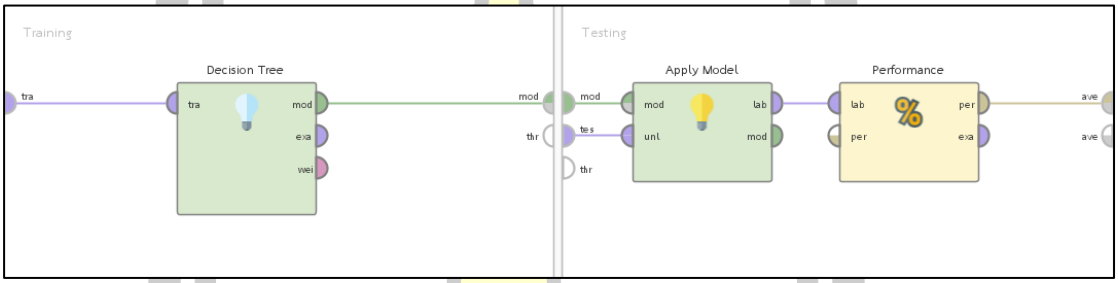
จากภาพประกอบ 26 แสดงขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) โดยมีขั้นตอน ดังนี้

1. Retrieve Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) ทั้ง 3 กลุ่ม

2. Numerical to Polynominal : ทำการแปลงข้อมูล Numerical ให้อยู่ในรูปแบบของ Polynominal

3. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Decision Tree ที่สร้างขึ้น

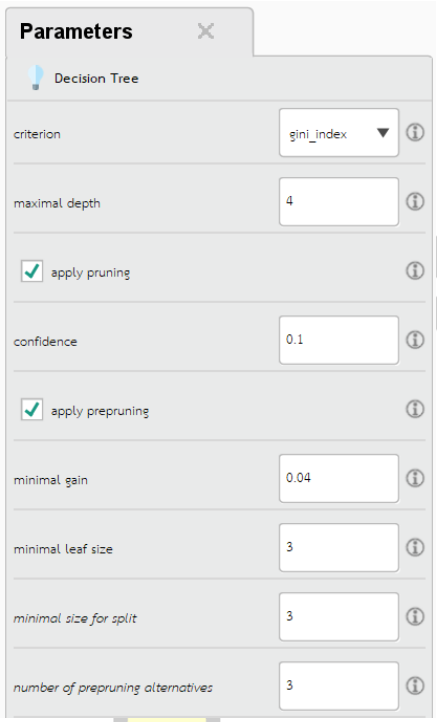
ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection) สามารถแสดงได้ดังภาพประกอบ 28



ภาพประกอบ 28 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection)

3.1 Decision Tree : เป็นการสร้างตัวแบบพยากรณ์ Decision Tree จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ดังภาพประกอบ 29

พหุน ปณ ฑิโต ชีเว

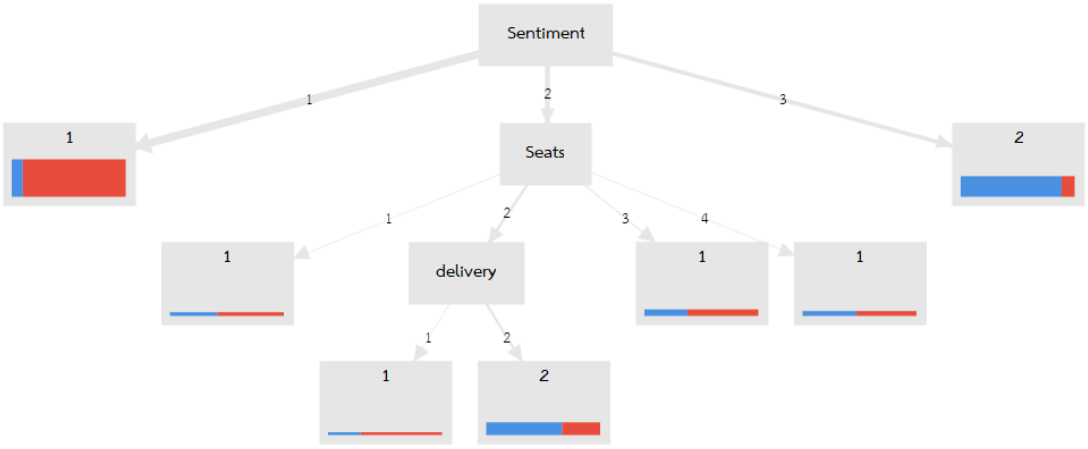


ภาพประกอบ 29 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีเพิ่มตัวแปร (forward selection)

3.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

3.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Decision Tree ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังภาพประกอบ 30



ภาพประกอบ 30 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) ซึ่งมีความคิดเห็นจากบทวิจารณ์ (Sentiment) เป็นโหนดรากสามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นจะพิจารณาจำนวนที่นั่ง(seats) เป็นลำดับถัดมา และการมีบริการดิลิเวอรี(delivery) หรือไม่จะถูกพิจารณาเป็นลำดับสุดท้าย

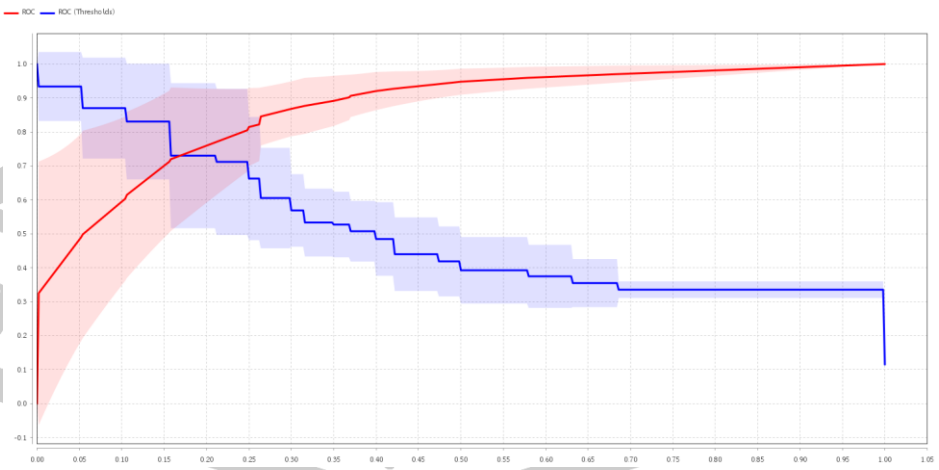
ผู้วิจัยได้ใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.22% ตามรายละเอียดที่แสดงไว้ในตาราง 5

พหุน ปณ ฑิโต ชีเว

ตาราง 5 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

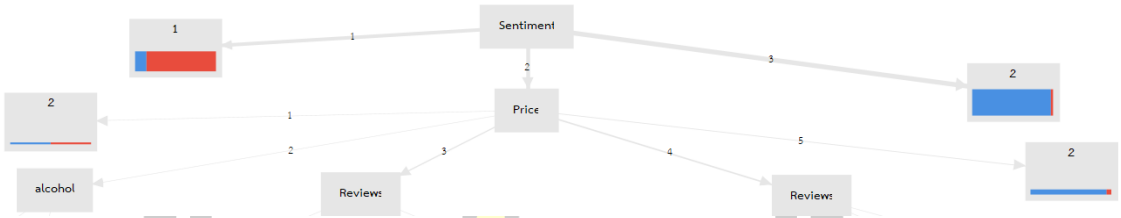
ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	145	44	76.72%
	คะแนนรีวิว อยู่ในระดับต่ำ	47	224	82.66%
	class recall	75.52%	83.58%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.55 ค่าความจำเพาะ (specificity) ร้อยละ 75.61 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.87 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 77.29 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.05 และมีค่า AUC เท่ากับ 0.869 ดังแสดงในภาพประกอบ 31



ภาพประกอบ 31 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังภาพประกอบ 32



ภาพประกอบ 32 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

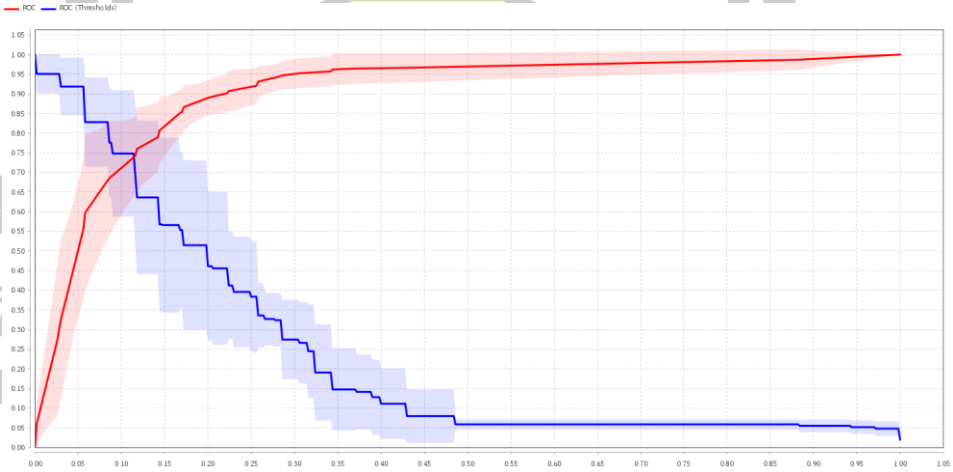
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) ซึ่งมีความคิดเห็นจากบทวิจารณ์ (Sentiment) เป็นโหนดรากสามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นจะพิจารณาช่วงราคา(price) เป็นลำดับถัดมา และจำนวนบทวิจารณ์(reviews) และการมีบริการเครื่องดื่มแอลกอฮอล์(alcohol) จะถูกพิจารณาเป็นลำดับสุดท้าย

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 83.96% ตามรายละเอียดที่แสดงไว้ในตาราง 6

ตาราง 6 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	311	52	85.67%
	คะแนนรีวิว อยู่ในระดับต่ำ	38	160	80.81%
	class recall	89.11%	75.47%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 75.50 ค่าความจำเพาะ (specificity) ร้อยละ 89.15 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 81.19 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 85.90 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 77.87 และมีค่า AUC เท่ากับ 0.903 ดังแสดงในภาพประกอบ 33



ภาพประกอบ 33 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังภาพประกอบ 34

ภาพประกอบ 34 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) ซึ่งมีความคิดเห็นจากบทวิจารณ์ (Sentiment) เป็นโหนดราก

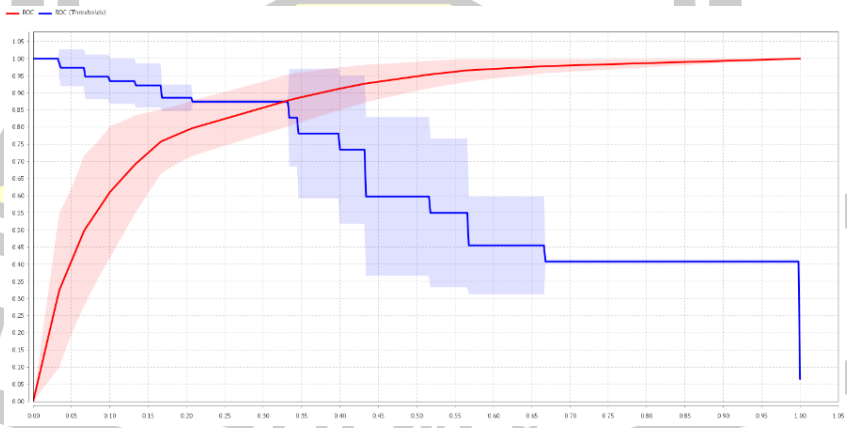
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มตัวแปร (forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.91% ตามรายละเอียดที่แสดงไว้ในตาราง 7

พหุน ปณฺฑิโต ชีเว

ตาราง 7 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	264	84	75.86%
	คะแนนรีวิว อยู่ในระดับต่ำ	34	236	87.41%
	class recall	88.59%	73.75%	

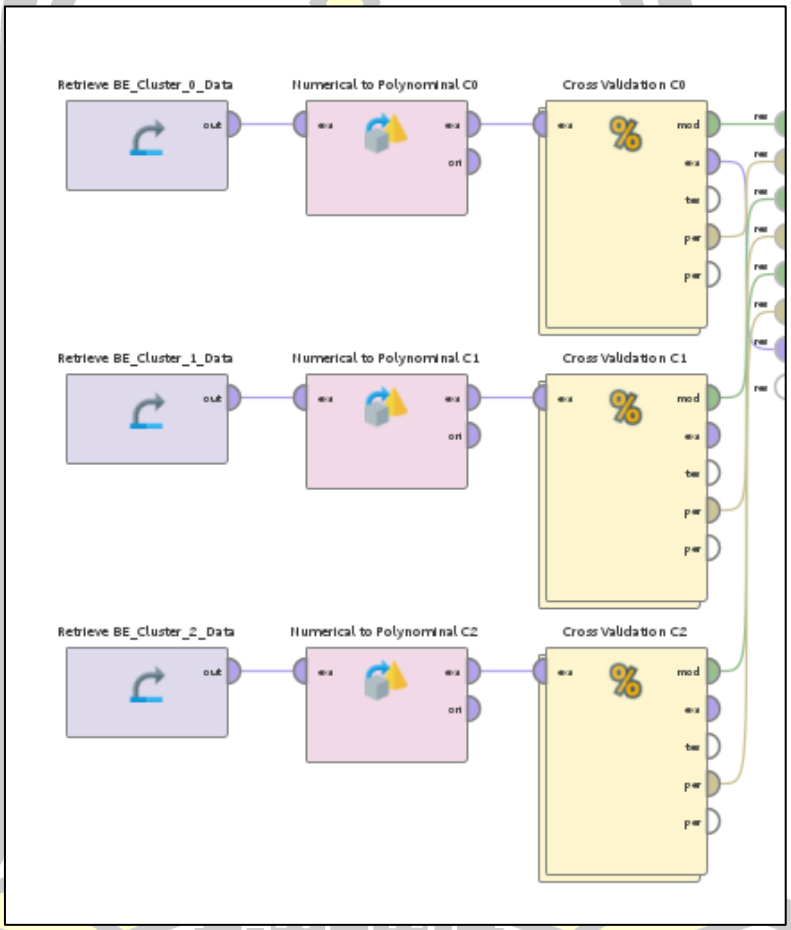
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.75 ค่าความจำเพาะ (specificity) ร้อยละ 88.59 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 87.64 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 76.14 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.89 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 35



ภาพประกอบ 35 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

1.2 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination)

ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) แสดงดังภาพประกอบ 36



ภาพประกอบ 36 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination)

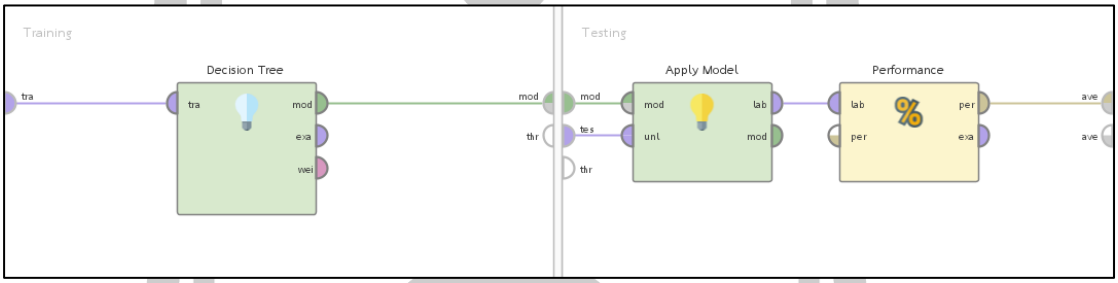
จากภาพประกอบ 35 แสดงขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) โดยมีขั้นตอน ดังนี้

1. Retrieve Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) ทั้ง 3 กลุ่ม

2. Numerical to Polynominal : ทำการแปลงข้อมูล Numerical ให้อยู่ในรูปแบบของ Polynominal

3. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Decision Tree ที่สร้างขึ้น

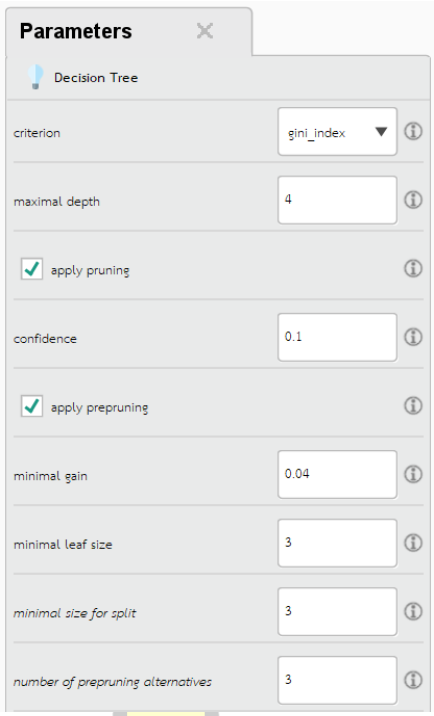
ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination) สามารถแสดงได้ดังภาพประกอบ 37



ภาพประกอบ 37 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination)

3.1 Decision Tree : เป็นการสร้างตัวแบบพยากรณ์ Decision Tree จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ดังภาพประกอบ 38

พหุน ปณ ฑิโต ชีเว

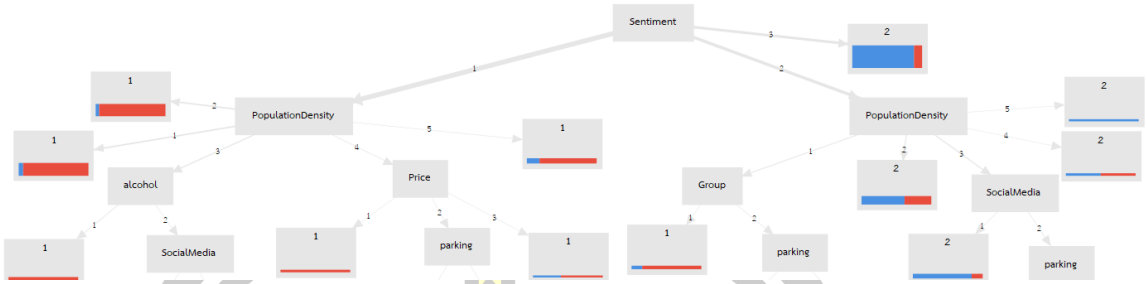


ภาพประกอบ 38 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีลดตัวแปร(Backward Elimination)

3.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

3.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Decision Tree ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังภาพประกอบ 39



ภาพประกอบ 39 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) ซึ่งมีความคิดเห็นจากบทวิจารณ์ (Sentiment) เป็นโหนดรากสามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นความหนาแน่นของประชากรในพื้นที่ (populationDensity) จะเป็นตัวแปรที่มีผลต่อการพิจารณาเป็นลำดับถัดมา จากนั้น ช่วงราคา (price) การมีโซเชียลมีเดียอื่นๆ(SocialMedia) การมีบริการเครื่องดื่มแอลกอฮอล์(alcohol) และเหมาะสำหรับมาเป็นกลุ่ม(group) จะถูกพิจารณาเป็นลำดับถัดมา และมีที่จอดรถ(parking) จะถูกพิจารณาเป็นลำดับสุดท้าย

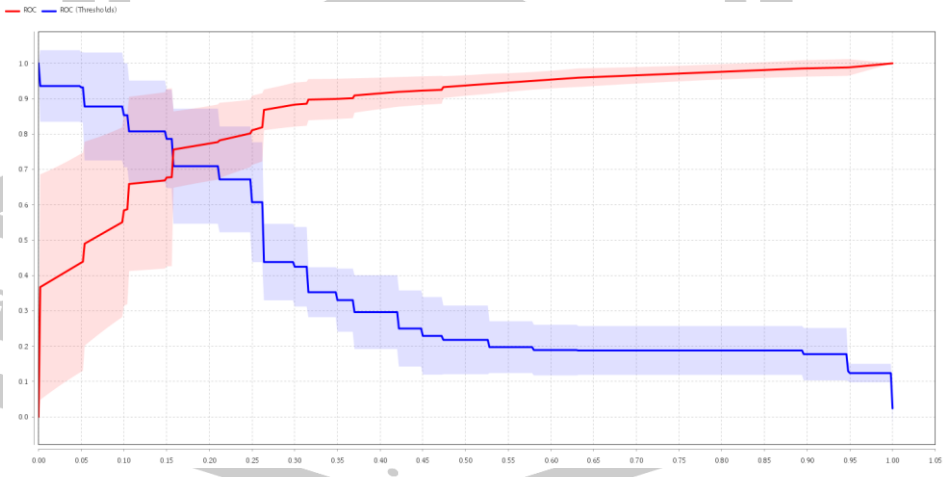
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.65% ตามรายละเอียดที่แสดงไว้ในตาราง 8

พหุน ปณ ฑิโต ชีเว

ตาราง 8 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	157	54	74.41%
	คะแนนรีวิว อยู่ในระดับต่ำ	35	214	85.94%
	class recall	81.77%	79.85%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 79.87 ค่าความจำเพาะ (specificity) ร้อยละ 81.79 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 86.14 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 74.49 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 82.79 และมีค่า AUC เท่ากับ 0.866 ดังแสดงในภาพประกอบ 40



ภาพประกอบ 40 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังภาพประกอบ 41

ภาพประกอบ 41 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

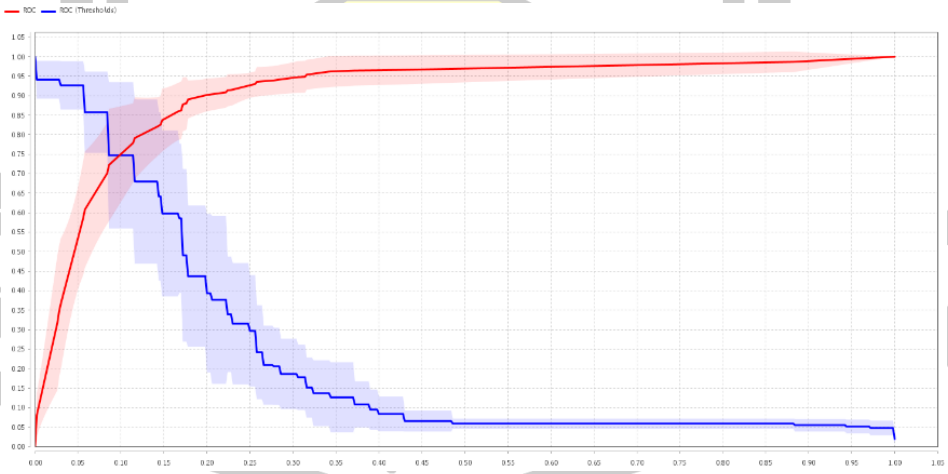
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) ซึ่งมีความคิดเห็นจากบทวิจารณ์ (Sentiment) เป็นโหนดรากสามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นช่วงราคา(price) จะเป็นตัวแปรที่มีผลต่อการพิจารณาเป็นลำดับถัดมา จากนั้น การมีบริการเครื่องดื่มแอลกอฮอล์(alcohol) จำนวนบทวิจารณ์ (reviews) และ การมีบริการ Wi-Fi(Wi-Fi) จะถูกพิจารณาเป็นลำดับสุดท้าย

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 84.32% ตามรายละเอียดที่แสดงไว้ในตาราง 9

ตาราง 9 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

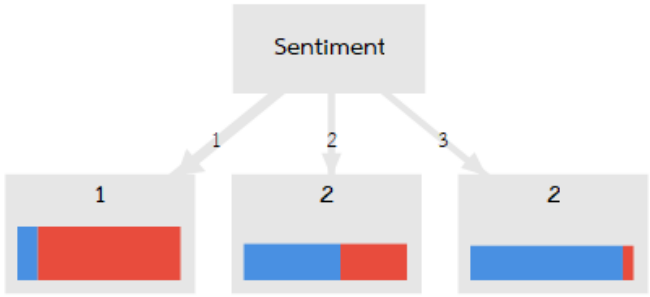
ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	300	39	88.50%
	คะแนนรีวิว อยู่ในระดับต่ำ	49	173	77.93%
	class recall	85.96%	81.60%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 81.62 ค่าความจำเพาะ (specificity) ร้อยละ 85.97 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 78.36 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 88.78 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.61 และมีค่า AUC เท่ากับ 0.908 ดังแสดงในภาพประกอบ 42



ภาพประกอบ 42 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังภาพประกอบ 43



ภาพประกอบ 43 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก

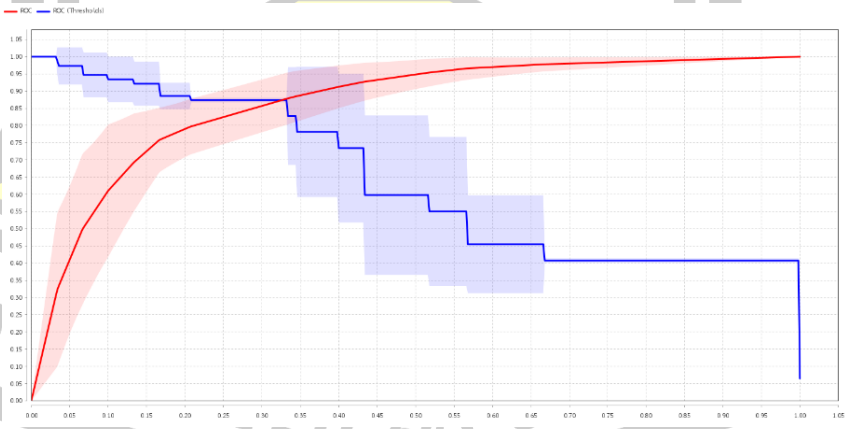
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 82.21% ตามรายละเอียดที่แสดงไว้ในตาราง 10

พหุน ปณฺ ฑิโต ชีเว

ตาราง 10 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	248	60	80.52%
	คะแนนรีวิว อยู่ในระดับต่ำ	50	260	83.87%
	class recall	83.22%	81.25%	

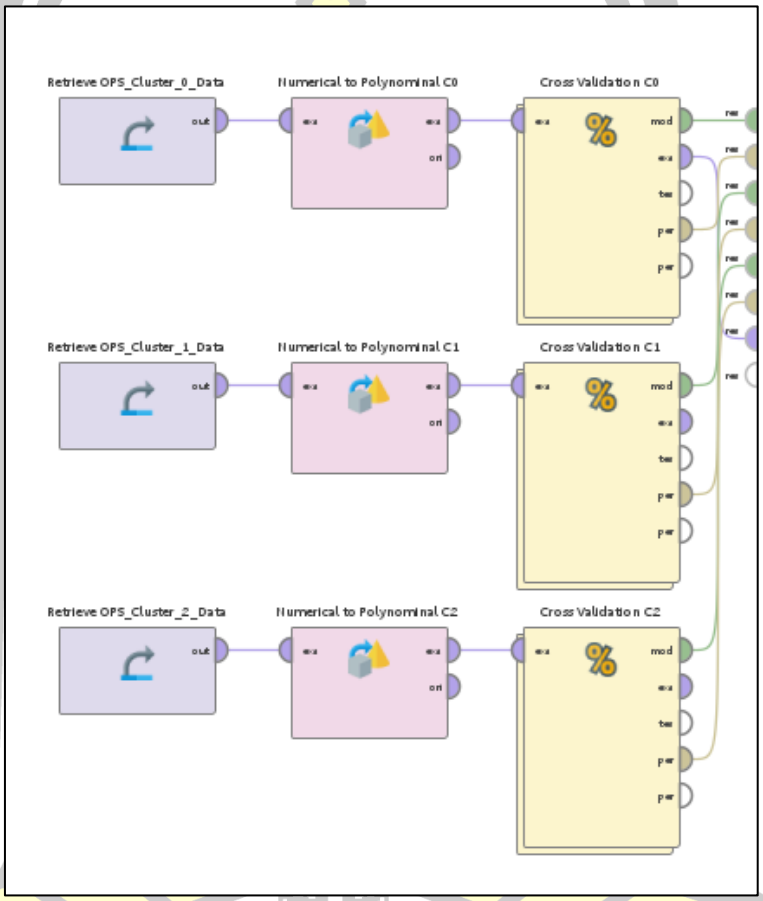
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 81.25 ค่าความจำเพาะ (specificity) ร้อยละ 83.26 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 84.44 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 80.61 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 82.58 และมีค่า AUC เท่ากับ 0.871 ดังแสดงในภาพประกอบ 44



ภาพประกอบ 44 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

1.3 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) แสดงดังภาพประกอบ 45



ภาพประกอบ 45 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

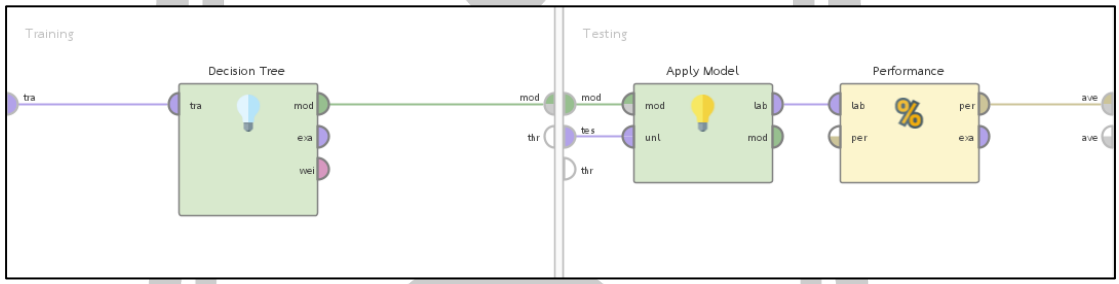
จากภาพประกอบ 44 แสดงขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) โดยมีขั้นตอน ดังนี้

1. Retrieve Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ทั้ง 3 กลุ่ม

2. Numerical to Polynominal : ทำการแปลงข้อมูล Numerical ให้อยู่ในรูปแบบของ Polynominal

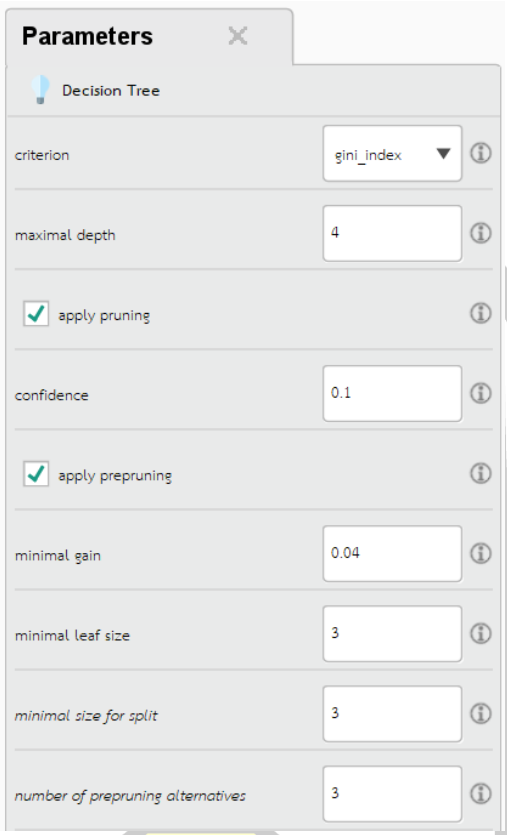
3. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Decision Tree ที่สร้างขึ้น

ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) สามารถแสดงได้ภาพประกอบ 46



ภาพประกอบ 46 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

3.1 Decision Tree : เป็นการสร้างตัวแบบพยากรณ์ Decision Tree จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ดังภาพประกอบ 47

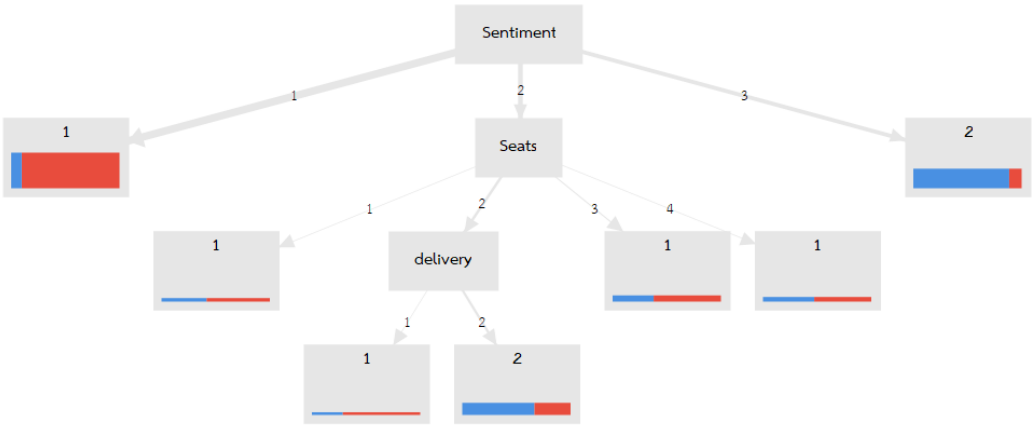


ภาพประกอบ 47 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

3.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

3.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Decision Tree ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังภาพประกอบ 48



ภาพประกอบ 48 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

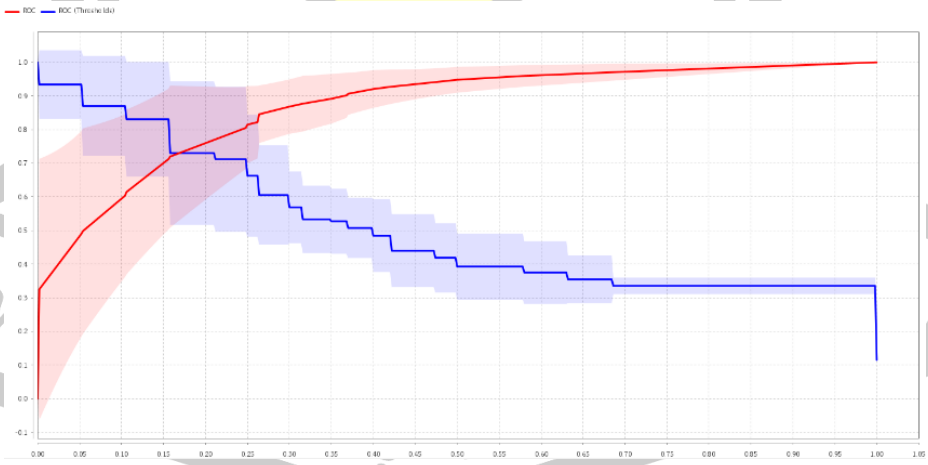
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นจำนวนที่นั่ง(seats) จะเป็นปัจจัยที่มีความสำคัญในลำดับถัดมา และบริการดิลิเวอรี(delivery) เป็นปัจจัยที่มีความสำคัญเป็นลำดับสุดท้าย

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.22 % ตามรายละเอียดที่แสดงไว้ในตาราง 11

ตาราง 11 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบต้นไม้ตัดสินใจ		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	145	44	76.72%
	คะแนนรีวิวอยู่ในระดับต่ำ	47	224	82.66%
	class recall	75.52%	83.58%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.55 ค่าความจำเพาะ (specificity) ร้อยละ 75.61 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.87 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 77.29 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.05 และมีค่า AUC เท่ากับ 0.869 ดังแสดงในภาพประกอบ 49



ภาพประกอบ 49 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังภาพประกอบ 49

ภาพประกอบ 50 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นช่วงราคา(price) จะเป็นปัจจัยที่มีความสำคัญในลำดับถัดมา และการมีบริการเครื่องดื่มแอลกอฮอล์ (alcohol) และจำนวนบทวิจารณ์(reviews) เป็นปัจจัยที่มีความสำคัญในลำดับสุดท้าย

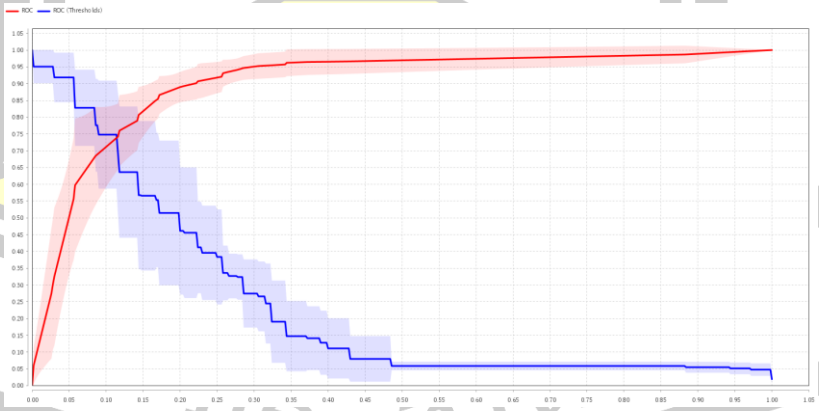
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 83.96% ตามรายละเอียดที่แสดงไว้ในตาราง 12

พหุน ปณฑิโต ชีเว

ตาราง 12 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

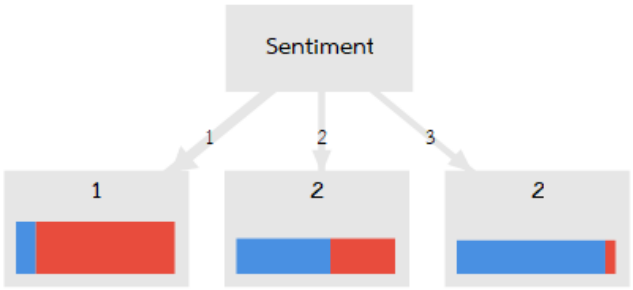
ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	311	52	85.67%
	คะแนนรีวิว อยู่ในระดับต่ำ	38	160	80.81%
	class recall	89.11%	75.47%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 75.50 ค่าความจำเพาะ (specificity) ร้อยละ 89.15 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 81.19 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 85.90 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.89 และมีค่า AUC เท่ากับ 77.87 ดังแสดงในภาพประกอบ 51



ภาพประกอบ 51 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังภาพประกอบ 52



ภาพประกอบ 52 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

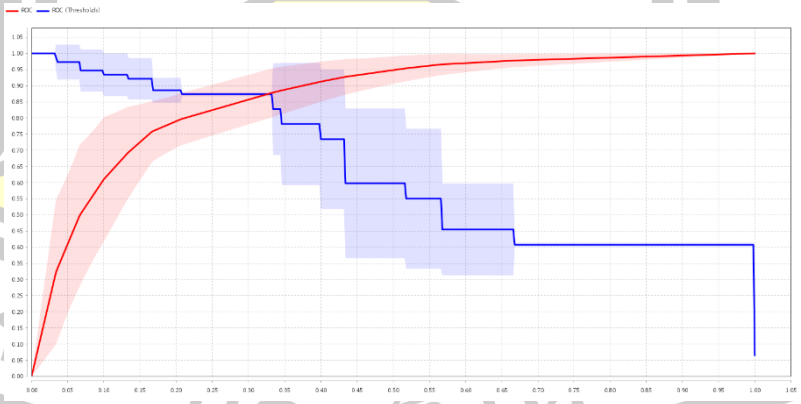
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์เป็นปัจจัยสำคัญเพียงปัจจัยเดียว

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.91% ตามรายละเอียดที่แสดงไว้ในตาราง 13

ตาราง 13 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบต้นไม้ตัดสินใจ		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	264	84	75.86%
	คะแนนรีวิว อยู่ในระดับต่ำ	34	236	87.41%
	class recall	88.59%	73.75%	

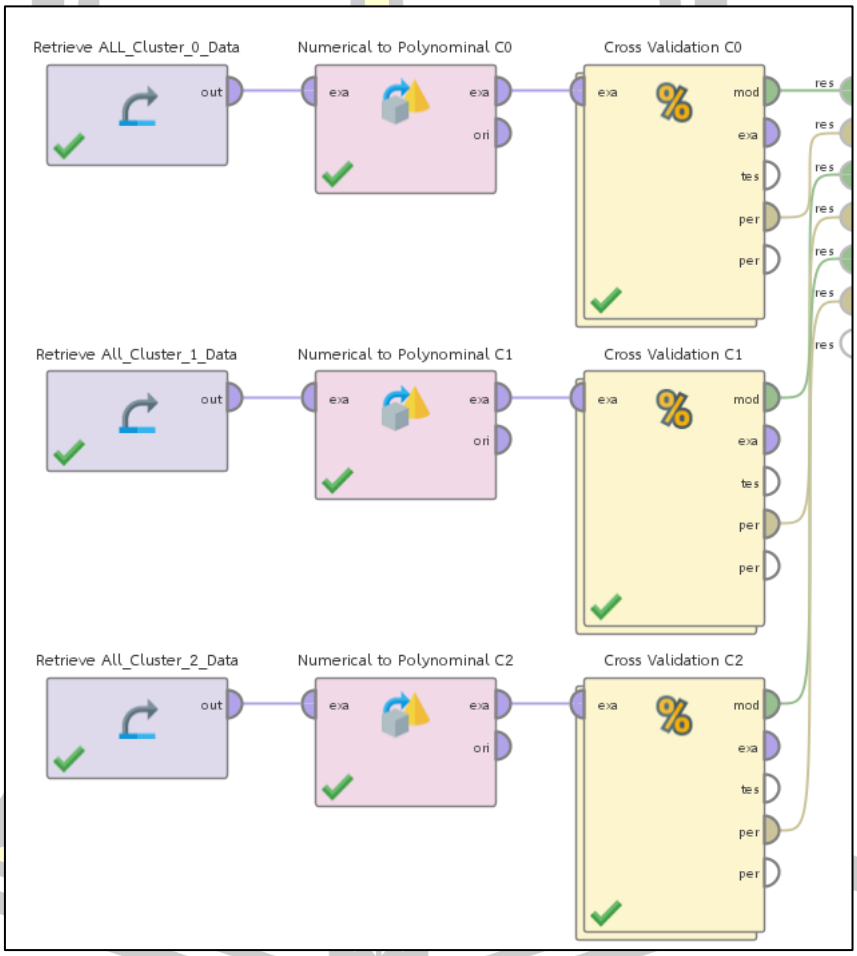
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.75 ค่าความจำเพาะ (specificity) ร้อยละ 88.59 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 87.64 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 76.14 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.89 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 53



ภาพประกอบ 53 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

1.4 การสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) แสดงดังภาพประกอบ 54



ภาพประกอบ 54 ขั้นตอนการสร้างตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

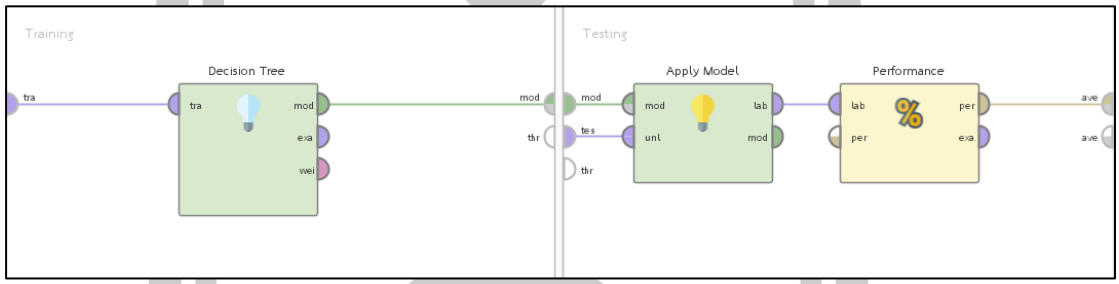
จากภาพประกอบ 44 แสดงขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) โดยมีขั้นตอน ดังนี้

1. Retrieve Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ทั้ง 3 กลุ่ม

2. Numerical to Polynominal : ทำการแปลงข้อมูล Numerical ให้อยู่ในรูปแบบของ Polynominal

3. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Decision Tree ที่สร้างขึ้น

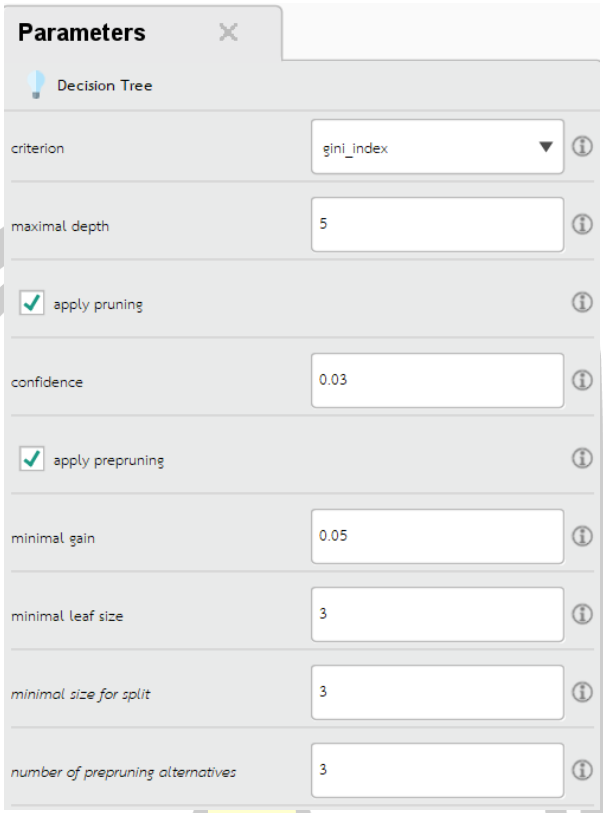
ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) สามารถแสดงได้ภาพประกอบ 55



ภาพประกอบ 55 ขั้นตอนการสร้างตัวแบบต้นไม้ตัดสินใจภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

3.1 Decision Tree : เป็นการสร้างตัวแบบพยากรณ์ Decision Tree จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ดังภาพประกอบ 56

พหุน ปณ ฑิโต ชีเว

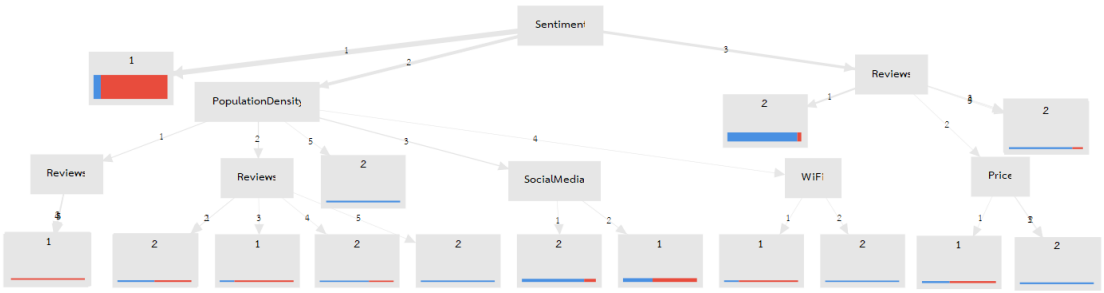


ภาพประกอบ 56 การกำหนดค่าพารามิเตอร์สำหรับ Decision Tree ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

3.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

3.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Decision Tree ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังภาพประกอบ 57



ภาพประกอบ 57 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นความหนาแน่นของประชากรในพื้นที่ (PopulationDensity) และจำนวนบทวิจารณ์ (Reviews) จะเป็นปัจจัยที่มีความสำคัญในลำดับถัดมา การมีโซเชียลมีเดียอื่นๆ (SocialMedia) มีบริการ (Wi-Fi) และช่วงราคา (Price) เป็นปัจจัยที่มีความสำคัญเป็นลำดับสุดท้าย

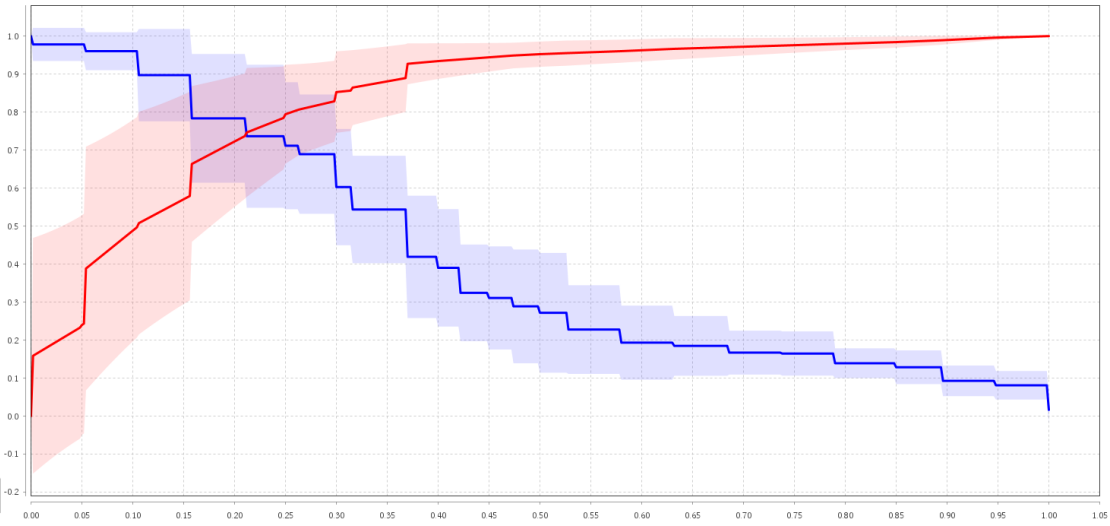
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 80.43% ตามรายละเอียดที่แสดงไว้ในตาราง 14

ตาราง 14 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมดในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	138	37	78.86%

ตัวแบบต้นไม้ตัดสินใจ		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
	คะแนนรีวิว อยู่ในระดับต่ำ	54	231	81.05%
	class recall	71.88%	86.19%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 86.21 ค่าความจำเพาะ (specificity) ร้อยละ 71.89 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 81.33 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 79.65 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.47 และมีค่า AUC เท่ากับ 0.842 ดังแสดงในภาพประกอบ 58



ภาพประกอบ 58 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมดในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังภาพประกอบ 59

ภาพประกอบ 59 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

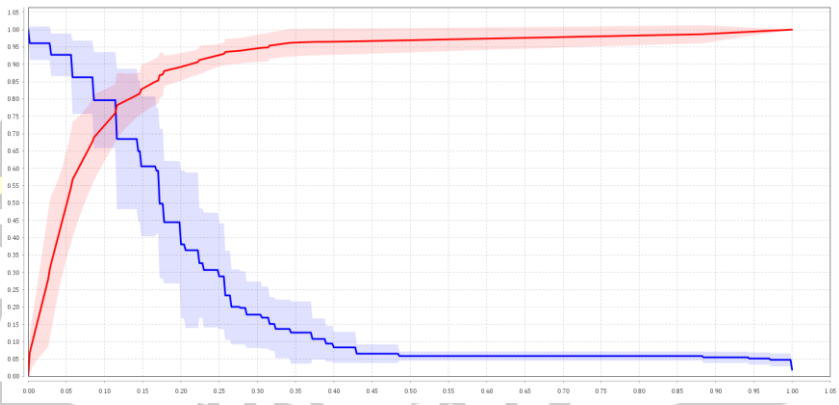
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นช่วงราคา (price) จะเป็นปัจจัยที่มีความสำคัญในลำดับถัดมา และการมีบริการเครื่องดื่มแอลกอฮอล์ (alcohol) จำนวนบทวิจารณ์ (Reviews) และการมีบริการ Wi-Fi เป็นปัจจัยที่มีความสำคัญในลำดับสุดท้าย

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 83.78% ตามรายละเอียดที่แสดงไว้ในตาราง 15

ตาราง 15 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบต้นไม้ตัดสินใจ		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	299	41	87.94%
	คะแนนรีวิวอยู่ในระดับต่ำ	50	171	77.38%
	class recall	85.67%	80.66%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 80.67 ค่าความจำเพาะ (specificity) ร้อยละ 85.69 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 77.79 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 88.25 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 78.85 และมีค่า AUC เท่ากับ 0.904 ดังแสดงในภาพประกอบ 60



ภาพประกอบ 60 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์ต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังภาพประกอบ 61

ภาพประกอบ 61 ผลการพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

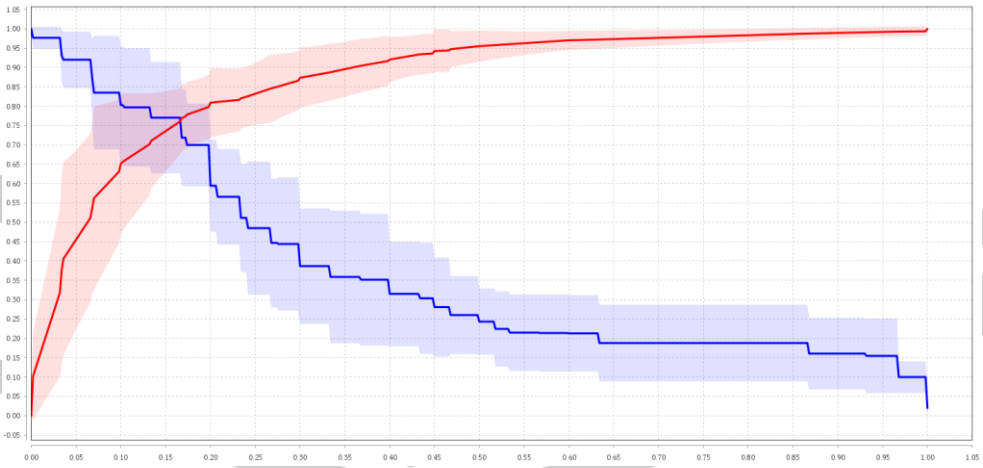
จากผลการวิเคราะห์การตัดสินใจด้วยตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถอธิบายได้ว่า ความคิดเห็นจากบทวิจารณ์จะถูกพิจารณาเป็นลำดับแรก จากนั้นจำนวนที่นั่ง (Seats) จะเป็นปัจจัยที่มีความสำคัญในลำดับถัดมา และความหนาแน่นของประชากรในพื้นที่ (PopulationDensity) ช่วงราคา (Price) และจำนวนเช็คอิน (Check-in) เป็นปัจจัยที่มีความสำคัญในลำดับสุดท้าย

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบต้นไม้ตัดสินใจ (Decision Tree) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 80.43% ตามรายละเอียดที่แสดงไว้ในตาราง 16

ตาราง 16 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบต้นไม้ตัดสินใจ		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	238	61	79.60%
	คะแนนรีวิว อยู่ในระดับต่ำ	60	259	81.19%
	class recall	79.87%	80.94%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 80.94 ค่าความจำเพาะ (specificity) ร้อยละ 79.91 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 81.75 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 79.76 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 81.12 และมีค่า AUC เท่ากับ 0.872 ดังแสดงในภาพประกอบ 62

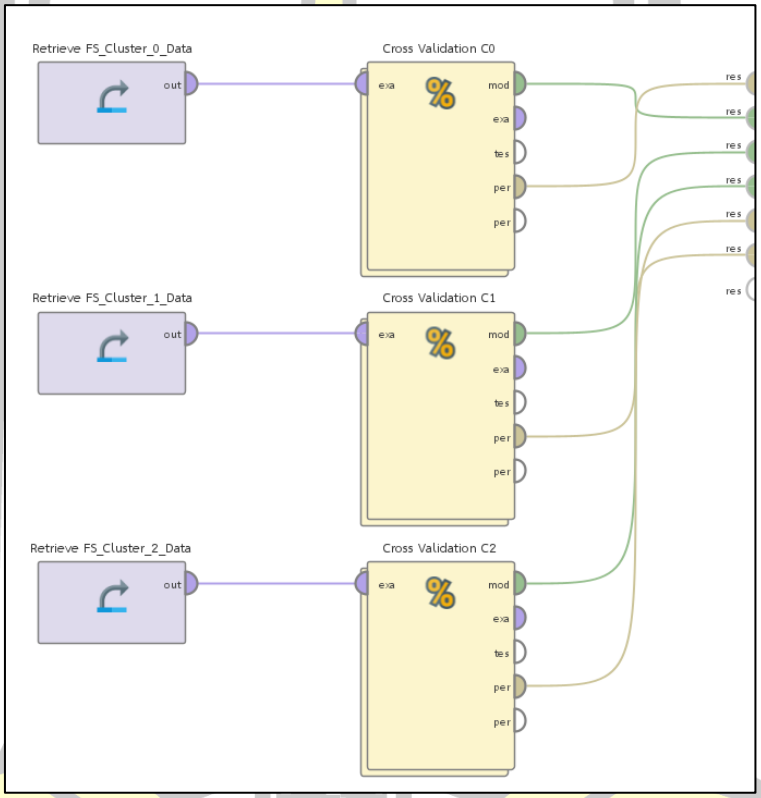


ภาพประกอบ 62 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์ต้นไม้ตัดสินใจด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

2. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression)

2.1 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection) แสดงดังภาพประกอบ 63



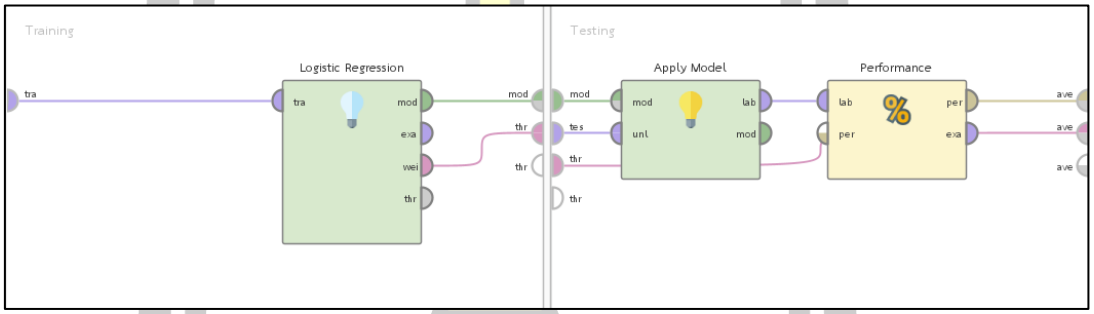
ภาพประกอบ 63 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection)

จากภาพประกอบ 63 แสดงขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (forward selection) โดยมีขั้นตอน ดังนี้

1. Retrieve_FS_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์จากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (forward selection) ทั้ง 3 กลุ่ม

2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Logistic regression ที่สร้างขึ้น

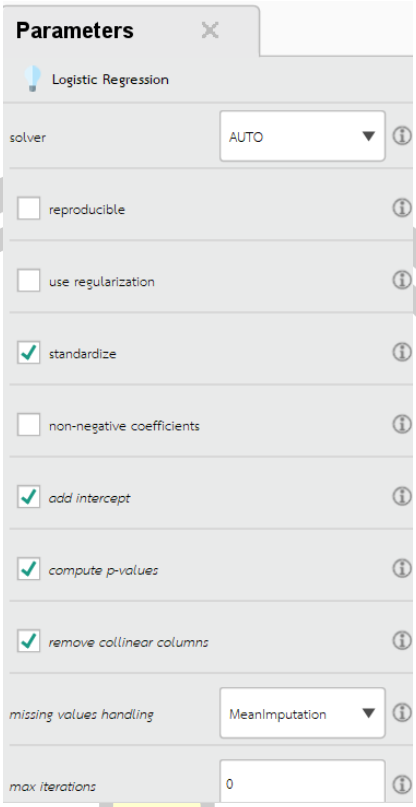
ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection) สามารถแสดงได้ดังภาพประกอบ 64



ภาพประกอบ 64 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (forward selection)

2.1 Logistic regression : เป็นการสร้างตัวแบบพยากรณ์ Logistic regression จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ดังภาพประกอบ 65

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 65 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีเพิ่มตัวแปร (forward selection)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Logistic regression ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังตาราง 17

ตาราง 17 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

Attribute	Coeffcient
Sentiment	-2.129586103731607
Seats	0.3475129109024351
delivery	-0.7852817169939895

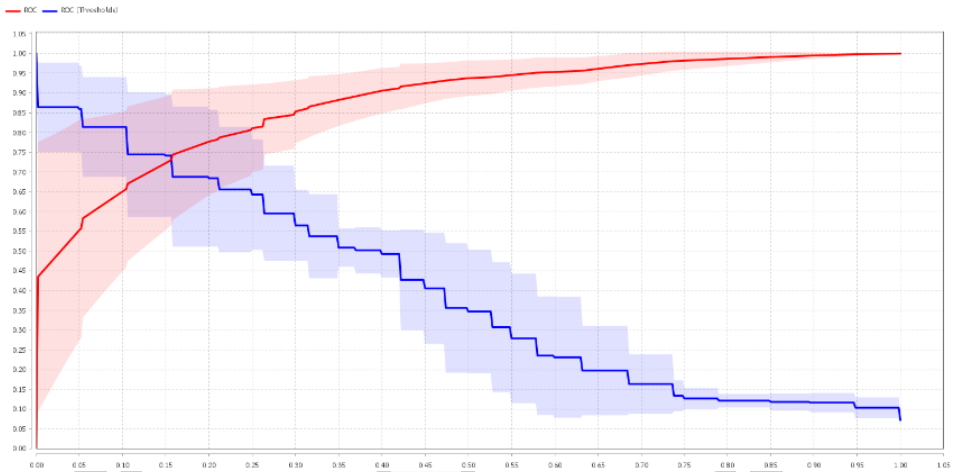
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 81.52% ตามรายละเอียดที่แสดงไว้ในตาราง 18

ตาราง 18 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

		ผลที่เกิดขึ้นจริง		
ตัวแบบการถดถอยโลจิสติกส์		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	149	42	78.01%
	คะแนนรีวิวอยู่ในระดับต่ำ	43	226	84.01%
	class recall	77.60%	84.33%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 84.29 ค่าความจำเพาะ (specificity) ร้อยละ 77.71 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 84.03 ค่าการทำนายผลลบ (negative predictive value)

ร้อยละ 78.32 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 84.10 และมีค่า AUC เท่ากับ 0.877 ดังแสดงในภาพประกอบ 66



ภาพประกอบ 66 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังตาราง 19

ตาราง 19 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

Attribute	Coeffcient
Sentiment	-2.5502952763247735
Price	-0.3766878855199519
alcohol	-0.5243974839543829
Reviews	0.061683316996407764
Group	-0.08793070066010084

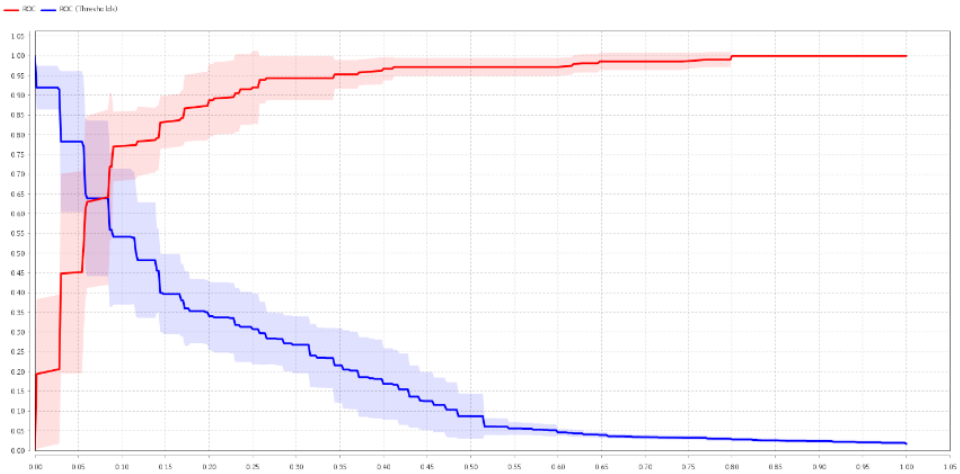
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 84.31% ตามรายละเอียดที่แสดงไว้ในตาราง 20

ตาราง 20 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	320	59	84.43%
	คะแนนรีวิวอยู่ในระดับต่ำ	29	153	84.07%
	class recall	91.69%	72.17%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 72.19 ค่าความจำเพาะ (specificity) ร้อยละ 91.71 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 84.79 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 84.73 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 77.36 และมีค่า AUC เท่ากับ 0.908 ดังแสดงในภาพประกอบ 67

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 67 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังตาราง 21

ตาราง 21 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

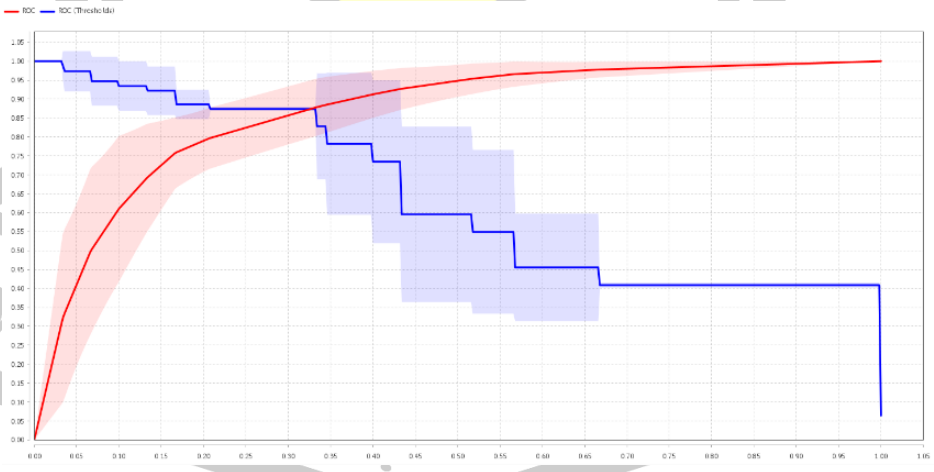
Attribute	Coeffcient
Sentiment	-2.3037637674405724

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.91% ตามรายละเอียดที่แสดงไว้ในตาราง 22

ตาราง 22 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	264	84	75.86%
	คะแนนรีวิว อยู่ในระดับต่ำ	34	236	87.41%
	class recall	88.59%	73.75%	

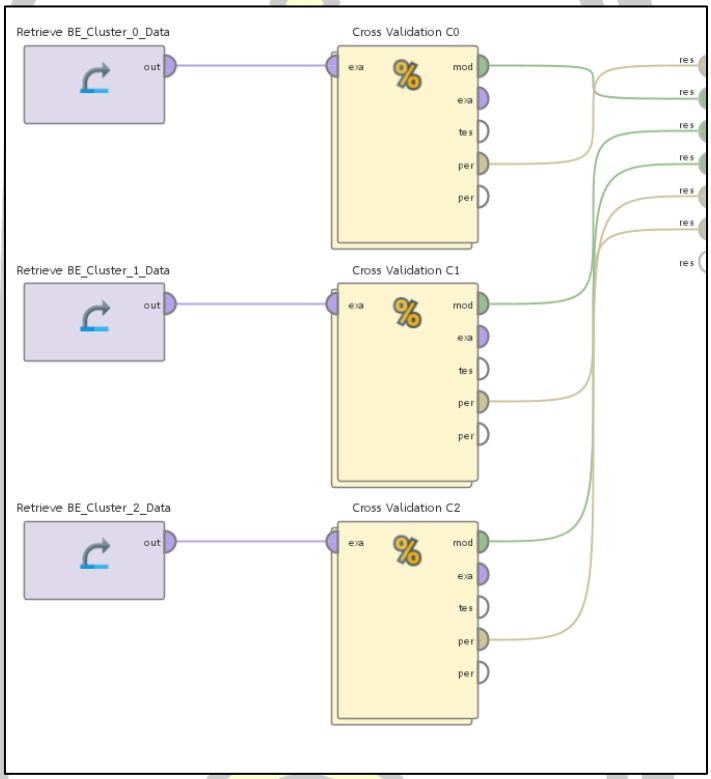
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.75 ค่าความจำเพาะ (specificity) ร้อยละ 88.59 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 87.64 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 76.14 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.89 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 68



ภาพประกอบ 68 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

2.2 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination)

ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) แสดงดังภาพประกอบ 69



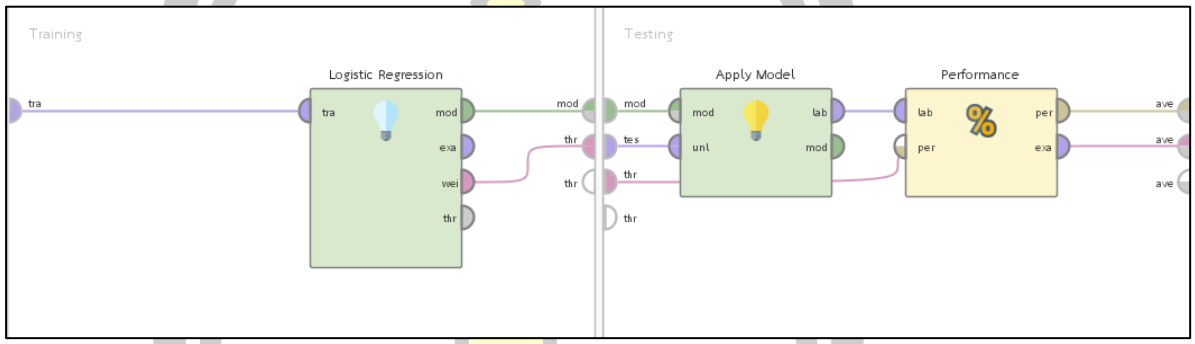
ภาพประกอบ 69 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination)

จากภาพประกอบ 69 แสดงขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) โดยมีขั้นตอน ดังนี้

1. Retrieve_BE_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์จากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) ทั้ง 3 กลุ่ม

2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Logistic regression ที่สร้างขึ้น

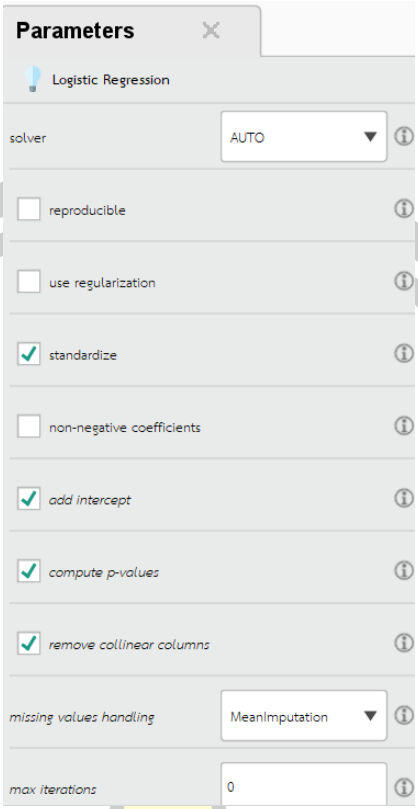
ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination) แสดงดังภาพประกอบ 70



ภาพประกอบ 70 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination)

2.1 Logistic regression : เป็นการสร้างตัวแบบพยากรณ์ Logistic regression จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 71

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 71 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีลดตัวแปร (Backward Elimination)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Logistic regression ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังตาราง 23

ตาราง 23 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

Attribute	Coeffcient
Price	-0.3763733517926911
Seats	0.5720998613242676
Check-in	-0.1804772135368417
Like	0.4001342711797505
PopulationDensity	-0.5527364811397888
Sentiment	-2.376610901484111
delivery	-0.7818504883707224
Group	-0.2300266071152017
alcohol	-0.3910083756190203
parking	-0.216396311949905
SocialMedia	0.20129875978390716

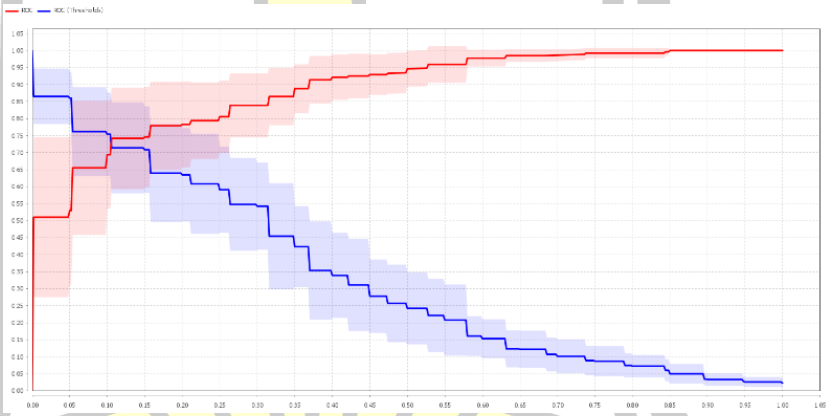
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 81.09% ตามรายละเอียดที่แสดงไว้ในตาราง 24

ตาราง 24 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

		ผลที่เกิดขึ้นจริง		
ตัวแบบการถดถอยโลจิสติกส์		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	144	39	78.69%

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
	คะแนนรีวิว อยู่ในระดับต่ำ	48	229	82.67%
	class recall	75.00%	85.45%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 85.41 ค่าความจำเพาะ (specificity) ร้อยละ 75.13 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.81 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 76.14 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 78.60 และมีค่า AUC เท่ากับ 0.888 ดังแสดงในภาพประกอบ 72



ภาพประกอบ 72 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังตาราง 25

ตาราง 25 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

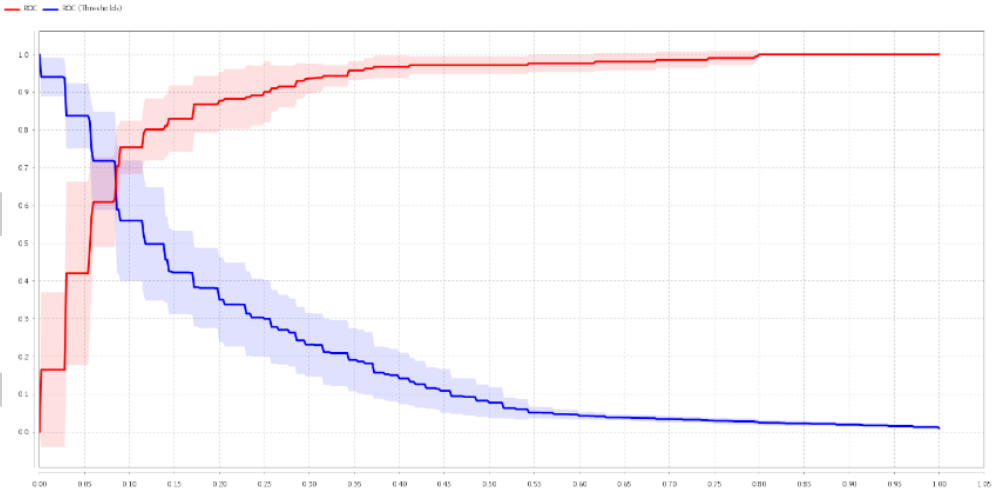
Attribute	Coeffcient
Price	-0.4336890915914939
Reviews	-0.0023279655036000766
Check-in	0.04778958272600127
Like	0.09679778172905854
PopulationDensity	0.007379404391427583
Sentiment	-2.5701362473805682
delivery	0.27358096034179796
Group	-0.1342942174620963
alcohol	-0.4701749167762231
parking	0.5634654211888388
creditCard	-6.320192549689422E-4
Website	-0.3698732148659093
SocialMedia	0.2583495835034037
WiFi	-0.2589637069298068

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 83.07% ตามรายละเอียดที่แสดงไว้ในตาราง 26

ตาราง 26 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	310	56	84.70%
	คะแนนรีวิว อยู่ในระดับต่ำ	39	156	80.00%
	class recall	88.83%	73.58%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.61 ค่าความจำเพาะ (specificity) ร้อยละ 88.83 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 80.38 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 85.01 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 76.40 และมีค่า AUC เท่ากับ 0.904 ดังแสดงในภาพประกอบ 73



ภาพประกอบ 73 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังตาราง 27

ตาราง 27 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

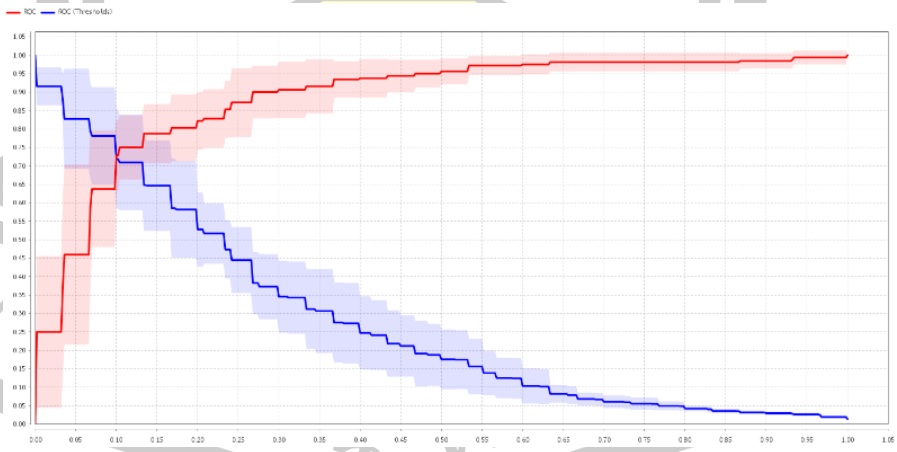
Attribute	Coeffcient
Price	-0.5604833854299905
Seats	0.39868856766556815
Reviews	0.14113480177496598
Check-in	-0.03602292325852113
PopulationDensity	-0.31735410851826024
Sentiment	-2.3806007906293876
delivery	-0.46205925618445276
alcohol	-0.40016609806916836
parking	0.19837505038796377
creditCard	0.31881812513979774
Website	0.4560225359028367
SocialMedia	-0.05011732866868885

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 82.22% ตามรายละเอียดที่แสดงไว้ในตาราง 28

ตาราง 28 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	241	53	81.97%
	คะแนนรีวิว อยู่ในระดับต่ำ	57	267	82.41%
	class recall	80.87%	83.44%	

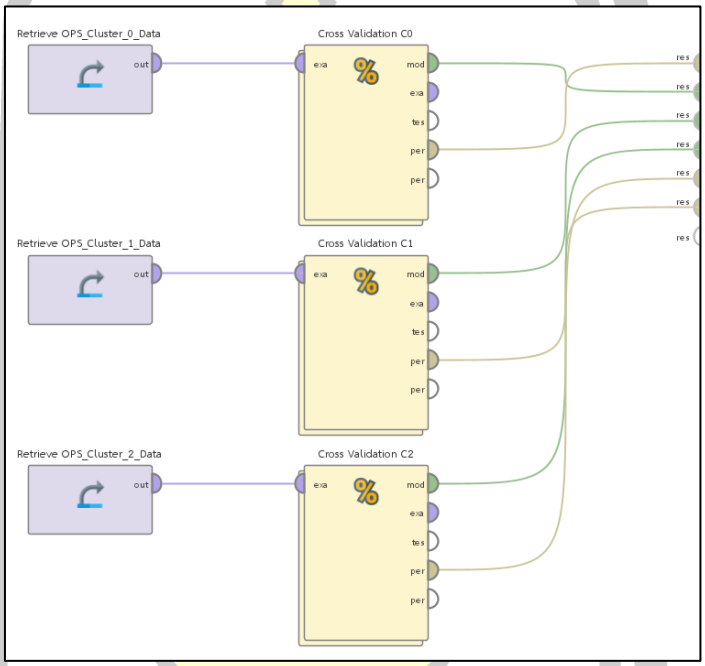
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.44 ค่าความจำเพาะ (specificity) ร้อยละ 80.91 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.85 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 82.78 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 82.82 และมีค่า AUC เท่ากับ 0.885 ดังแสดงในภาพประกอบ 74



ภาพประกอบ 74 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

2.3 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและเพิ่มและลดตัวแปร (Optimize selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) แสดงดังภาพประกอบ 75

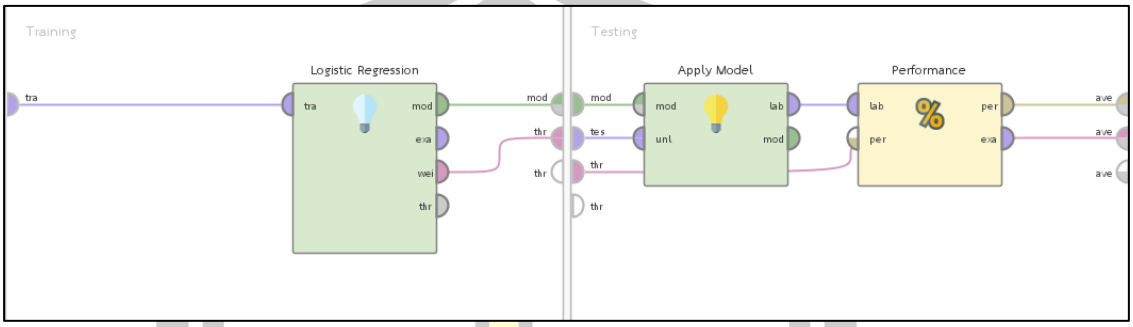


ภาพประกอบ 75 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)

จากภาพประกอบ 75 แสดงขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) โดยมีขั้นตอน ดังนี้

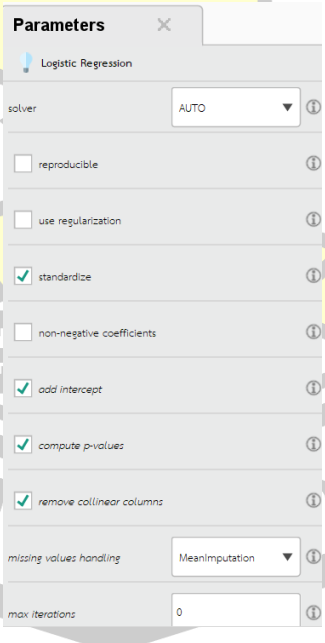
1. Retrieve_OPS_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์จากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ทั้ง 3 กลุ่ม
2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Logistic regression ที่สร้างขึ้น

ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) แสดงดังภาพประกอบ 76



ภาพประกอบ 76 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)

2.1 Logistic regression : เป็นการสร้างตัวแบบพยากรณ์ Logistic regression จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 77



ภาพประกอบ 77 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Logistic regression ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังตาราง 29

ตาราง 29 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

Attribute	Coeffcient
Seats	0.3475129109024362
Sentiment	-2.1295861037316155
delivery	-0.7852817169939917

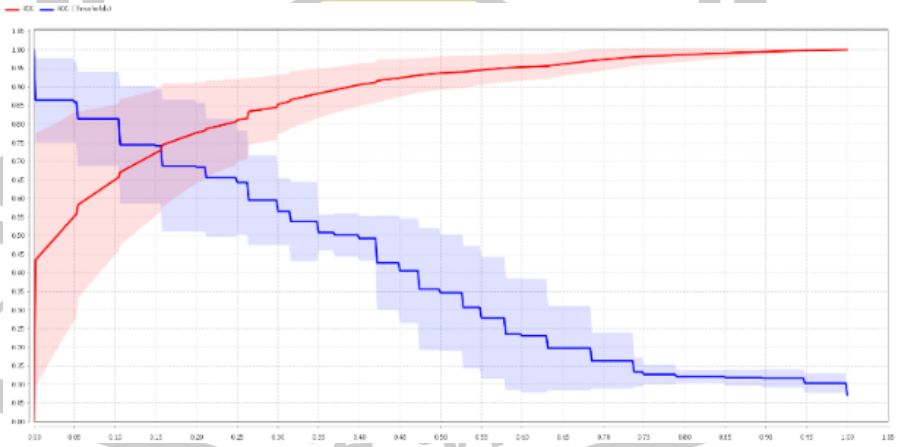
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 81.52% ตามรายละเอียดที่แสดงไว้ในตาราง 30

พหุน ปณ ฑิโต ชีเว

ตาราง 30 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	149	42	78.01%
	คะแนนรีวิว อยู่ในระดับต่ำ	43	226	84.01%
	class recall	77.60%	84.33%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 84.29 ค่าความจำเพาะ (specificity) ร้อยละ 77.71 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 84.03 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 78.32 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 84.10 และมีค่า AUC เท่ากับ 0.877 ดังแสดงในภาพประกอบ 78



ภาพประกอบ 78 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังตาราง 31

ตาราง 31 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

Attribute	Coeffcient
Price	-0.37668788551995147
Reviews	0.06168331699640766
Sentiment	-2.550295276324772
Group	-0.08793070066010086
alcohol	-0.5243974839543826

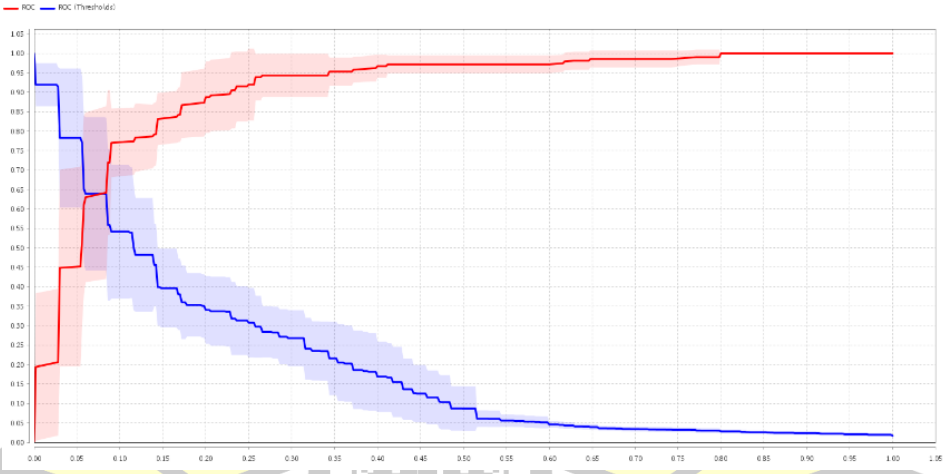
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 84.31% ตามรายละเอียดที่แสดงไว้ในตาราง 32

ตาราง 32 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	320	59	84.43%
	คะแนนรีวิว	29	153	8 4.07%

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
	อยู่ในระดับต่ำ			
	class recall	91.69%	72.17%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 72.19 ค่าความจำเพาะ (specificity) ร้อยละ 91.71 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 84.79 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 84.73 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 77.36 และมีค่า AUC เท่ากับ 0.908 ดังแสดงในภาพประกอบ 79



ภาพประกอบ 79 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังตาราง 33

ตาราง 33 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

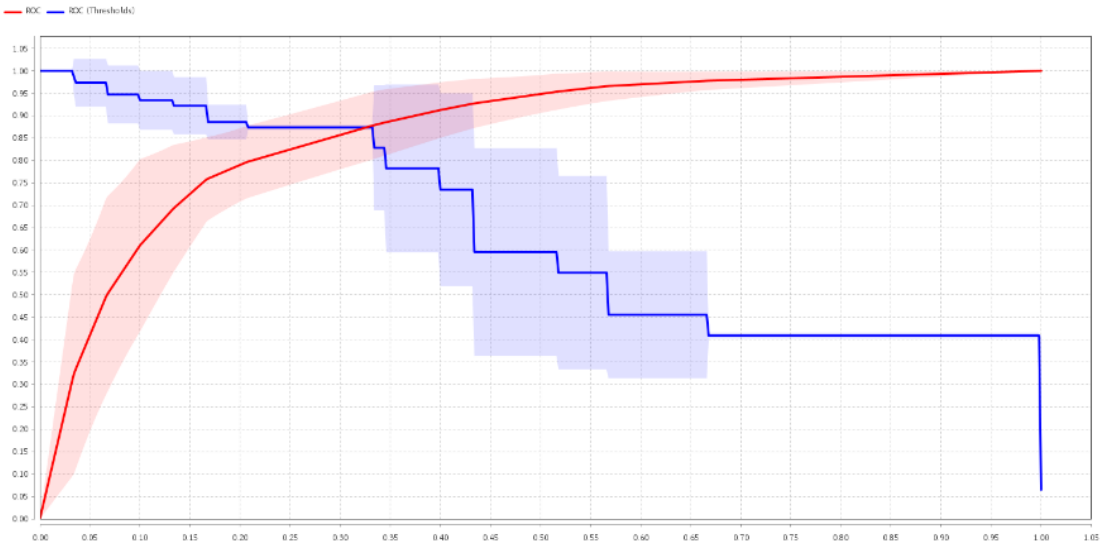
Attribute	Coeffcient
Sentiment	-2.3037637674405724

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.91% ตามรายละเอียดที่แสดงไว้ในตาราง 34

ตาราง 34 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	264	84	75.86%
	คะแนนรีวิวอยู่ในระดับต่ำ	34	236	87.41%
	class recall	88.59%	73.75%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.75 ค่าความจำเพาะ (specificity) ร้อยละ 88.59 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 87.64 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 76.14 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 79.89 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 80

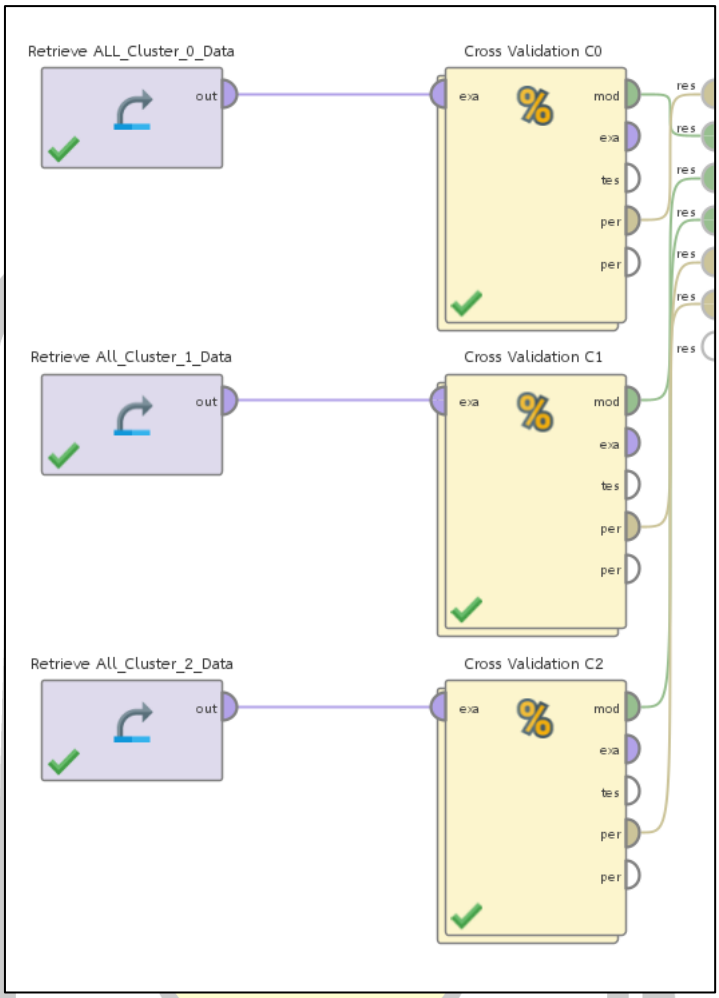


ภาพประกอบ 80 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

2.4 การสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) แสดงดังภาพประกอบ 81

พหุน ปณ ฑิโต ชีเว

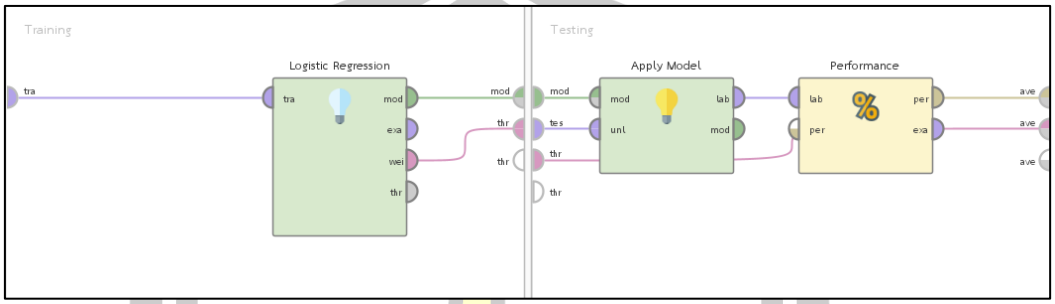


ภาพประกอบ 81 ขั้นตอนการสร้างตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

จากภาพประกอบ 81 แสดงขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) โดยมีขั้นตอน ดังนี้

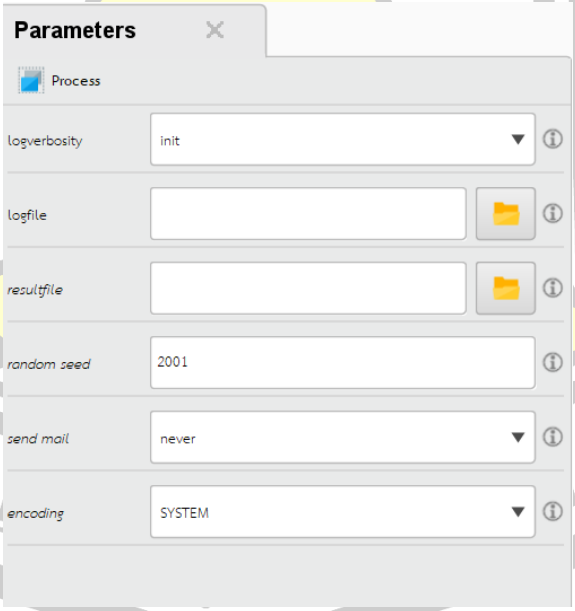
1. Retrieve_All_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์จากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ทั้ง 3 กลุ่ม
2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Logistic regression ที่สร้างขึ้น

ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) แสดงดังภาพประกอบ 82



ภาพประกอบ 82 ขั้นตอนการสร้างตัวแบบการถดถอยโลจิสติกส์ภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

2.1 Logistic regression : เป็นการสร้างตัวแบบพยากรณ์ Logistic regression จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 83



ภาพประกอบ 83 การกำหนดค่าพารามิเตอร์สำหรับ Logistic regression ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Logistic regression ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) สามารถแสดงได้ดังตาราง 35

ตาราง 35 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

Attribute	Coeffcient
Seats	0.571
Like	0.382
SocialMedia	0.299
Reviews	0.080
Website	0.073
creditCard	-0.043
Group	-0.081
parking	-0.156
Check-in	-0.184
alcohol	-0.277
Price	-0.368
WiFi	-0.539
PopulationDensity	-0.542
delivery	-0.874
Sentiment	-2.377

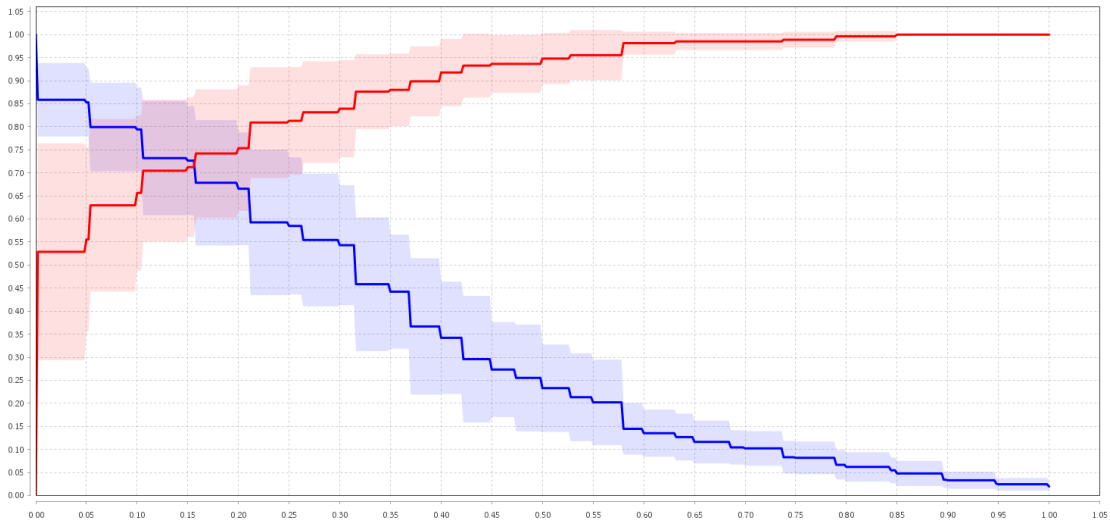
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 81.09% ตามรายละเอียดที่แสดงไว้ในตาราง 36

ตาราง 36 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบการถดถอยโลจิสติกส์		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	143	38	79.01%
	คะแนนรีวิวอยู่ในระดับต่ำ	49	230	82.44%
	class recall	74.48%	85.82%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 85.80 ค่าความจำเพาะ (specificity) ร้อยละ 74.61 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.59 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 78.98 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 84.11 และมีค่า AUC เท่ากับ 0.885 ดังแสดงในภาพประกอบ 84

พหุน ปณฺฑิโต ชีเว



ภาพประกอบ 84 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) สามารถแสดงได้ดังตาราง 37

ตาราง 37 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

Attribute	Coeffcient
parking	0.5658699962386653
delivery	0.27438611288042625
SocialMedia	0.25594618316746925
Like	0.09788354639507008
Check-in	0.04898997598600827
PopulationDensity	0.007150824502338128
creditCard	0.0020900016575968435
Reviews	-0.0017715758224614085

Attribute	Coeffcient
Seats	-0.013260811723665387
Group	-0.13453967453566731
WiFi	-0.25585193319886923
Website	-0.36715886874435766
Price	-0.43494013930880104
alcohol	-0.4684179401335637
Sentiment	-2.5709825860844733

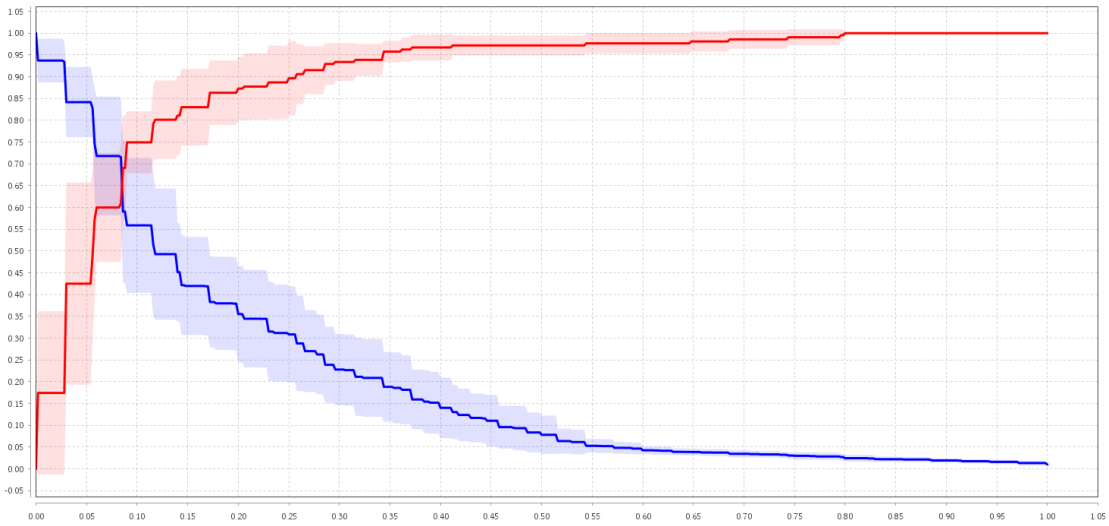
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 82.71% ตามรายละเอียดที่แสดงไว้ในตาราง 38

ตาราง 38 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

		ผลที่เกิดขึ้นจริง		
ตัวแบบการถดถอยโลจิสติกส์		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	308	56	84.62%
	คะแนนรีวิวอยู่ในระดับต่ำ	41	156	79.19%
	class recall	88.25%	73.58%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 73.61 ค่าความจำเพาะ (specificity) ร้อยละ 88.24 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 79.74 ค่าการทำนายผลลบ (negative predictive value)

ร้อยละ 84.92 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 76.07 และมีค่า AUC เท่ากับ 0.903 ดังแสดงในภาพประกอบ 85



ภาพประกอบ 85 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์การถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) สามารถแสดงได้ดังตาราง 39

ตาราง 39 ผลการพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

Attribute	Coeffcient
Website	0.4584583913374028
Seats	0.39828276945814955
creditCard	0.29066393717953976
parking	0.217882353160596
Reviews	0.13161977406175976

Attribute	Coeffcient
WiFi	0.10361037274723899
Like	0.0855083872218786
Check-in	-0.044039816345727484
SocialMedia	-0.04887315217485365
Group	-0.14151719751486208
PopulationDensity	-0.32123721012383183
alcohol	-0.39123519075563007
delivery	-0.4517997812945683
Price	-0.5657810138034306
Sentiment	-2.387730681122083

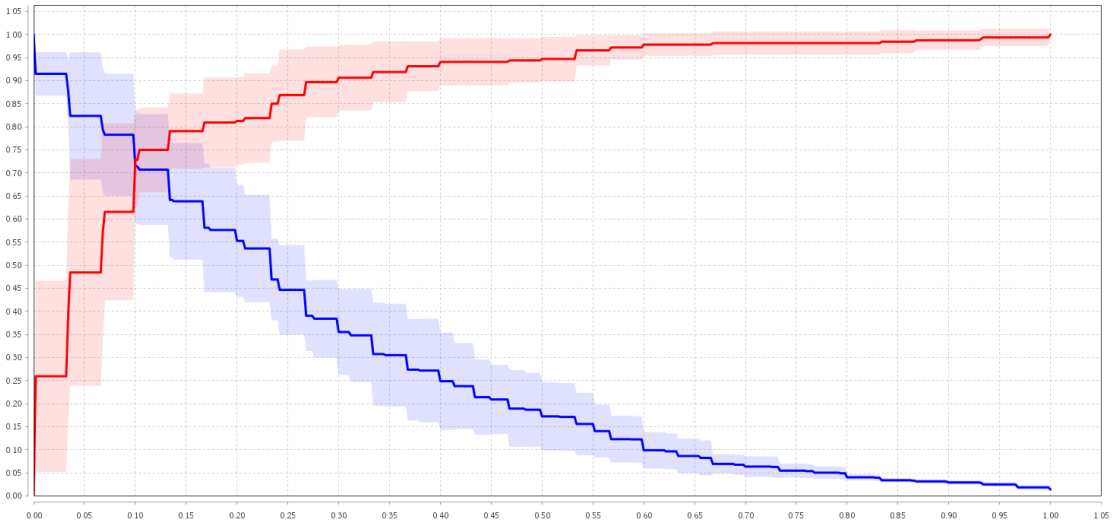
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบการถดถอยโลจิสติกส์ (Logistic regression) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 81.57% ตามรายละเอียดที่แสดงไว้ในตาราง 40

ตาราง 40 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบการถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	240	56	81.08%
	คะแนนรีวิวอยู่ในระดับต่ำ	58	264	81.99%

ตัวแบบการถดถอยโลจิสติกส์		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
	class recall	80.54%	82.50%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 82.50 ค่าความจำเพาะ (specificity) ร้อยละ 80.57 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.35 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 81.82 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 82.12 และมีค่า AUC เท่ากับ 0.885 ดังแสดงในภาพประกอบ 86

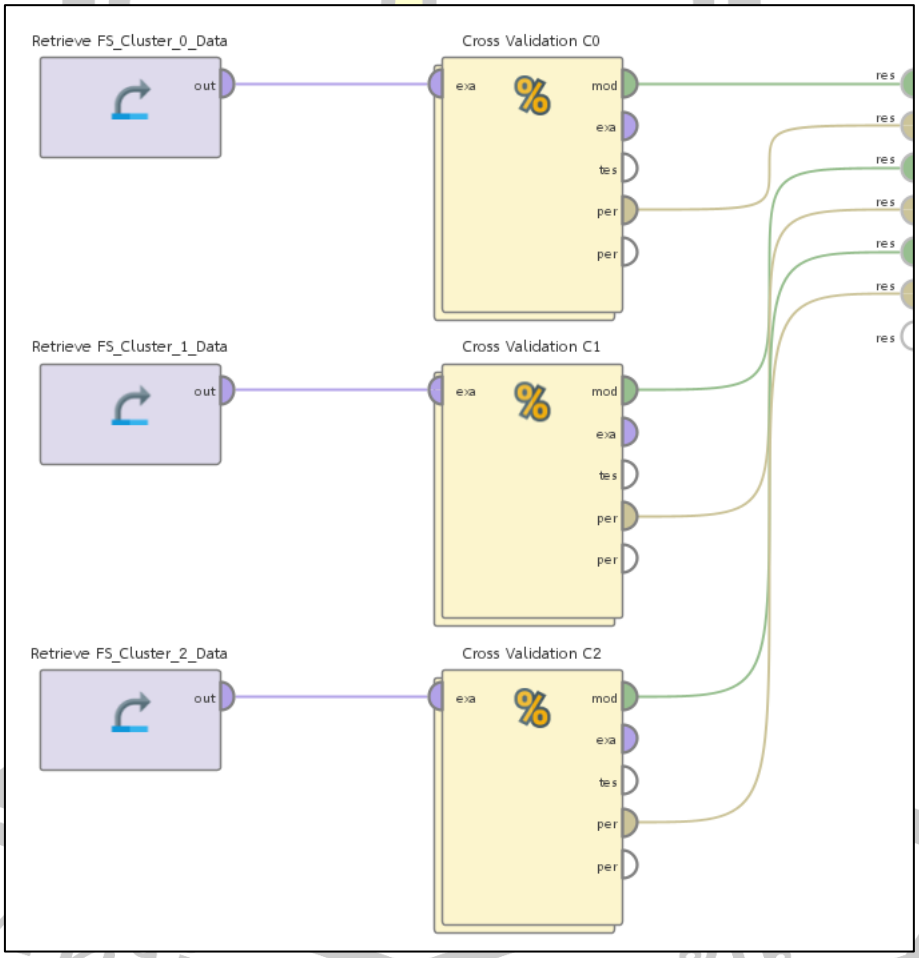


ภาพประกอบ 86 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์การถดถอยโลจิสติกส์ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

3. การสร้างตัวแบบพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)

3.1 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) แสดงดังภาพประกอบ 87



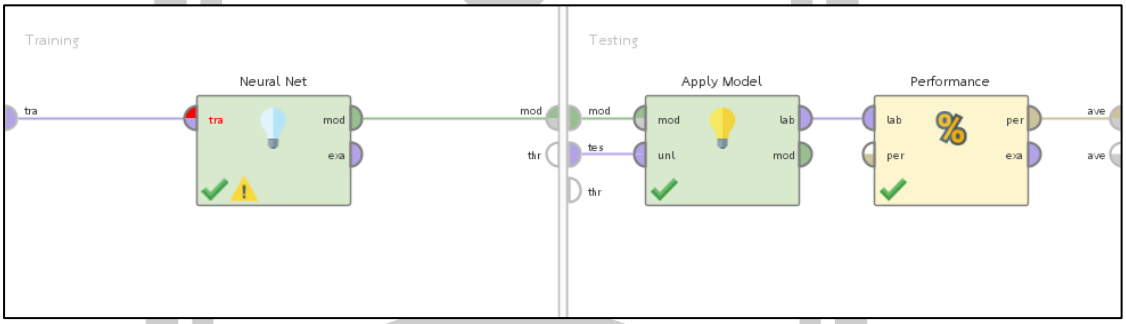
ภาพประกอบ 87 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection)

จากภาพประกอบ 87 แสดงขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีเพิ่มตัวแปร (Forward selection) โดยมีขั้นตอน ดังนี้

1. Retrieve_FS_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียมจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปร (Forward selection) ทั้ง 3 กลุ่ม

2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Neural Networks ที่สร้างขึ้น

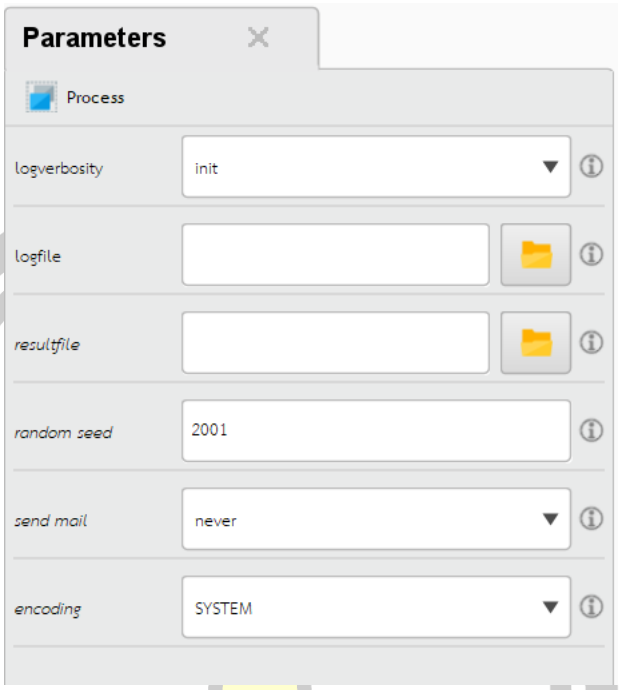
ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (Forward selection) แสดงดังภาพประกอบ 88



ภาพประกอบ 88 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มตัวแปร (Forward selection)

2.1 Neural Net : เป็นการสร้างตัวแบบพยากรณ์ Neural Networks จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 89

พหุน ปณ ฑิโต ชีเว

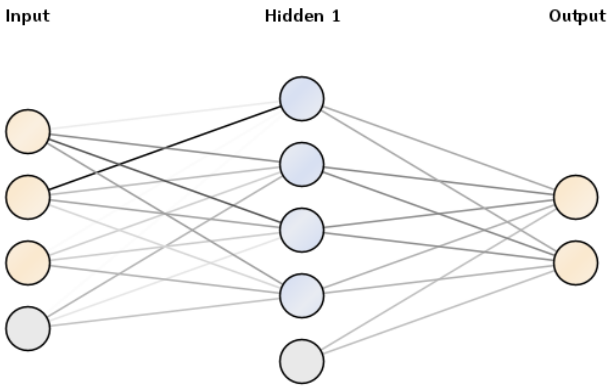


ภาพประกอบ 89 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีเพิ่มตัวแปร (Forward selection)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Neural Networks ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) แสดงในภาพที่ 90



ภาพประกอบ 90 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ แสดงดังตาราง 41

ตาราง 41 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

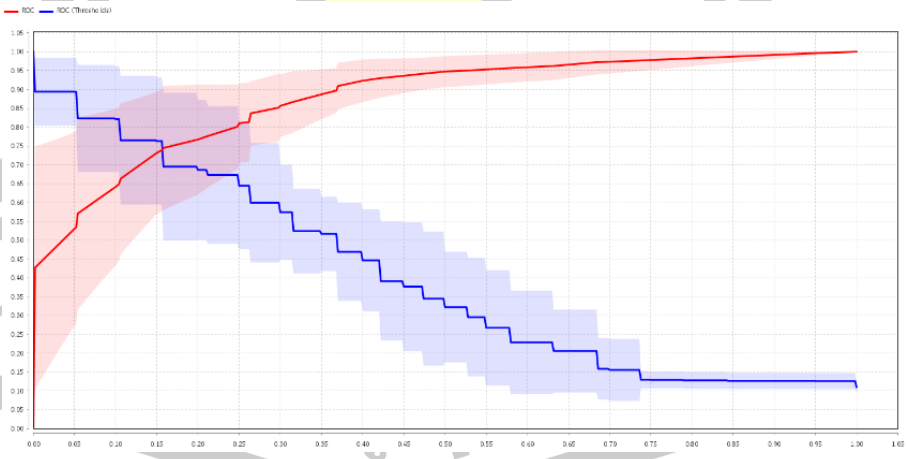
	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)
Sentiment	-0.432	2.707	-4.03	-2.283
Seats	5.947	1.613	-2.178	0.96
delivery	-0.082	-1.14	-1.216	-1.839
Bias	0.065	-1.766	-0.573	-1.45

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีเพิ่มตัวแปร (Forward selection)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.65% ตามรายละเอียดที่แสดงไว้ในตาราง 42

ตาราง 42 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

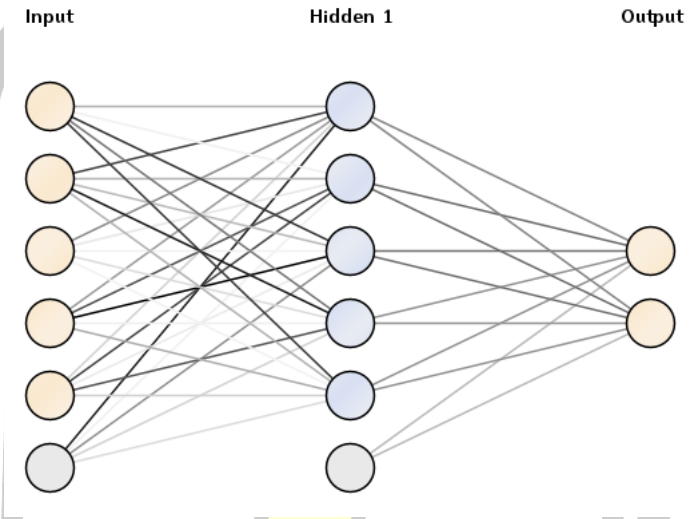
ตัวแบบโครงข่ายประสาทเทียม		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	146	43	77.25%
	คะแนนรีวิวอยู่ในระดับต่ำ	46	225	83.03%
	class recall	76.04%	83.96%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.92 ค่าความจำเพาะ (specificity) ร้อยละ 76.13 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 83.14 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 77.39 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.46 และมีค่า AUC เท่ากับ 0.878 ดังแสดงในภาพประกอบ 91



ภาพประกอบ 91 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) แสดงในภาพที่ 92



ภาพประกอบ 92 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงแสดงดังตาราง 43

ตาราง 43 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)
Sentiment:	-6.195	1.358	-2.859	-6.489	-4.45
Price:	-5.747	-3.492	0.803	-4.955	0.289
alcohol:	-1.453	-0.371	-1.578	2.928	-2.375
Reviews:	-2.672	1.65	7.301	-3.013	0.916
Group:	-3.796	1.223	4.335	0.767	-1.343

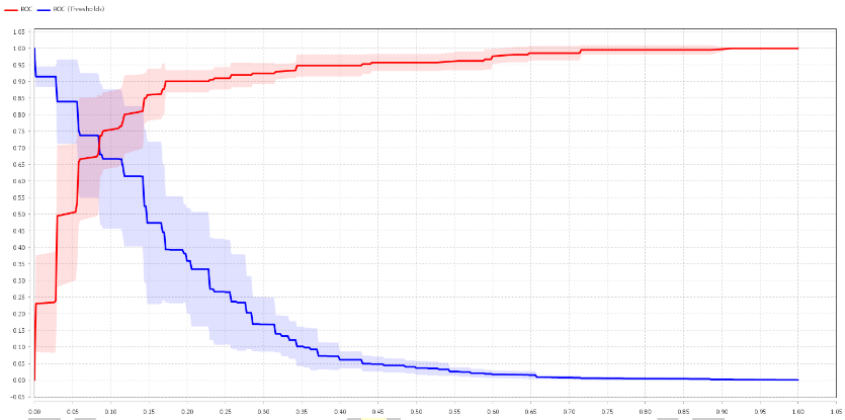
	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)
Bias:	-1.713	-0.664	-0.941	-1.348	-0.196

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีเพิ่มตัวแปร (Forward selection)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 85.92% ตามรายละเอียดที่แสดงไว้ในตาราง 44

ตาราง 44 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

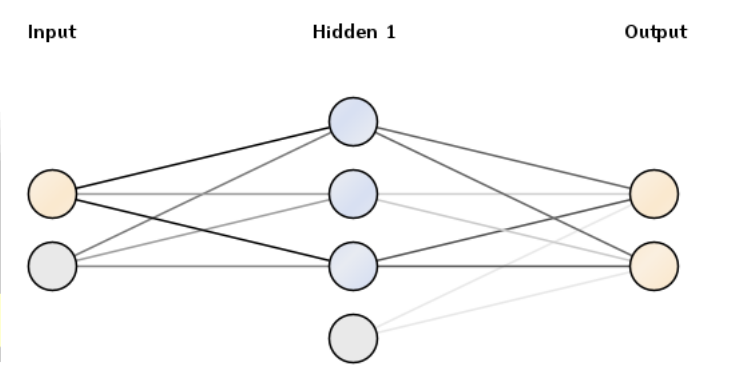
		ผลที่เกิดขึ้นจริง		
ตัวแบบโครงข่ายประสาทเทียม		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	304	34	89.94%
	คะแนนรีวิวอยู่ในระดับต่ำ	45	178	79.82%
	class recall	87.11%	83.96%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.96 ค่าความจำเพาะ (specificity) ร้อยละ 87.15 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 80.94 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 90.36 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 81.76 และมีค่า AUC เท่ากับ 0.909 ดังแสดงในภาพประกอบ 93



ภาพประกอบ 93 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) แสดงในภาพที่ 94



ภาพประกอบ 94 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง แสดงดังตาราง 45

ตาราง 45 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)
Sentiment:	3.352	3.256	3.388
Bias:	-0.855	-0.728	-0.859

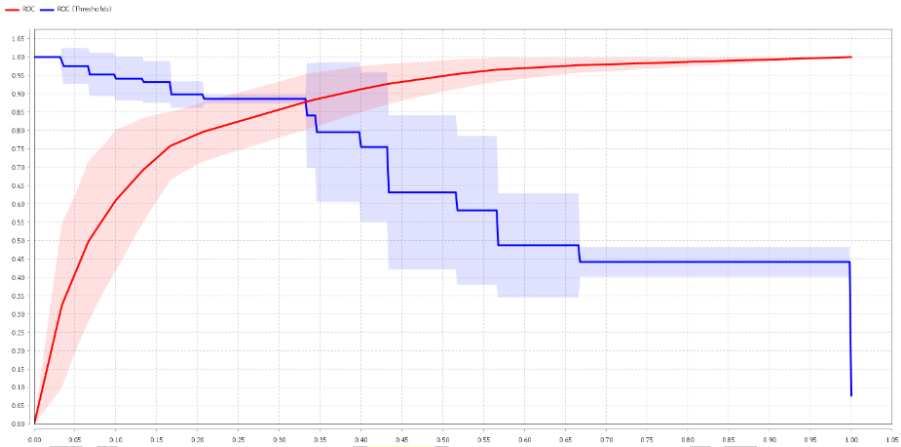
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มตัวแปร (Forward selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.10% ตามรายละเอียดที่แสดงไว้ในตาราง 46

ตาราง 46 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

		ผลที่เกิดขึ้นจริง		
ตัวแบบโครงข่ายประสาทเทียม		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	245	70	77.78%
	คะแนนรีวิวอยู่ในระดับต่ำ	53	250	82.51%
	class recall	82.21%	78.12%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 78.12 ค่าความจำเพาะ (specificity) ร้อยละ 82.25 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 83.99 ค่าการทำนายผลลบ (negative predictive value)

ร้อยละ 78.82 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 80.20 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 95

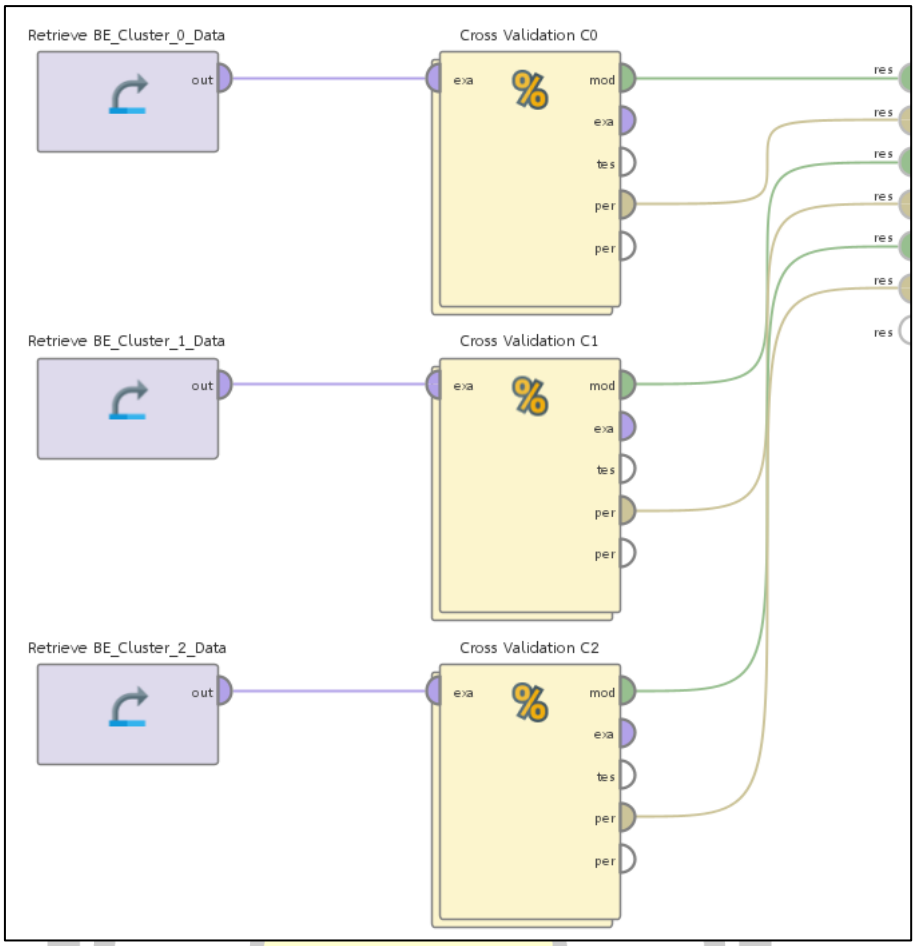


ภาพประกอบ 95 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

3.2 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination)

ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) แสดงดังภาพประกอบ 96

พหุน ปณ ฑิโต ชีเว

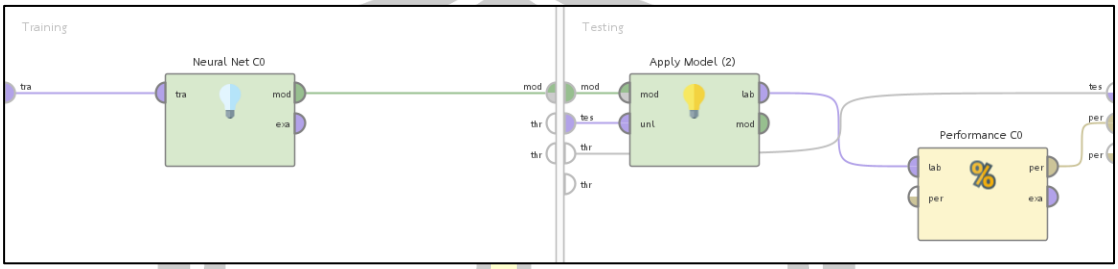


ภาพประกอบ 96 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination)

จากภาพประกอบ 96 แสดงขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) โดยมีขั้นตอน ดังนี้

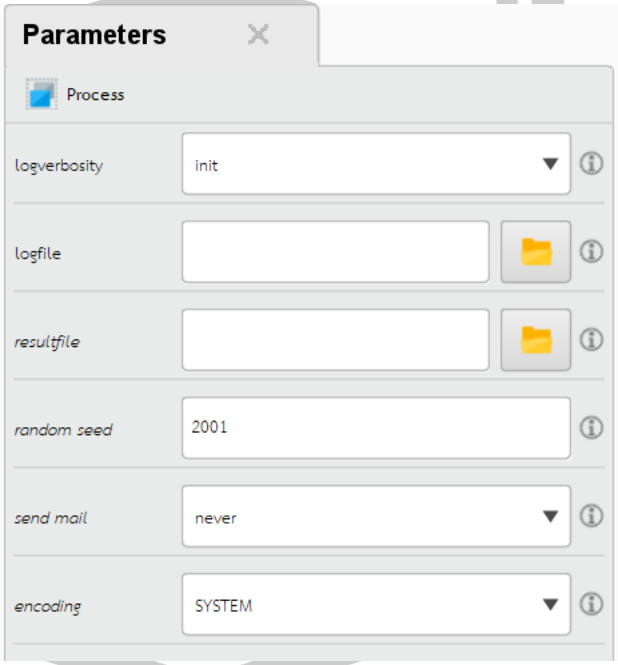
1. Retrieve_BE_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียมจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) ทั้ง 3 กลุ่ม
2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Neural Networks ที่สร้างขึ้น

ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination) แสดงดังภาพประกอบ 97



ภาพประกอบ 97 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีลดตัวแปร (Backward Elimination)

2.1 Neural Networks : เป็นการสร้างตัวแบบพยากรณ์ Neural Networks จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 98

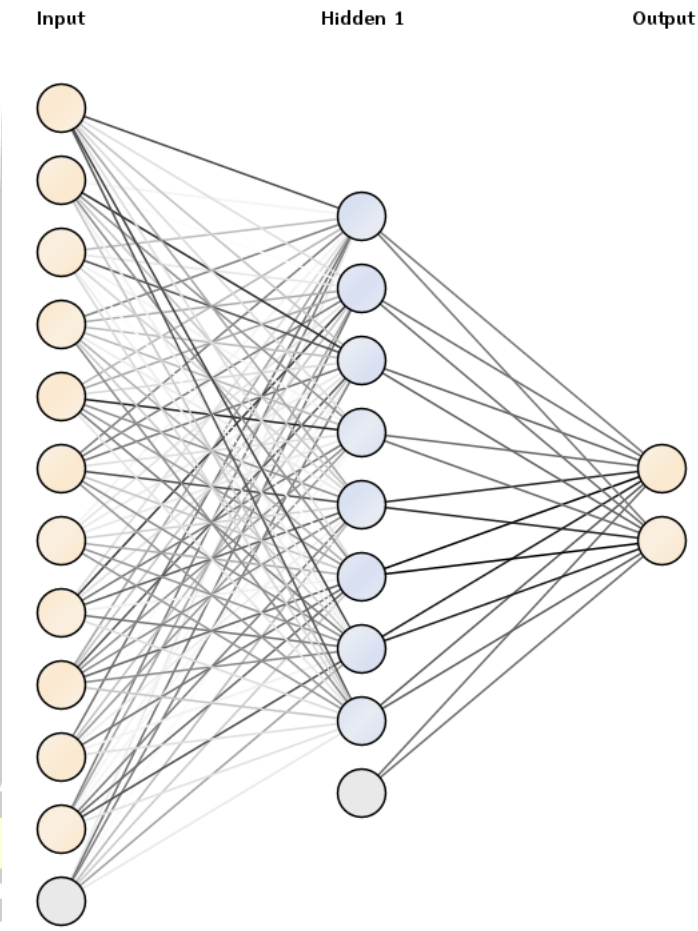


ภาพประกอบ 98 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีลดตัวแปร (Backward Elimination)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Neural Networks ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) แสดงในภาพที่ 99



ภาพประกอบ 99 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำสามารถแสดงได้ดังตาราง 47

ตาราง 47 ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)	Node 6 (Sigmoid)	Node 7 (Sigmoid)	Node 8 (Sigmoid)
Price:	-0.54	-2.72	1.81	2.39	0.4	0.16	-0.32	0.97
Seats:	-3.41	-2.15	4.04	-2.17	-2.41	-0.45	0.57	-2.18
Check-in:	1.77	2.57	-2.89	3.54	1.75	0.72	-1.26	-1.89
Like:	3.24	2.85	-0.48	-0.73	-1.57	1.42	-0.73	-2.87
Population Density:	2.2	0.25	-1.22	3.31	-5.36	0.72	6.3	5
Sentiment:	-1.28	0.67	6.11	0.82	5.19	0.73	0.18	8.01
delivery:	0.45	0	1.28	-0.96	-1.67	0.16	2.32	1.9
Group:	2.93	2.03	-0.17	-1.24	-1.32	0.86	0.52	0.78
alcohol:	2.73	1.76	1.38	1.79	-2.59	0.45	-0.07	-0.64
parking:	1.82	-0.92	4.89	-0.02	-1.95	0.17	-2.05	-2.87
SocialMedia :	3.45	1.58	1.16	-4.07	-3.84	0.49	4.01	0.08
Bias:	-3.74	-3.47	-6.43	-1.5	-2.08	-1.69	-2.64	0.29

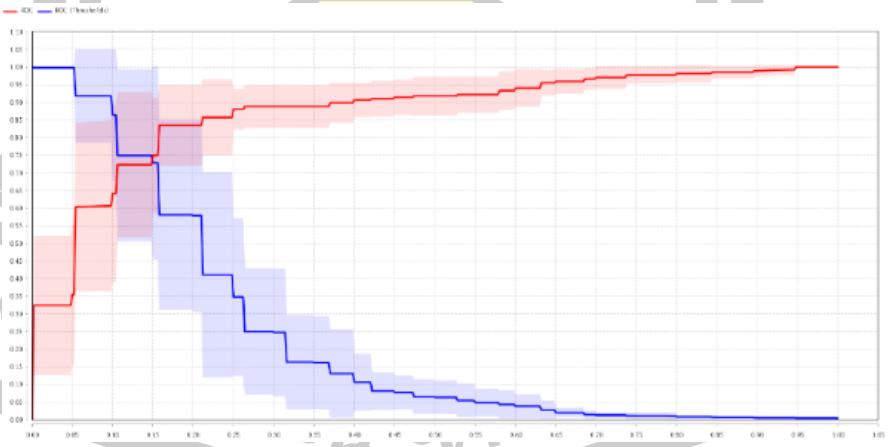
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีลดตัวแปร (Backward Elimination)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 84.13% ตามรายละเอียดที่แสดงไว้ในตาราง 48

พหุน ปณฺ ฑิโต ชีเว

ตาราง 48 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

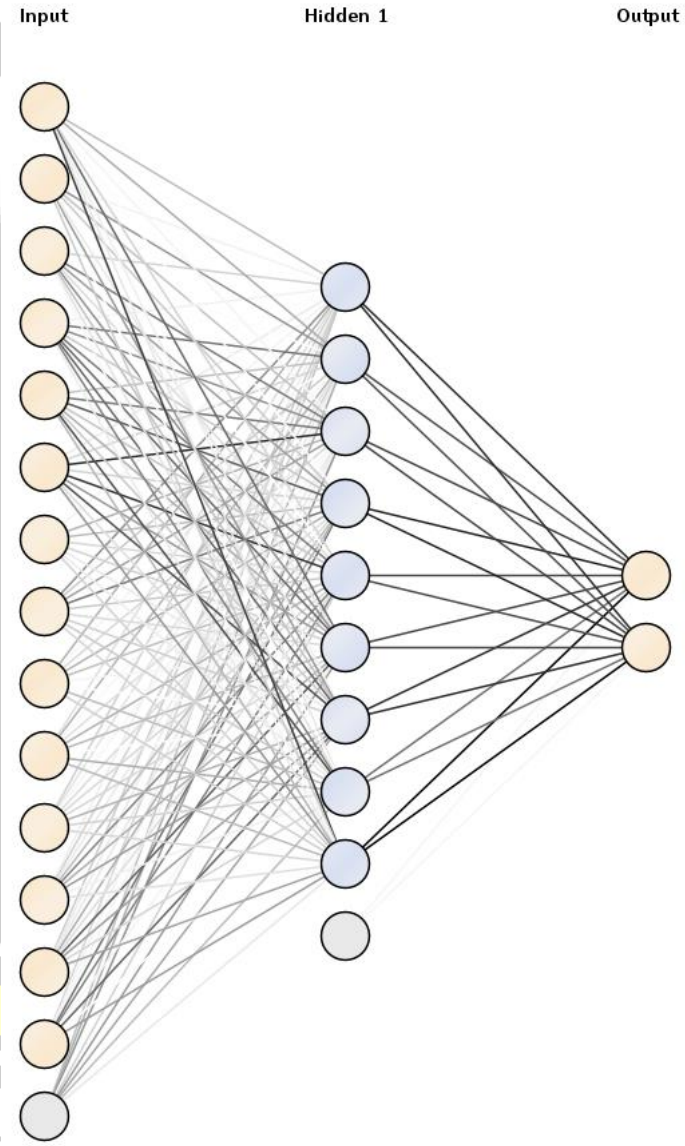
ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	159	40	79.90%
	คะแนนรีวิว อยู่ในระดับต่ำ	33	228	87.36%
	class recall	82.81%	85.07%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 85.04 ค่าความจำเพาะ (specificity) ร้อยละ 82.82 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 87.44 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 80.50 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 86.10 และมีค่า AUC เท่ากับ 0.873 ดังแสดงในภาพประกอบ 100



ภาพประกอบ 100 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) แสดงในภาพที่ 101



ภาพประกอบ 101 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงสามารถแสดงได้ดังตาราง 49

ตาราง 49 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

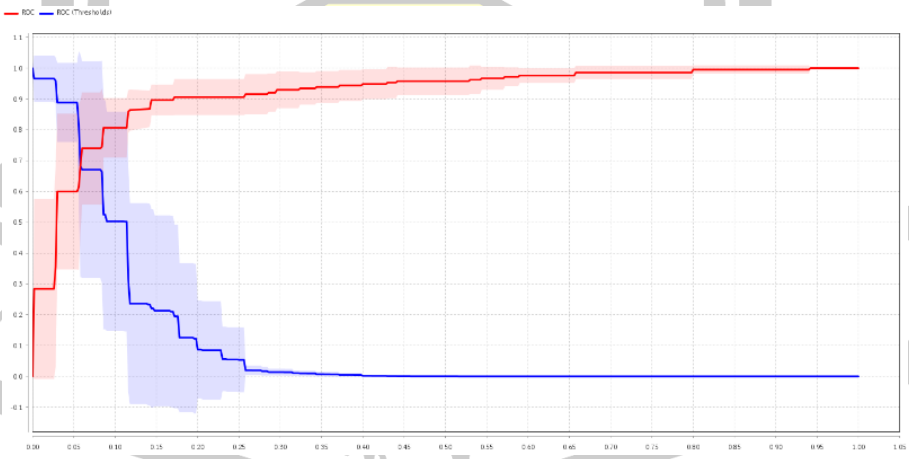
	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8	Node 9
Price:	-1.542	-0.649	-0.376	2.052	-0.303	-1.231	-2.413	0.431	-1.934
Reviews:	0.702	1.279	0.317	-1.548	-0.169	-1.674	0.504	-1.067	-0.193
Check-in:	-0.494	-0.498	-0.753	1.795	-0.154	-2.739	2.878	2.526	-1.416
Like:	0.363	-0.194	-1.314	-2.423	-0.781	-3.454	1.322	3.523	1.212
Population Density:	-0.991	2.415	0.153	3.11	3.161	-1.389	4.234	2.576	1.342
Sentiment:	-3.331	-6.443	-4.038	-1.008	-1.491	-5.883	-4.893	-4.221	-1.668
delivery:	1.244	-2.849	1.455	-2.4	-0.868	-2.4	-1.203	-0.454	0.068
Group:	-2.859	2.498	-0.154	0.309	-2.586	-2.367	-3.404	1.267	0.121
alcohol:	0.063	-1.314	1.689	-1.155	0.295	-2.876	0.407	3.3	-0.847
parking:	-2.98	1.132	0.968	0.447	0.374	-1.201	0.115	0.303	0.508
creditCard:	1.07	1.568	0.352	0.995	-0.824	-0.808	1.903	-1.487	-0.831
Website:	1.895	-3.134	1.176	0.783	0.372	3.091	-1.065	-3.384	1.988
SocialMedia:	0.094	2.121	0.605	-2.12	1.22	1.604	-3.148	0.183	-2.263
WiFi:	-1.436	-0.745	1.466	1.372	0.822	0.701	1.079	-1.245	-0.698
Bias:	-2.589	-2.395	-0.28	0.594	-0.841	-1.829	-0.802	-0.496	-1.258

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีลดตัวแปร (Backward Elimination)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 87.51% ตามรายละเอียดที่แสดงไว้ในตาราง 50

ตาราง 50 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

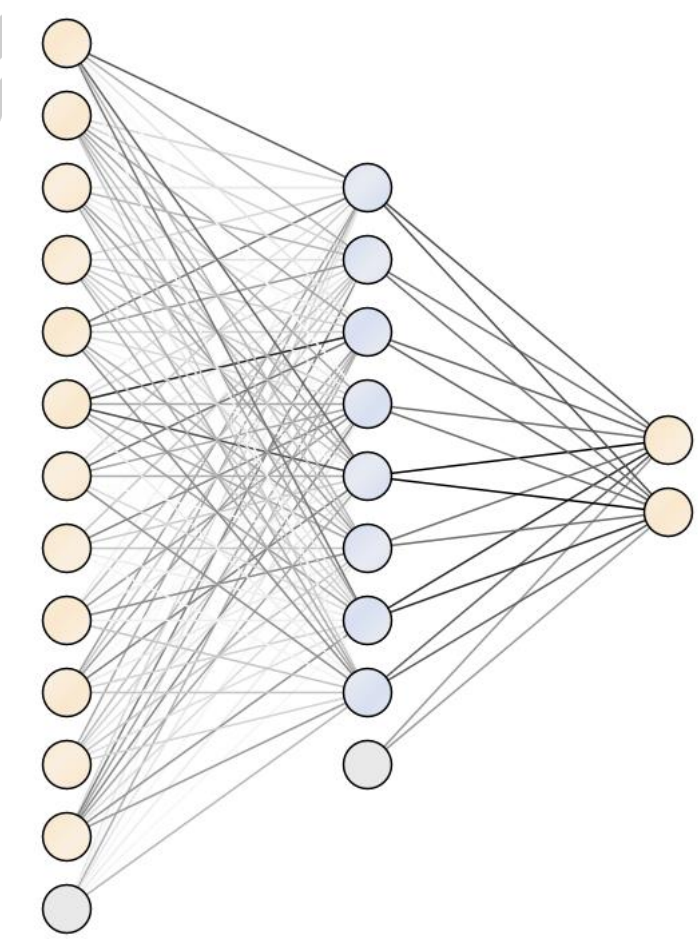
ตัวแบบโครงข่ายประสาทเทียม		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	318	39	89.08%
	คะแนนรีวิว อยู่ในระดับต่ำ	31	173	84.80%
	class recall	91.12%	81.60%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 81.60 ค่าความจำเพาะ (specificity) ร้อยละ 91.09 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 85.34 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 89.43 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.02 และมีค่า AUC เท่ากับ 0.918 ดังแสดงในภาพประกอบ 102



ภาพประกอบ 102 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) แสดงในภาพที่ 103



ภาพประกอบ 103 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลางสามารถแสดงได้ดังตาราง 51

ตาราง 51 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

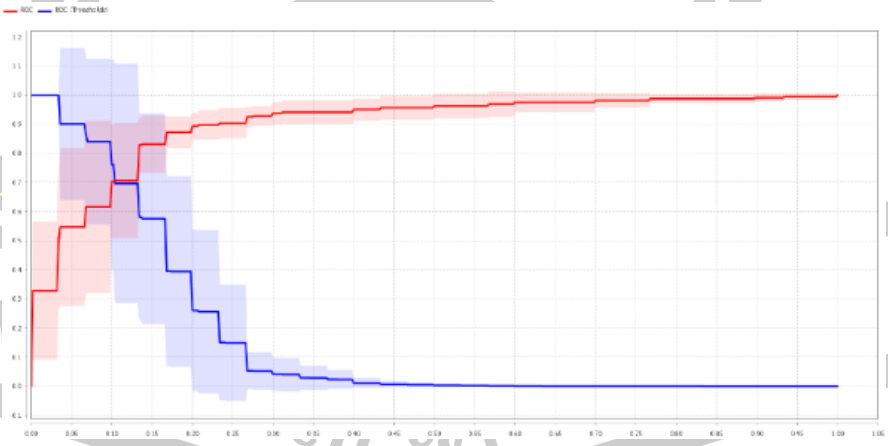
	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
Price:	0.295	5.055	2.299	0.286	2.382	1.28	-0.114	4.754
Seats:	0.543	0.861	-2.803	1.494	-4.185	-0.631	-3.189	-2.865
Reviews:	-2.461	-0.415	-0.267	1.793	-1.914	-1.936	2.583	0.663
Check-in:	-2.937	-0.592	-1.613	2.219	2.717	-2.436	0.523	1.184
Population Density:	-0.445	3.04	-1.367	-2.982	2.926	2.754	3.439	3.676
Sentiment:	3.752	-0.025	-0.225	5.923	4.13	5.117	-0.301	5.357
delivery:	1.654	-3.542	1.457	2.964	3.893	-3.356	-2.713	1.774
alcohol:	-2.224	1.237	0.447	0.391	0.867	3.191	1.358	-1.159
parking:	-3.674	0.786	-1.158	-2.756	2.881	2.291	-1.09	-1.345
creditCard:	-3.287	-2.66	0.508	3.264	1.961	-0.593	-4.447	3.399
Website:	0.529	-4.965	0.949	3.009	0.762	-0.027	-4.795	-3.325
SocialMedia:	-0.138	0.402	1.596	-3.47	2.981	3.533	-1.581	-0.765
Bias:	-4.131	-2.338	-5.949	-2.693	-0.206	-0.334	-2.065	-1.913

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีลดตัวแปร (Backward Elimination) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 85.46% ตามรายละเอียดที่แสดงไว้ในตาราง 52

ตาราง 52 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	255	47	84.44%
	คะแนนรีวิว อยู่ในระดับต่ำ	43	273	86.39%
	class recall	85.57%	85.31%	

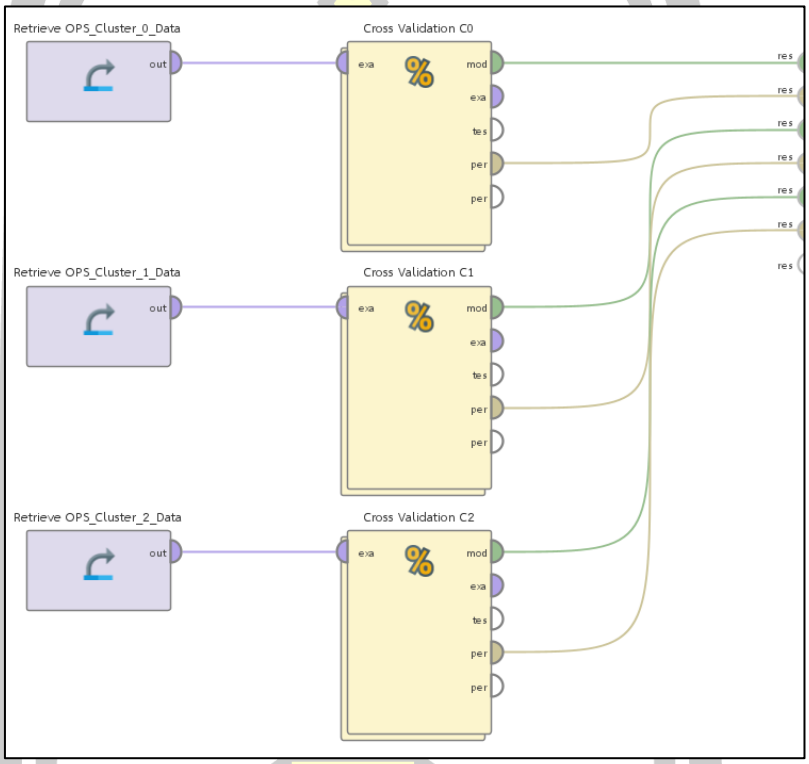
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 85.31 ค่าความจำเพาะ (specificity) ร้อยละ 85.63 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 86.62 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 84.49 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 85.89 และมีค่า AUC เท่ากับ 0.901 ดังแสดงในภาพประกอบ 104



ภาพประกอบ 104 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

3.3 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) แสดงดังภาพประกอบ 105



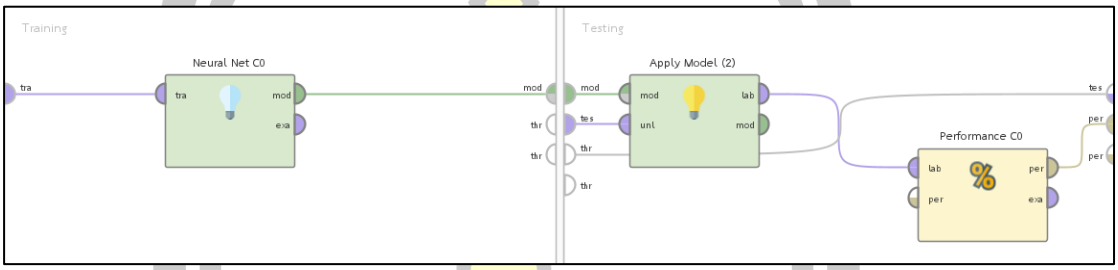
ภาพประกอบ 105 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

จากภาพประกอบ 105 แสดงขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) โดยมีขั้นตอน ดังนี้

1. Retrieve_BE_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียมจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ทั้ง 3 กลุ่ม

2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Neural Networks ที่สร้างขึ้น

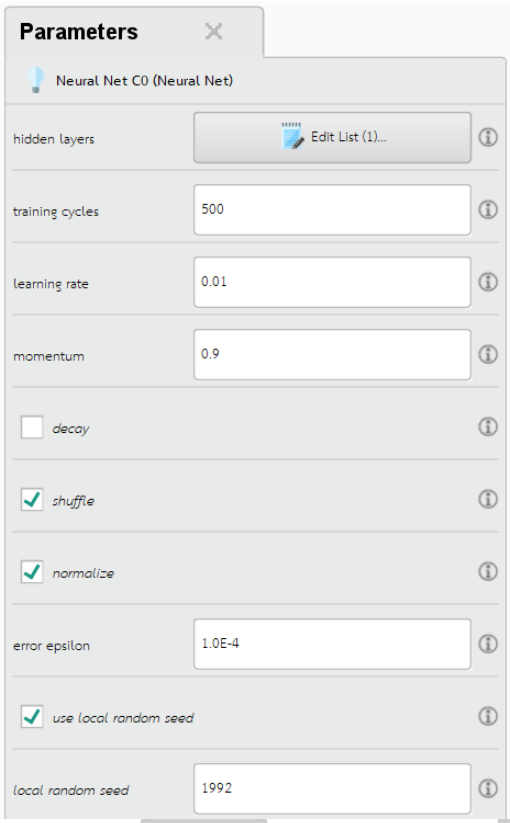
ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) แสดงดังภาพประกอบ 106



ภาพประกอบ 106 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

2.1 Neural Networks : เป็นการสร้างตัวแบบพยากรณ์ Neural Networks จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 107

พหุน ปณ ฑิโต ชีเว

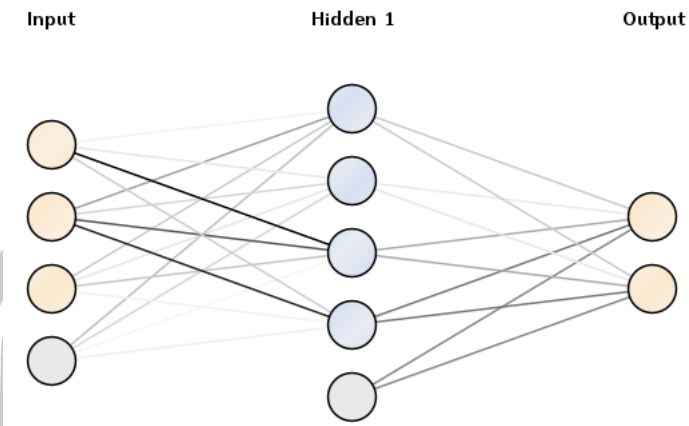


ภาพประกอบ 107 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Neural Networks ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) แสดงในภาพที่ 108



ภาพประกอบ 108 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำสามารถแสดงได้ดังตาราง 53

ตาราง 53 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)
Sentiment:	-6.195	1.358	-2.859	-6.489	-4.45
Price:	-5.747	-3.492	0.803	-4.955	0.289
alcohol:	-1.453	-0.371	-1.578	2.928	-2.375
Reviews:	-2.672	1.65	7.301	-3.013	0.916
Group:	-3.796	1.223	4.335	0.767	-1.343
Bias:	-1.713	-0.664	-0.941	-1.348	-0.196

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วย

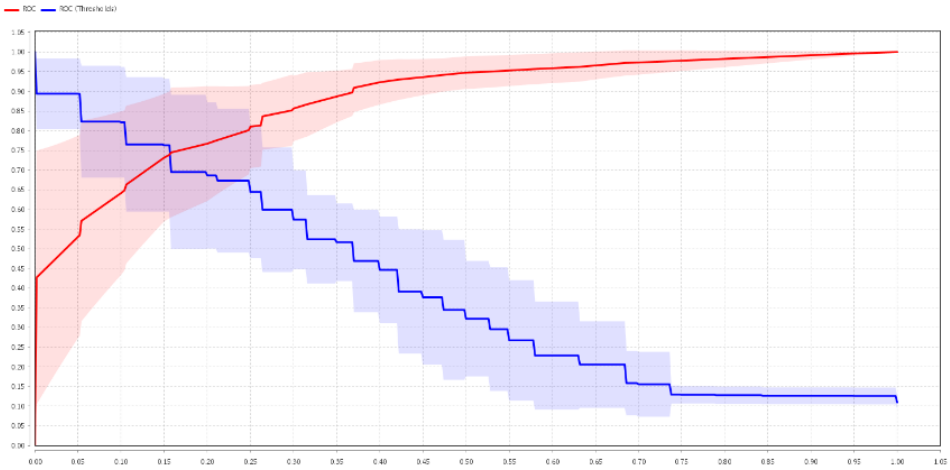
วิธีเพิ่มและลดตัวแปร (Optimize Selection)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.65% ตามรายละเอียดที่แสดงไว้ในตาราง 54

ตาราง 54 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	146	43	77.25%
	คะแนนรีวิว อยู่ในระดับต่ำ	46	225	83.03%
	class recall	76.04%	83.96%	

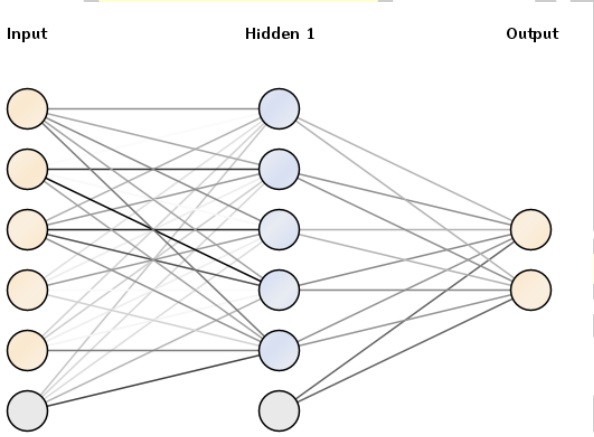
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.92 ค่าความจำเพาะ (specificity) ร้อยละ 76.13 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 83.14 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 77.39 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.46 และมีค่า AUC เท่ากับ 0.878 ดังแสดงในภาพประกอบ 109

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 109 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) แสดงในภาพที่ 110



ภาพประกอบ 110 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงสามารถแสดงได้ตาราง 55

ตาราง 55 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)
Sentiment:	-6.195	1.358	-2.859	-6.489	-4.45
Price:	-5.747	-3.492	0.803	-4.955	0.289
alcohol:	-1.453	-0.371	-1.578	2.928	-2.375
Reviews:	-2.672	1.65	7.301	-3.013	0.916
Group:	-3.796	1.223	4.335	0.767	-1.343
Bias:	-1.713	-0.664	-0.941	-1.348	-0.196

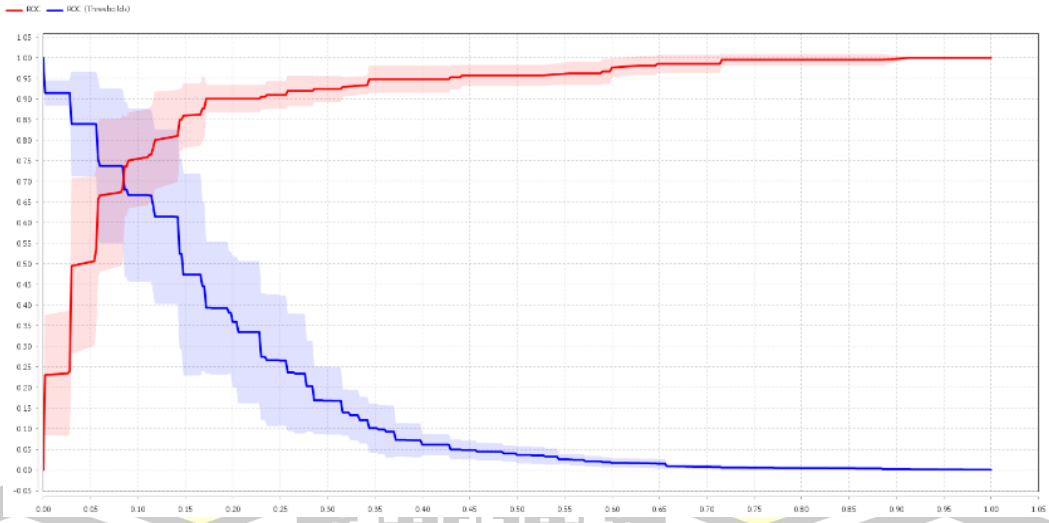
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks)ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection)ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 85.92% ตามรายละเอียดที่แสดงไว้ในตาราง 56

ตาราง 56 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	304	34	89.94%

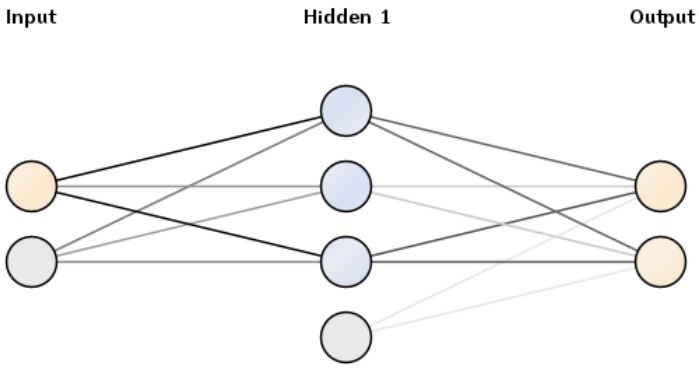
ตัวแบบโครงข่ายประสาทเทียม		ผลที่เกิดขึ้นจริง		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
	คะแนนรีวิว อยู่ในระดับต่ำ	45	178	79.82%
	class recall	87.11%	83.96%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 83.96 ค่าความจำเพาะ (specificity) ร้อยละ 87.15 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 80.94 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 90.36 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 81.76 และมีค่า AUC เท่ากับ 0.909 ดังแสดงในภาพประกอบ 111



ภาพประกอบ 111 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) แสดงในภาพที่ 112



ภาพประกอบ 112 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลางสามารถแสดงได้ดังตาราง 57

ตาราง 57 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

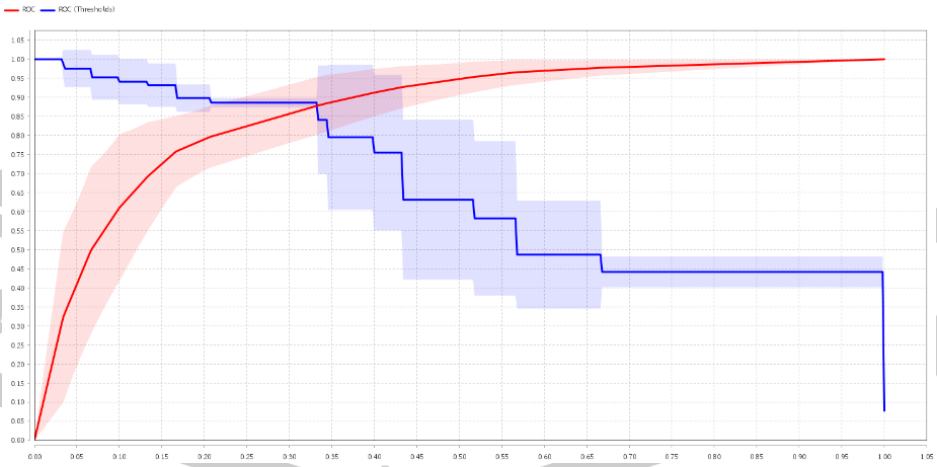
	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)
Sentiment:	3.352	3.256	3.388
Bias:	-0.855	-0.728	-0.859

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีเพิ่มและลดตัวแปร (Optimize Selection) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ(Accuracy) ของแบบจำลองมีเท่ากับ 80.10% ตามรายละเอียดที่แสดงไว้ในตาราง 58

ตาราง 58 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	245	70	77.78%
	คะแนนรีวิว อยู่ในระดับต่ำ	53	250	82.51%
	class recall	82.21%	78.12%	

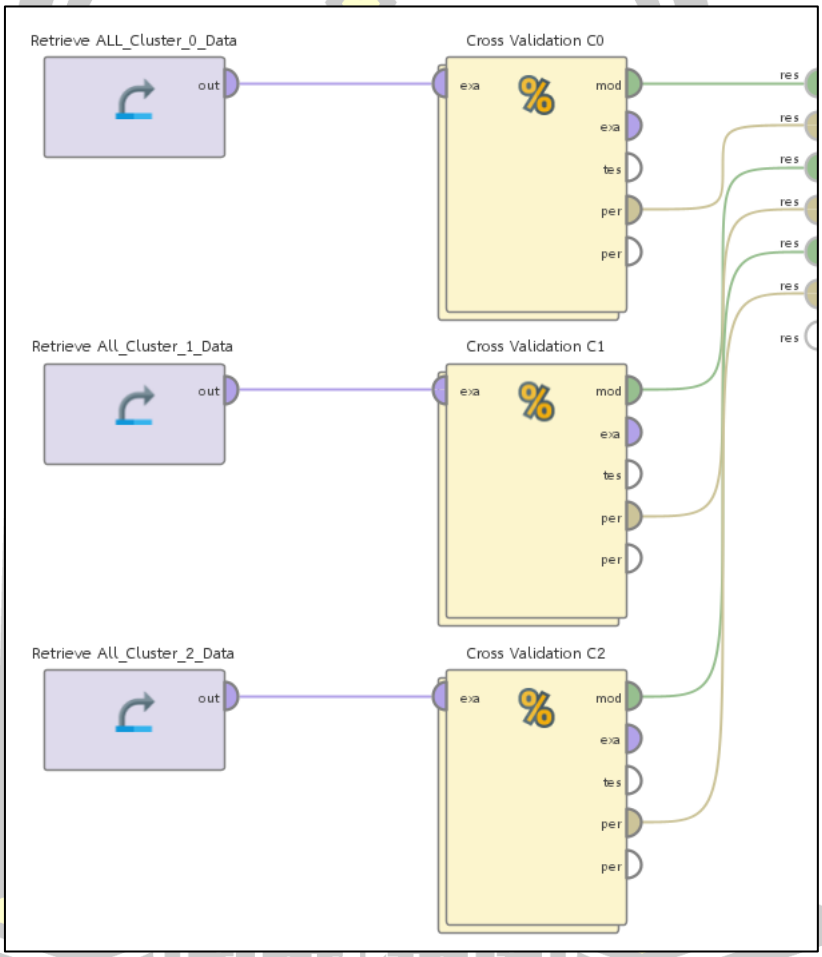
ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 78.12 ค่าความจำเพาะ (specificity) ร้อยละ 82.25 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 83.99 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 78.82 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 80.20 และมีค่า AUC เท่ากับ 0.864 ดังแสดงในภาพประกอบ 113



ภาพประกอบ 113 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีเพิ่มและลดตัวแปรในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

3.4 การสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) แสดงดังภาพประกอบ 114



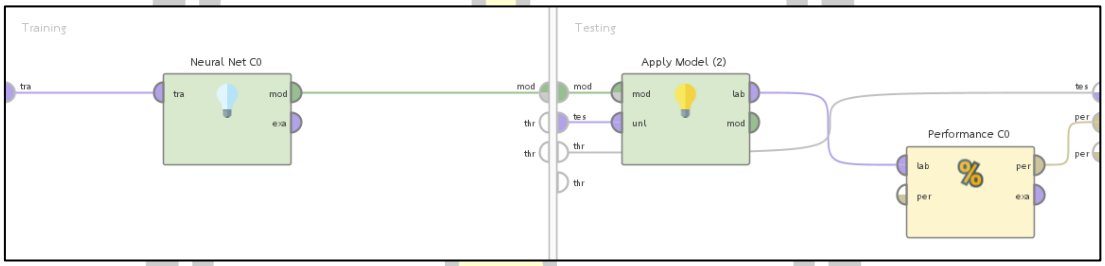
ภาพประกอบ 114 ขั้นตอนการสร้างตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

จากภาพประกอบ 114 แสดงขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) โดยมีขั้นตอน ดังนี้

1. Retrieve_All_Cluster_x_Data : เชื่อมต่อกับชุดข้อมูลสำหรับการพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียมจากขั้นตอนการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ทั้ง 3 กลุ่ม

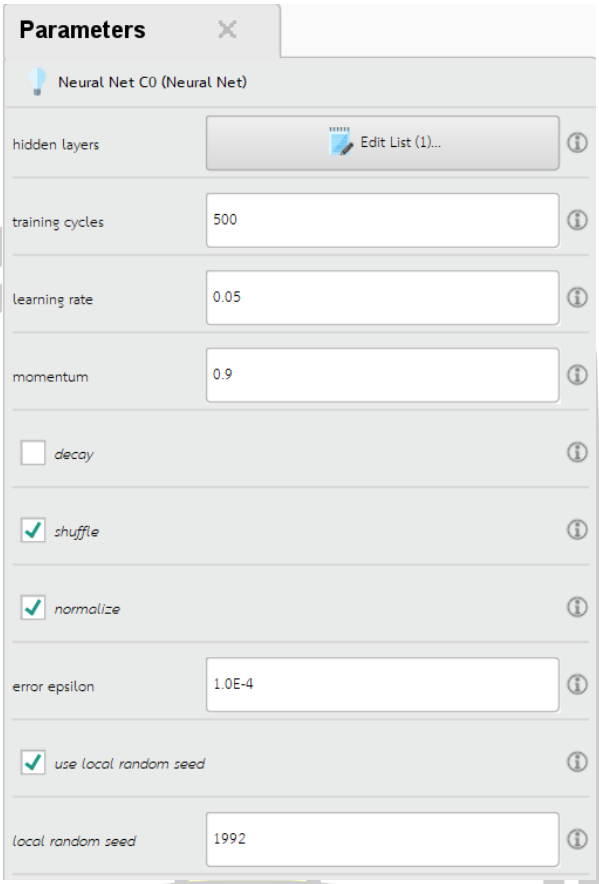
2. Cross Validation : ตรวจสอบประสิทธิภาพของตัวแบบพยากรณ์ Neural Networks ที่สร้างขึ้น

ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) แสดงดังภาพประกอบ 115



ภาพประกอบ 115 ขั้นตอนการสร้างตัวแบบโครงข่ายประสาทเทียมภายใน Cross Validation ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

2.1 Neural Networks : เป็นการสร้างตัวแบบพยากรณ์ Neural Networks จากชุดข้อมูลฝึกฝน (Training dataset) เพื่อใช้สำหรับพยากรณ์ข้อมูลต่อไป โดยกำหนดค่าพารามิเตอร์ ดังภาพประกอบ 116

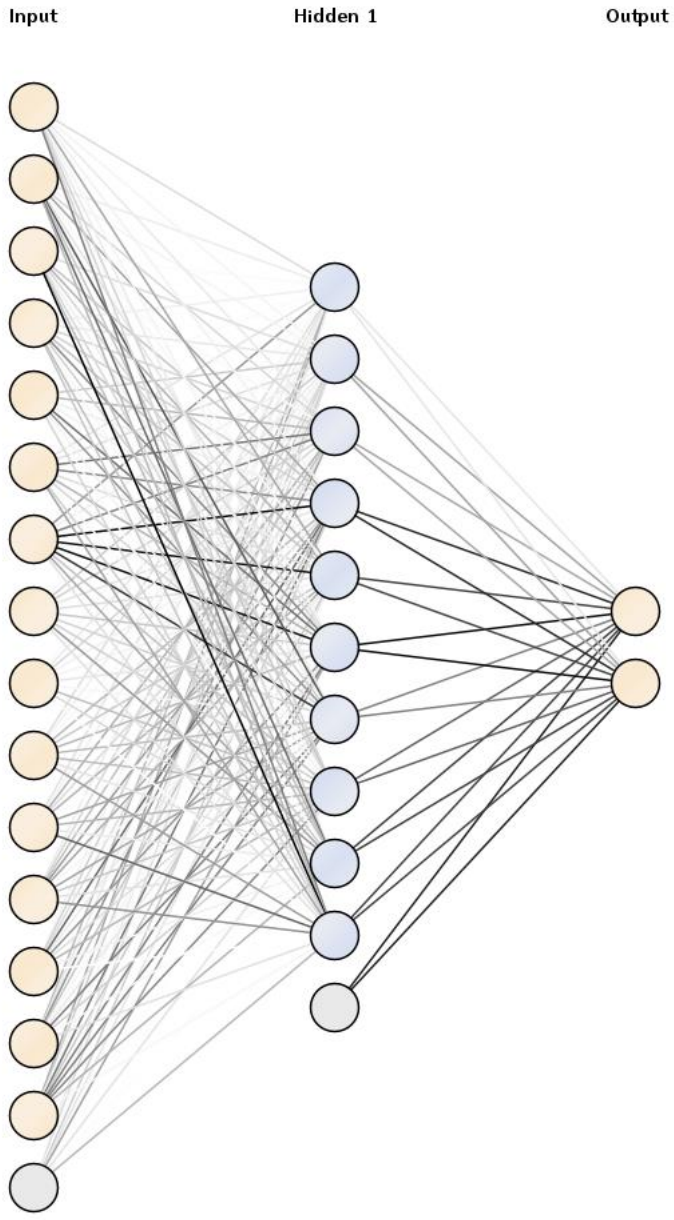


ภาพประกอบ 116 การกำหนดค่าพารามิเตอร์สำหรับ Neural Networks ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)

2.2 Apply Model : ตัวแบบพยากรณ์ที่ได้นำมาใช้กับชุดข้อมูลทดสอบ (Testing Dataset)

2.3 Performance : ตรวจสอบประสิทธิภาพของตัวแบบ Neural Networks ที่ได้สร้างขึ้น

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) แสดงในภาพที่ 117



ภาพประกอบ 117 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำสามารถแสดงได้ดังตาราง 59

ตาราง 59 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)	Node 6 (Sigmoid)	Node 7 (Sigmoid)
Price:	-1.676	0.512	1.054	-0.654	-4.324	-0.503	-1.006
Seats:	0.907	0.471	1.922	4.361	0.355	1.341	-8.249
Reviews:	0.198	1.704	0.250	2.820	-3.710	1.973	-0.736
Check-in:	-0.760	1.128	2.091	-2.238	-2.977	5.728	2.301
Like:	0.196	2.457	-3.511	-0.772	-0.640	7.051	5.203
Population Density:	-1.278	-1.598	6.720	5.706	1.722	-3.821	-2.072
Sentiment:	-5.108	0.741	6.306	12.387	11.745	-11.987	-9.915
delivery:	-0.251	2.464	-2.255	-1.321	1.588	-2.751	-2.250
Group:	0.291	-1.057	-1.927	-0.661	0.532	-1.188	0.419
alcohol:	-1.614	-0.694	0.313	-3.458	2.695	3.582	-3.255
parking:	1.400	-0.602	-3.923	1.828	3.770	-0.078	4.434
Credit Card:	0.894	1.982	3.118	4.948	4.304	-0.440	-4.435
Website:	-0.583	2.224	1.622	7.415	-0.992	5.198	1.639
Social Media:	-3.506	-1.084	-5.448	-2.888	0.449	1.024	-6.353
WiFi:	4.419	2.431	-0.396	6.768	2.646	-6.353	6.193
Bias:	1.007	-1.613	-0.080	2.404	1.231	-4.560	-0.108

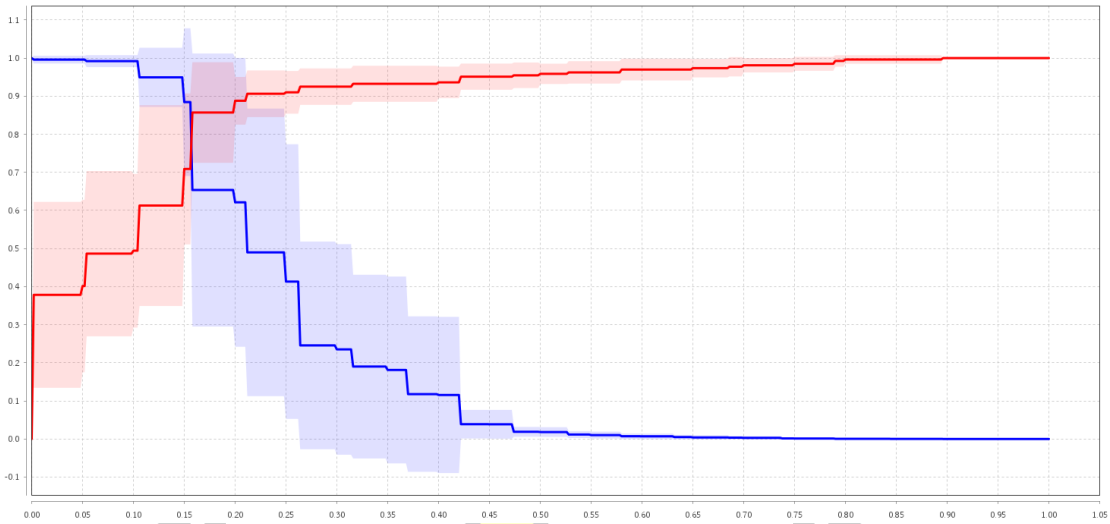
เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 85.22% ตามรายละเอียดที่แสดงไว้ในตาราง 60

ตาราง 60 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	153	29	84.07%
	คะแนนรีวิวอยู่ในระดับต่ำ	39	239	85.97%
	class recall	79.69%	89.18%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 89.13 ค่าความจำเพาะ (specificity) ร้อยละ 79.66 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 86.22 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 84.91 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 87.47 และมีค่า AUC เท่ากับ 0.886 ดังแสดงในภาพประกอบ 118

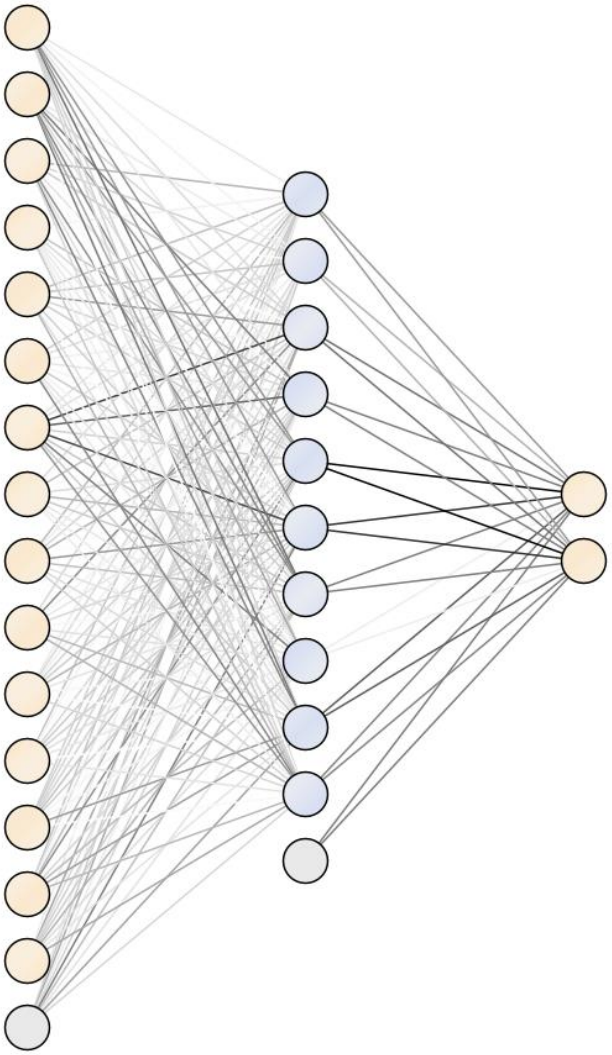
พหุน ปณฺฑิโต ชีเว



ภาพประกอบ 118 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) แสดงในภาพที่ 119

พหุน ปณ ฑิโต ชีเว



ภาพประกอบ 119 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

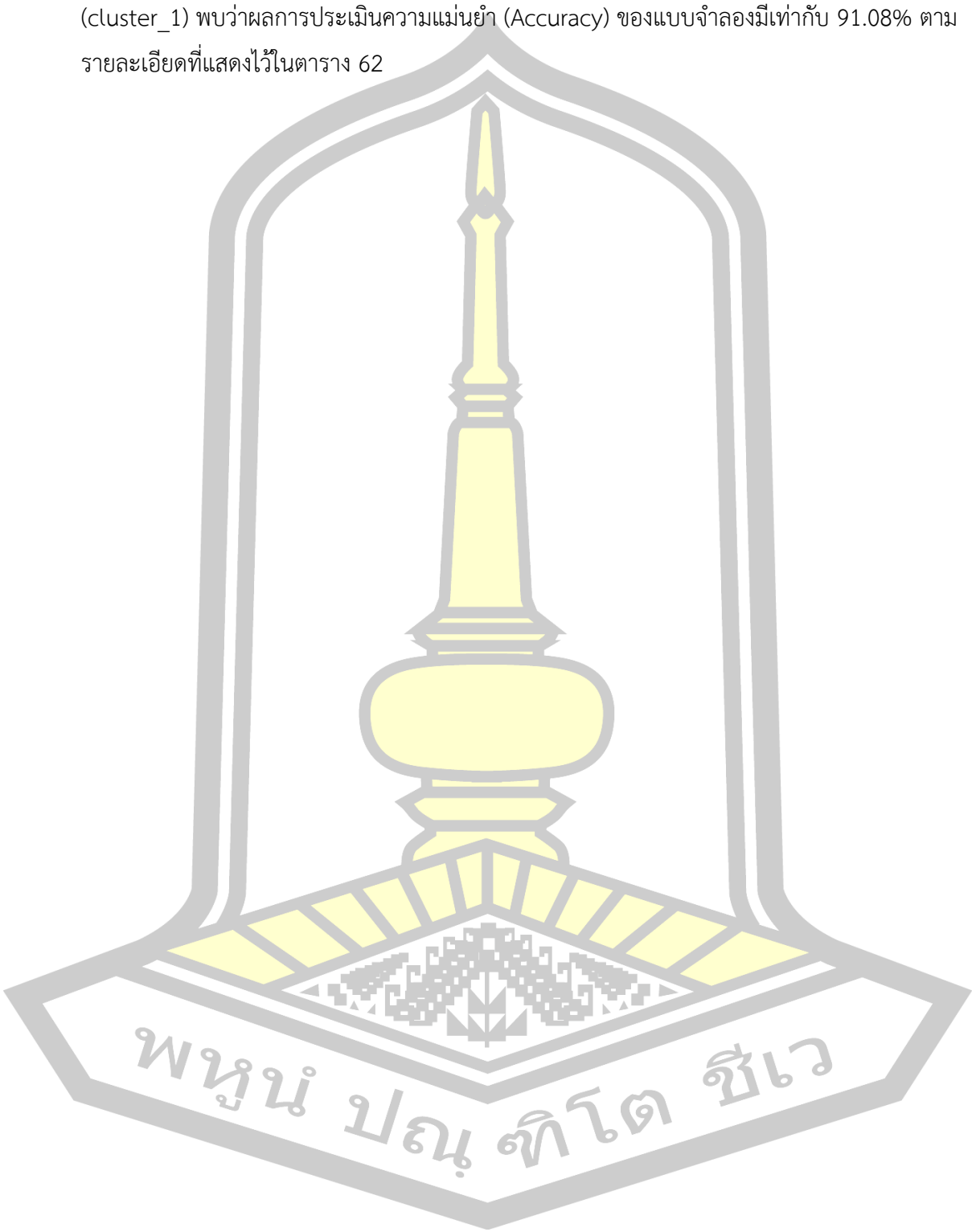
ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงสามารถแสดงได้ตาราง 61

ตาราง 61 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)	Node 6 (Sigmoid)	Node 7 (Sigmoid)	N (Si
Price:	-1.136	-0.578	1.743	-1.088	5.197	-4.843	-6.873	2.2
Seats:	0.015	0.965	-1.971	-6.219	-4.849	-1.576	5.817	-4.
Reviews:	3.501	2.040	-2.443	-0.197	-1.331	0.255	5.350	1.5
Check-in:	-0.645	-0.133	-0.071	-2.734	1.776	1.031	-1.063	-0.
Like:	3.073	2.808	-4.850	1.958	-2.392	0.230	-2.288	-0.
Population Density:	2.344	1.536	0.776	-1.249	0.491	-2.464	-0.562	-1.
Sentiment:	1.183	1.544	9.314	7.564	2.220	-8.777	2.748	6.1
delivery:	-1.671	-0.088	-2.963	-0.196	-1.737	-3.502	-2.515	1.1
Group:	-3.283	-0.837	-1.992	3.624	-0.159	-4.938	-3.069	1.4
alcohol:	1.938	-0.106	5.446	0.628	-1.136	1.938	0.350	1.4
parking:	-0.924	-1.808	0.817	-2.060	-2.347	-3.536	1.243	0.4
Credit Card:	-2.808	-0.586	0.167	1.904	-2.457	-2.193	1.516	-1.
Website:	-2.248	2.235	2.062	2.120	-2.871	2.350	-0.491	0.5
Social Media:	2.790	1.517	-0.286	-3.174	-5.346	0.136	1.906	-3.
WiFi:	-1.079	-0.155	4.248	1.816	-2.265	-2.314	-3.607	1.6
Bias:	-1.074	-2.421	-4.031	-0.898	-2.025	6.163	-3.145	-1.

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วย

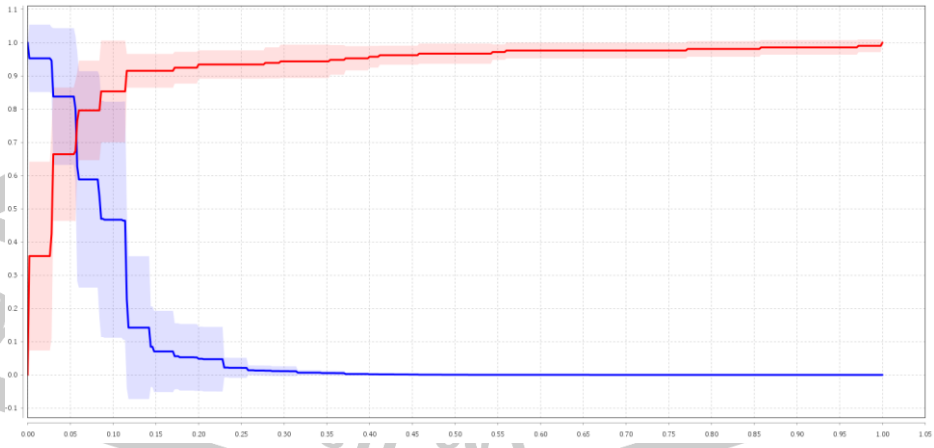
วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 91.08% ตามรายละเอียดที่แสดงไว้ในตาราง 62



ตาราง 62 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

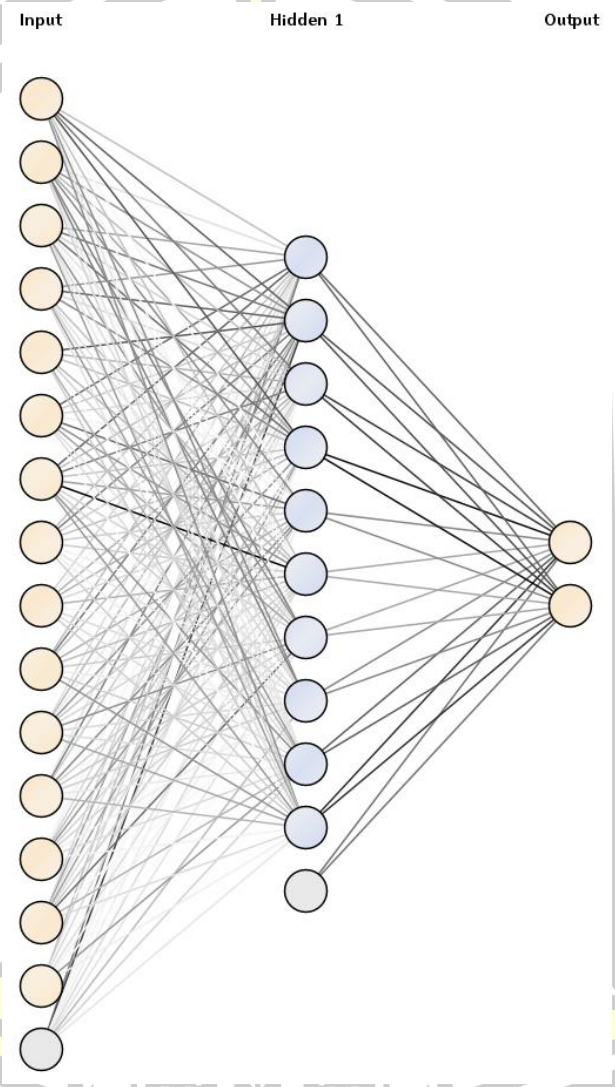
ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิว อยู่ในระดับสูง	คะแนนรีวิว อยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิว อยู่ในระดับสูง	325	26	92.59%
	คะแนนรีวิว อยู่ในระดับต่ำ	24	186	88.57%
	class recall	93.12%	87.74%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 87.71 ค่าความจำเพาะ (specificity) ร้อยละ 93.13 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 88.98 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 92.81 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 88.06 และมีค่า AUC เท่ากับ 0.929 ดังแสดงในภาพประกอบ 120



ภาพประกอบ 120 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

จากผลการประเมินประสิทธิภาพตัวแบบพยากรณ์โครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) แสดงในภาพที่ 121



ภาพประกอบ 121 โครงข่ายประสาทเทียมที่ได้จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลางสามารถแสดงได้ดังตาราง 63

ตาราง 63 แสดงค่าน้ำหนักตัวแปร Hidden 1 Layer จากการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

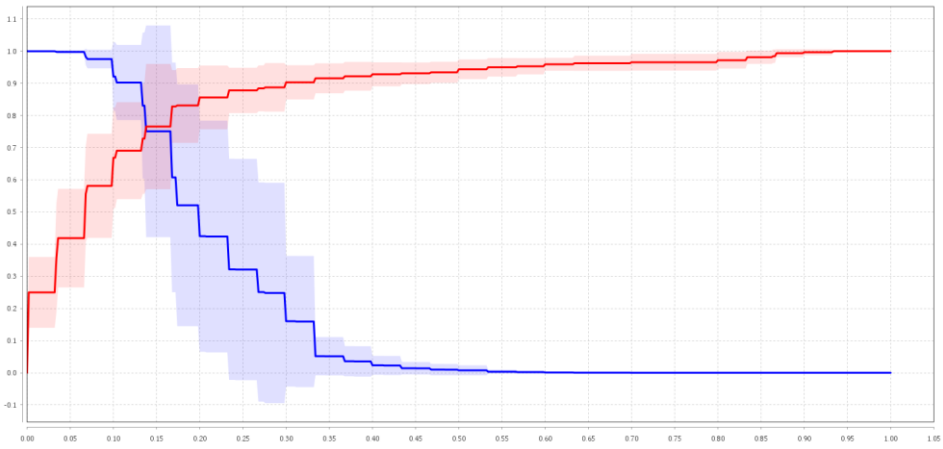
	Node 1 (Sigmoid)	Node 2 (Sigmoid)	Node 3 (Sigmoid)	Node 4 (Sigmoid)	Node 5 (Sigmoid)	Node 6 (Sigmoid)	Node 7 (Sigmoid)	N (Si
Price:	3.759	-9.442	7.839	-10.236	-0.474	3.55	0.544	6.7
Seats:	-1.821	6.372	-3.952	5.299	-6.969	-0.587	7.448	-7.
Reviews:	5.391	-8.14	2.58	8.564	-1.654	0.704	6.462	1.9
Check-in:	6.237	-9.374	2.334	6.921	0.021	-3.708	-1.676	2.4
Like:	0.286	9.988	4.359	3.826	-0.066	4.164	-2.769	0.0
Population Density:	-8.777	-3.095	-2.241	0.963	8.767	-6.432	-5.378	1.7
Sentiment:	-9.229	3.948	7.636	3.509	5.654	-15.11	-2.732	2.7
delivery:	3.679	-7.461	2.855	-3.663	-2.976	-3.047	-0.84	1.5
Group:	2.545	0.648	4.604	-4.38	-2.45	-4.526	-0.675	-2.
alcohol:	-3.966	-7.882	-3.703	0.254	4.819	3.917	-3.108	1.1
parking:	-2.039	4.122	-1.19	-6.39	-3.343	2.409	-6.587	2.2
Credit Card:	1.655	7.39	3.066	-0.047	-7.014	2.854	-3.256	2.8
Website:	-0.613	-0.69	-2.433	-3.686	-6.239	2.593	-2.529	-0.
Social Media:	-2.614	5.579	-1.603	-5.743	2.151	-3.013	-8.65	0.4
WiFi:	3.875	-2.613	-0.247	2.194	-0.093	-2.028	-0.838	-0.
Bias:	-7.779	-11.452	-0.639	-0.635	-1.069	-2.214	-1.678	-4.

เมื่อใช้เมทริกซ์ความสับสน (Confusion matrix) ในการวิเคราะห์ผลการทดสอบความแม่นยำของแบบจำลองพยากรณ์ด้วยตัวแบบโครงข่ายประสาทเทียม (Neural Networks) ด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) พบว่าผลการประเมินความแม่นยำ (Accuracy) ของแบบจำลองมีเท่ากับ 82.54% ตามรายละเอียดที่แสดงไว้ในตาราง 64

ตาราง 64 ผลการประเมินค่าความแม่นยำ (Accuracy) ของตัวแบบโครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบโครงข่ายประสาทเทียม		**ผลที่เกิดขึ้นจริง**		
		คะแนนรีวิวอยู่ในระดับสูง	คะแนนรีวิวอยู่ในระดับต่ำ	class precision
ผลจากการพยากรณ์	คะแนนรีวิวอยู่ในระดับสูง	241	51	82.53%
	คะแนนรีวิวอยู่ในระดับต่ำ	57	269	82.52%
	class recall	80.87%	84.06%	

ผลจากการทดสอบประสิทธิภาพของตัวแบบพยากรณ์พบว่า ค่าความไว (sensitivity) ร้อยละ 84.06 ค่าความจำเพาะ (specificity) ร้อยละ 80.91 ค่าการทำนายผลบวก (positive predictive value) ร้อยละ 82.71 ค่าการทำนายผลลบ (negative predictive value) ร้อยละ 82.92 ค่าประสิทธิภาพโดยรวม (f_measure) พบว่ามีค่าร้อยละ 83.21 และมีค่า AUC เท่ากับ 0.875 ดังแสดงในภาพประกอบ 122



ภาพประกอบ 122 ค่า Area under the curve (AUC) จากการตรวจสอบความแม่นยำของตัวแบบพยากรณ์โครงข่ายประสาทเทียมด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

การประเมินประสิทธิภาพตัวแบบพยากรณ์

1. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0)

ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) แสดงดังตาราง 65

ตาราง 65 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0)

ตัวแบบพยากรณ์	วิธีการคัดเลือกตัวแปร	Accuracy	F-measure
ต้นไม้ตัดสินใจ (Decision Tree)	วิธีเพิ่มตัวแปร (forward selection)	80.22	83.05
	วิธีลดตัวแปร (Backward Elimination)	80.65	82.79
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	80.22	83.05

ตัวแบบพยากรณ์	วิธีการ คัดเลือกตัวแปร	Accuracy	F-measure
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	80.43	83.47
การถดถอยโลจิสติกส์ (Logistic regression)	วิธีเพิ่มตัวแปร (forward selection)	81.52	84.10
	วิธีลดตัวแปร (Backward Elimination)	81.09	78.60
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	81.52	84.10
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	81.09	84.11
โครงข่ายประสาทเทียม (Neural Networks)	วิธีเพิ่มตัวแปร (forward selection)	80.65	83.46
	วิธีลดตัวแปร (Backward Elimination)	84.13	86.10
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	80.65	83.46
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	**85.22**	**87.47**

จากตาราง 65 พบว่าอัลกอริทึมที่ให้ประสิทธิภาพในการพยากรณ์สูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) โดยใช้วิธีการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ได้ค่าความถูกต้อง (Accuracy) เท่ากับ 85.22% และเมื่อพิจารณาประสิทธิภาพในการพยากรณ์โดยรวม (F-measure) อัลกอริทึมที่ให้ประสิทธิภาพสูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) เช่นกัน โดยได้ค่าประสิทธิภาพการพยากรณ์โดยรวมเท่ากับ 87.47%

จากผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ (cluster_0) ตัวแปรที่ผ่านการคัดเลือกด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ส่งผลต่อคะแนนรีวิว สามารถเปรียบเทียบค่าสหสัมพันธ์ (Weight by Correlation) เพื่อหาระดับของความสัมพันธ์ระหว่างตัวแปรดังกล่าวและตัวแปรตาม ผลการเปรียบเทียบค่าสหสัมพันธ์จากตัวแปรที่ผ่านการคัดเลือกด้วยวิธีนำตัวแปรเข้าทั้งหมดสำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ แสดงดังตาราง 66

พหุน ปณฺฑิโต ชีเว

ตาราง 66 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ

ตัวแปร	Weight
ระดับความคิดเห็น (Sentiment)	0.6561198611383731
จำนวนที่นั่ง (Seats)	0.21176887890680576
เหมาะสำหรับมาเป็นกลุ่ม (Group)	0.1640635516853657
มีบริการ Wi-Fi (Wi-Fi)	0.1405297505506546
ความหนาแน่นของประชากรในพื้นที่ (PopulationDensity)	0.1071870608202514
จำนวนผู้ถูกใจเพจ (Like)	0.09543089599293758
จำนวนบทวิจารณ์ (Reviews)	0.08904389188542278
บริการดิลิเวอรี (Delivery)	0.0679809058794426
มีโซเชียลมีเดียอื่นๆ (SocialMedia)	0.05737028885891772
มีบริการเครื่องดื่มแอลกอฮอล์ (Alcohol)	0.05616766294731035
ช่วงราคา (Price)	0.050764707070020634
รับบัตรเครดิต (CreditCard)	0.03587920088796787
มีที่จอดรถ (Parking)	0.03235879067643812
จำนวนเช็คอิน (Check-in)	0.020428369340400133
การมีเว็บไซต์ (Website)	0.002469203545608166

จากตาราง 66 พบว่า ปัจจัยที่มีผลต่อคะแนนรีวิวซึ่งสะท้อนถึงประสิทธิภาพของร้านอาหารบนโซเชียลคอมเมิร์ซสำหรับร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ 5 อันดับแรก ได้แก่ ระดับความคิดเห็น รองลงมาคือ จำนวนที่นั่ง เหมาะสำหรับมาเป็นกลุ่ม มีบริการ Wi-Fi และความหนาแน่นของประชากรในพื้นที่ตามลำดับ ปัจจัยที่ส่งผลน้อยที่สุดคือ การมีเว็บไซต์

2. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) แสดงดังตาราง 67

ตาราง 67 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแบบพยากรณ์	วิธีการคัดเลือกตัวแปร	Accuracy	F-measure
ต้นไม้ตัดสินใจ (Decision Tree)	วิธีเพิ่มตัวแปร (forward selection)	83.96	77.87
	วิธีลดตัวแปร (Backward Elimination)	84.32	79.61
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	83.96	77.87
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	83.78	78.85
การถดถอยโลจิสติกส์ (Logistic regression)	วิธีเพิ่มตัวแปร (forward selection)	84.31	77.36
	วิธีลดตัวแปร (Backward Elimination)	83.07	76.40
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	84.31	77.36
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	82.71	76.07
โครงข่ายประสาทเทียม (Neural Networks)	วิธีเพิ่มตัวแปร (forward selection)	85.92	81.76
	วิธีลดตัวแปร (Backward Elimination)	87.51	83.02
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	85.92	81.76
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	**91.08**	**88.06**

จากตาราง 67 พบว่าอัลกอริทึมที่ให้ประสิทธิภาพในการพยากรณ์สูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) โดยใช้วิธีการคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ได้ค่าความถูกต้อง (Accuracy) เท่ากับ 91.08% และเมื่อพิจารณาประสิทธิภาพในการพยากรณ์โดยรวม (F-measure) อัลกอริทึมที่ให้ประสิทธิภาพสูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) เช่นกัน โดยได้ค่าประสิทธิภาพการพยากรณ์โดยรวมเท่ากับ 88.06%

จากผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง (cluster_1) ตัวแปรที่ผ่านการคัดเลือกด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) ส่งผลต่อคะแนนรีวิว สามารถเปรียบเทียบค่าสหสัมพันธ์ (Weight by Correlation) เพื่อหาระดับของความสัมพันธ์ระหว่างตัวแปรดังกล่าวและตัวแปรตาม ผลการเปรียบเทียบค่าสหสัมพันธ์จากตัวแปรที่ผ่านการคัดเลือกด้วยวิธีนำตัวแปรเข้าทั้งหมดสำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง แสดงดังตาราง 68

ตาราง 68 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ตัวแปร	Weight
ระดับความคิดเห็น (Sentiment)	0.6954411847976828
ช่วงราคา (Price)	0.2364985717759491
มีบริการเครื่องดื่มแอลกอฮอล์ (Alcohol)	0.1880592239750462
รับบัตรเครดิต (CreditCard)	0.11298360143234373
จำนวนผู้ถูกใจเพจ (Like)	0.11091743814609674
บริการดิลิเวอรี (Delivery)	0.09697425012628968
มีบริการ Wi-Fi (Wi-Fi)	0.08948380533564136
เหมาะสำหรับมาเป็นกลุ่ม (Group)	0.06851063024118471
จำนวนที่นั่ง (Seats)	0.05917345723977308
มีที่จอดรถ (Parking)	0.0393758389621038
จำนวนเช็คอิน (Check-in)	0.03269077302717302
การมีเว็บไซต์ (Website)	0.013992155536301856
จำนวนบทวิจารณ์ (Reviews)	0.008546628202570954
มีโซเชียลมีเดียอื่นๆ (SocialMedia)	0.007088130139944237
ความหนาแน่นของประชากรในพื้นที่ (PopulationDensity)	0.0043240852106172095

จากตาราง 68 พบว่า ปัจจัยที่มีผลต่อคะแนนรีวิวซึ่งสะท้อนถึงประสิทธิภาพของร้านอาหารบนโซเชียลคอมเมิร์ซสำหรับร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง 5 อันดับแรก ได้แก่ ระดับความคิดเห็น รองลงมาคือ ช่วงราคา มีบริการเครื่องดื่มแอลกอฮอล์ รับบัตรเครดิต จำนวนผู้ถูกใจเพจ ตามลำดับ ปัจจัยที่ส่งผลน้อยที่สุดคือ ความหนาแน่นของประชากรในพื้นที่

3. ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) แสดงดังตาราง 69

ตาราง 69 เปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแบบพยากรณ์	วิธีการคัดเลือกตัวแปร	Accuracy	F-measure
ต้นไม้ตัดสินใจ (Decision Tree)	วิธีเพิ่มตัวแปร (forward selection)	80.91	79.89
	วิธีลดตัวแปร (Backward Elimination)	82.21	82.58
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	80.91	79.89
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	80.43	81.12
การถดถอยโลจิสติกส์ (Logistic regression)	วิธีเพิ่มตัวแปร (forward selection)	80.91	79.89
	วิธีลดตัวแปร (Backward Elimination)	82.22	82.82
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	80.91	79.89
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	81.57	82.12
โครงข่ายประสาทเทียม (Neural Networks)	วิธีเพิ่มตัวแปร (forward selection)	80.10	80.20
	วิธีลดตัวแปร (Backward Elimination)	**85.46**	**85.89**
	วิธีเพิ่มและลดตัวแปร (Optimize Selection)	80.10	80.20
	วิธีนำตัวแปรเข้าทั้งหมด (Enter Regression)	82.54	83.21

จากตาราง 69 พบว่าอัลกอริทึมที่ให้ประสิทธิภาพในการพยากรณ์สูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) โดยใช้วิธีการคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) ได้ค่าความถูกต้อง (Accuracy) เท่ากับ 85.46% และเมื่อพิจารณาประสิทธิภาพในการพยากรณ์โดยรวม (F-measure) อัลกอริทึมที่ให้ประสิทธิภาพสูงที่สุดคือโครงข่ายประสาทเทียม (Neural Networks) เช่นกัน โดยได้ค่าประสิทธิภาพการพยากรณ์โดยรวมเท่ากับ 85.89%

จากผลการเปรียบเทียบประสิทธิภาพตัวแบบพยากรณ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง (cluster_2) ตัวแปรที่ผ่านการคัดเลือกด้วยวิธีลดตัวแปร (Backward Elimination) ส่งผลต่อคะแนนรีวิว สามารถเปรียบเทียบค่าสหสัมพันธ์ (Weight by Correlation) เพื่อหาระดับของความสัมพันธ์ระหว่างตัวแปรดังกล่าวและตัวแปรตาม ผลการเปรียบเทียบค่าสหสัมพันธ์จากตัวแปรที่ผ่านการคัดเลือกด้วยวิธีลดตัวแปรสำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง แสดงดังตาราง 70

ตาราง 70 ผลการเปรียบเทียบค่าสหสัมพันธ์สำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง

ตัวแปร	Weight
ระดับความคิดเห็น (Sentiment)	0.6763930252704548
จำนวนที่นั่ง (Seats)	0.15286114309149804
ช่วงราคา (Price)	0.10852022780572947
ความหนาแน่นของประชากรในพื้นที่ (PopulationDensity)	0.07487075231854502
จำนวนบทวิจารณ์ (Reviews)	0.06710157830537254
มีบริการเครื่องดื่มแอลกอฮอล์ (Alcohol)	0.061437906780790745
การมีเว็บไซต์ (Website)	0.04716407055113735
มีโซเชียลมีเดียอื่นๆ (SocialMedia)	0.0401756730730093
มีที่จอดรถ (Parking)	0.029653113722536838
บริการดิลิเวอรี (Delivery)	0.028140845078864718
จำนวนเช็คอิน (Check-in)	0.01829268348013048
รับบัตรเครดิต (CreditCard)	0.001806236606482678

จากตาราง 70 พบว่า ปัจจัยที่มีผลต่อคะแนนรีวิวซึ่งสะท้อนถึงประสิทธิภาพของร้านอาหารบนโซเชียลคอมเมิร์ซสำหรับร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง 5 อันดับแรก ได้แก่ ระดับความคิดเห็น รองลงมาคือ จำนวนที่นั่ง ช่วงราคา ความหนาแน่นของประชากรในพื้นที่ จำนวนบทวิจารณ์ ตามลำดับ ปัจจัยที่ส่งผลน้อยที่สุดคือ รับบัตรเครดิต

พหุน ปณ ฑิโต ชีเว

บทที่ 5
สรุปผล อภิปรายผล และข้อเสนอแนะ

1. ความมุ่งหมายของการวิจัย
2. เครื่องมือที่ใช้ในการวิจัย
3. การเก็บรวบรวมข้อมูล
4. การวิเคราะห์ข้อมูล
5. อภิปรายผล
6. ข้อเสนอแนะ

ความมุ่งหมายของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

เพื่อเปรียบเทียบประสิทธิภาพอัลกอริธึมในการพยากรณ์ของการทำเหมืองข้อมูลสำหรับธุรกิจร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

เพื่อพัฒนาแบบจำลองสำหรับการวิเคราะห์ปัจจัยที่ส่งผลต่อประสิทธิภาพของร้านอาหารบนแพลตฟอร์มโซเชียลคอมเมิร์ซ

เครื่องมือที่ใช้ในการวิจัย

เครื่องมือที่ใช้ในงานวิจัยสำหรับการวิเคราะห์ข้อมูล ได้แก่ โปรแกรม Rapidminer เวอร์ชัน 10.0 ในการจัดกลุ่มด้วยเทคนิค K-means และการวิเคราะห์ข้อมูลโดยใช้อัลกอริธึมแมชชีนเลิร์นนิ่งเพื่อเปรียบเทียบประสิทธิภาพความแม่นยำของตัวแบบการพยากรณ์ (Prediction Model) ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) โครงข่ายประสาทเทียม (Neural Networks) และการวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis)

เครื่องมือที่ใช้ในการรวบรวมข้อมูลได้แก่ โปรแกรม Microsoft Visual Studio 2010 สำหรับพัฒนาโปรแกรมในการอ่านและแยกข้อมูลจากข้อมูลร้านอาหารที่รวบรวมจากเว็บไซต์ wongnai.com ที่อยู่ในเขตกรุงเทพมหานคร ข้อมูลจาก facebook fanpage ของร้านที่มีข้อมูลบน

เว็บไซต์ wongnai โปรแกรม Microsoft SQL Server 2008 R2 สำหรับจัดเก็บข้อมูล และโปรแกรม Microsoft Excel 2010 สำหรับแปลงข้อมูลและเป็นแหล่งข้อมูลนำเข้าให้กับโปรแกรม Rapidminer

การเก็บรวบรวมข้อมูล

ชุดข้อมูลที่ใช้ในงานวิจัยนี้ประกอบไปด้วยข้อมูลทุติยภูมิจากแหล่งต่างๆ นำมาเป็นปัจจัยนำเข้าเพื่อใช้ในการวิเคราะห์ข้อมูล ดังนี้

ข้อมูลร้านอาหารที่รวบรวมจากเว็บไซต์ wongnai.com ที่อยู่ในเขตกรุงเทพมหานคร ได้แก่ ประเภทร้านอาหาร ช่วงราคา ที่จอดรถ จำนวนบทวิจารณ์ การรับบัตรเครดิต บริการ Wi-Fi มีโซเชียลมีเดียอื่นๆ เหมาะสำหรับมาเป็นกลุ่ม การมีบริการเครื่องดื่มแอลกอฮอล์ มีเว็บไซต์(Website) บริการดิลิเวอรี และคะแนนรีวิว

ข้อมูลจาก facebook fanpage ของร้านที่มีข้อมูลบนเว็บไซต์ wongnai.com ที่อยู่ในเขตกรุงเทพมหานคร ได้แก่ จำนวนเช็คอิน จำนวนผู้ถูกใจเพจ จำนวนผู้ติดตามเพจ

ข้อมูลความหนาแน่นของประชากรในพื้นที่จากศูนย์เทคโนโลยีสารสนเทศภูมิศาสตร์กรุงเทพมหานคร ซึ่งเป็นหน่วยงานที่มีหน้าที่ดูแลรับผิดชอบ วางแผน จัดการกิจกรรมต่าง ๆ เกี่ยวข้องกับการพัฒนาและส่งเสริมการปฏิบัติงานด้านเทคโนโลยีสารสนเทศภูมิศาสตร์ของกรุงเทพมหานคร จากนั้นทำการบันทึกข้อมูลลงในฐานข้อมูลเพื่อนำไปใช้ในขั้นตอนการเตรียมข้อมูลต่อไป

ระดับความคิดเห็น โดยใช้ข้อมูลรีวิวจากลูกค้าผู้ใช้บริการซึ่งจะถูกนำไปหาค่าระดับความคิดเห็นโดยใช้ Natural Language AI ซึ่งให้บริการโดย Google สำหรับวิเคราะห์ความรู้สึกด้วยภาษาธรรมชาติ (Natural Language) โดยใช้ AI

การวิเคราะห์ข้อมูล (Data Analysis)

งานวิจัยนี้มุ่งเน้นการเปรียบเทียบประสิทธิภาพความแม่นยำของตัวแบบการพยากรณ์ (Prediction Model) ที่สร้างขึ้นโดยใช้อัลกอริธึมแมชชีนเลิร์นนิ่งด้วยโปรแกรม RapidMiner ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) โครงข่ายประสาทเทียม (Neural Networks) และการวิเคราะห์ถดถอยเชิงเส้น (Linear Regression Analysis) ซึ่งจากการศึกษางานวิจัยที่เกี่ยวข้องพบว่าเป็นอัลกอริธึมที่มีความแม่นยำสำหรับการสร้างตัวแบบพยากรณ์ที่มีประสิทธิภาพสูง โดยทำการแบ่งข้อมูลเพื่อใช้ในการทดสอบออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลเรียนรู้ (Training Data set) และทำการแบ่ง

ข้อมูลที่เหลือสำหรับใช้เป็นข้อมูลในการทดสอบ (Testing Data set) โดยใช้วิธี K-fold Cross validation

อภิปรายผล

จากผลการวิจัยพบว่า ในทุกกลุ่มประเภทร้านอาหารที่ถูกจัดกลุ่มโดยใช้ช่วงราคา เทคนิคโครงข่ายประสาทเทียม (Neural Networks) สามารถให้ประสิทธิภาพดีที่สุดในการพยากรณ์ ซึ่งสอดคล้องกับงานวิจัยของ Yi Luo (2021) จึงสามารถนำมาใช้ในการสร้างตัวแบบในการพยากรณ์ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซได้

ในกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำและสูง การคัดเลือกตัวแปรด้วยวิธีนำตัวแปรเข้าทั้งหมด (Enter Regression) สามารถให้ประสิทธิภาพดีที่สุดในการพยากรณ์ และสำหรับกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลาง การคัดเลือกตัวแปรด้วยวิธีลดตัวแปร (Backward Elimination) สามารถให้ประสิทธิภาพดีที่สุดในการพยากรณ์

ปัจจัยสำคัญที่ส่งผลต่อระดับคะแนนรีวิวของร้านอาหารซึ่งสะท้อนถึงระดับความพึงพอใจของลูกค้าในแต่ละกลุ่มประเภทร้านอาหารในแต่ละช่วงราคา เมื่อพิจารณาจากปัจจัยที่ผ่านการคัดเลือกและมีระดับของความสัมพันธ์กับระดับคะแนนรีวิวของร้านอาหาร 5 อันดับแรกของแต่ละกลุ่มสามารถสรุปได้ดังต่อไปนี้

ในด้านปัจจัยภายใน ระดับราคามีผลต่อระดับความพึงพอใจของลูกค้า ซึ่งสอดคล้องกับงานวิจัยของ Susskind (2019) โดยเฉพาะกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ปานกลางและสูง จำนวนที่นั่งส่งผลต่อกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำและปานกลาง เนื่องจากทั้งสองกลุ่มนี้ความหนาแน่นของประชากรในพื้นเป็นอีกหนึ่งปัจจัยที่ส่งผลต่อระดับความพึงพอใจของลูกค้าเช่นกัน จึงต้องสามารถรองรับจำนวนลูกค้าในปริมาณมากจึงควรมีจำนวนที่นั่งให้บริการอย่างเพียงพอ ซึ่งคล้ายคลึงกับงานวิจัยของ Wu (2021) และ Cantele (2020) ที่กล่าวถึงการพิจารณาทำแลที่ตั้งของร้านอาหารเป็นสิ่งสำคัญรวมถึงผู้มีส่วนได้ส่วนเสีย เช่น ผู้บริโภค และชุมชน

ในด้านการให้บริการอื่นๆ เพื่อเพิ่มความสะดวกสบายแก่ลูกค้า การมีบริการ Wi-Fi เป็นอีกหนึ่งปัจจัยที่สามารถดึงดูดลูกค้าให้เข้ามาใช้บริการในร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์ต่ำ ส่วนร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูงควรพิจารณาการรับบัตรเครดิต และการมีบริการเครื่องดื่มแอลกอฮอล์

ในด้านปัจจัยที่เกี่ยวข้องกับการตลาดผ่านโซเชียลมีเดีย ระดับความคิดเห็นของบทวิจารณ์ออนไลน์ส่งผลมากที่สุดต่อการตัดสินใจใช้บริการของลูกค้าในทุกกลุ่มร้านอาหาร ซึ่งสอดคล้องกับ

งานวิจัยของ Gao (2018) และ Hussain (2018) โดยเฉพาะอย่างยิ่ง จำนวนผู้ถูกใจเพจซึ่งเป็นข้อมูลจากแพลตฟอร์ม Facebook Fanpage ที่ร้านอาหารสามารถใช้เป็นช่องทางในการประชาสัมพันธ์ มีผลต่อระดับความพึงพอใจของลูกค้า ซึ่งสอดคล้องกับงานวิจัยของ P. Liu (2021) และ Sohn (2020) โดยเฉพาะกลุ่มประเภทร้านอาหารที่มีช่วงราคาอยู่ในเกณฑ์สูง

ข้อเสนอแนะ

ระดับความคิดเห็น (sentiment) เป็นปัจจัยที่ส่งผลต่อระดับคะแนนรีวิวของร้านอาหารซึ่งสะท้อนถึงระดับความพึงพอใจของลูกค้าซึ่งจะนำไปสู่ประสิทธิภาพของธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซในทุกระดับราคาแสดงให้เห็นว่าการตลาดแบบปากต่อปากมีผลอย่างมากสำหรับธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซ ดังนั้นในงานวิจัยจึงควรมุ่งเน้นที่การวิเคราะห์ความคิดเห็นในแง่มุมต่างโดยควรที่จะสามารถให้ระดับคะแนนตามประเภทเนื้อหาของบทวิจารณ์เพื่อให้สามารถนำข้อมูลดังกล่าวมาใช้ให้เกิดประโยชน์ในทางธุรกิจร้านอาหารบนโซเชียลคอมเมิร์ซได้อย่างมีประสิทธิภาพ นอกจากนี้ยังควรเพิ่มความหลากหลายของประเภทอาหารซึ่งอาจเป็นอีกปัจจัยสำคัญในการพิจารณาใช้บริการร้านอาหาร รวมถึงควรพิจารณากลยุทธ์การตลาด เช่นการสนับสนุนการขาย (Promotion Strategy) โดยการประชาสัมพันธ์ผ่านช่องทางโซเชียลมีเดียอื่นๆ เช่น Facebook Fanpage เนื่องจากลูกค้านิยมเลือกร้านอาหารผ่านแพลตฟอร์มดังกล่าวเช่นกัน รวมถึงกลยุทธ์ด้านสถานที่ (Place-Oriented Marketing) ซึ่งมีผลต่อความตั้งใจเข้าใช้บริการของลูกค้า

พหุน ปณ ฑิโต ชีเว

บรรณานุกรม

กม จารุเศรนี. (2022). ปัจจัยเชิงสาเหตุที่มีอิทธิพลต่อความสำเร็จของธุรกิจร้านอาหารขนาดกลางในเขตกรุงเทพมหานคร. *Journal of Management Science NakhonPathom Rajabhat University, 9*(2), 158-173.

กรุงศรีวิจัย. (2021). Social Commerce คลื่นลูกใหม่ของพาณิชย์อิเล็กทรอนิกส์. Retrieved from https://www.krungsri.com/th/research/research-intelligence/social-commerce-21

จักรรัตน์ กนะกาศัย, อุทัยรัตน์ เมืองแสน. (2022). การสร้างระบบการรักษาฐานของลูกค้าของธุรกิจร้านอาหารแบบดั้งเดิมเพื่อผลสำเร็จขององค์กร. *Rajapark Journal, 46*, 405-416.

จิราพัชร จุรุเทียบ. (2022). ช่องทางโซเชียลมีเดียและกลยุทธ์การตลาดที่มีผลต่อการตั้งใจ เข้าร้านอาหาร. *Journal of MCU Ubon Review, 7*(2), 673-682.

ชาลิสา จิตบุญญาพินิจ, ปราลี มณีรัตน์, นิเวศ จิระวิชิตชัย. (2022). การพัฒนาแบบจำลองการวิเคราะห์ความรู้สึกบนสื่อสังคมออนไลน์ไทยโดยใช้เทคนิคการเรียนรู้เชิงลึก. วารสารวิทยาศาสตร์และเทคโนโลยีหัวเฉียวเฉลิมพระเกียรติ, *8*(2), 8-18.

ณัฐวัฒน์ เอกธรรมนิตย์, วรพลพงษ์เพ็ชร. (2021). การจัดการวัตถุดิบคงคลังของร้านอาหาร โดยใช้ Proximal Policy Optimization. *Thai Journal of Operations Research, 9*, 45-54.

ปราโมชย์ นามวงศ์, ศุภเทพ สติมั่น,. (2022). การจำแนกความคิดเห็นการใช้บริการร้านอาหารบนเครือข่ายสังคมออนไลน์ ในจังหวัดอุบลราชธานีด้วยเทคนิคเหมืองข้อความคิดเห็น. วารสารการจัดการและการพัฒนา มหาวิทยาลัยราชภัฏอุบลราชธานี, *9*(2), 1-14.

ปริญญา สงวนสัตย์. (2019). *Artificial Intelligence with Machine Learning, AI* สร้างได้ด้วยแมชชีนเลิร์นนิ่ง. นนทบุรี: ไอดีซี พรีเมียร์, บจก.

พิศิษฐ์ บวรเลิศสุธี, วรภัทร ไพรีเกรง. (2022). ตัวแบบการวิเคราะห์ความรู้สึกทางอารมณ์สำหรับจำแนกประเภทบทความแนะนำสินค้าออนไลน์. *Journal of Engineering and Digital Technology (JEDT), 10*(1), 71-79.

Ba, J., Wang, P., Yang, X., Yu, H., & Yu, D. (2023). Glee: A granularity filter for feature selection. *Engineering Applications of Artificial Intelligence, 122*, 106080.

Bhadra, T., & Bandyopadhyay, S. (2021). Supervised feature selection using integration of densest subgraph finding with floating forward–backward search. *Information Sciences, 566*, 1-18.

Boardman, R. (2019). *Social Commerce: Consumer Behaviour in Online Environments*. Switzerland: Springer International Publishing.

Cantele, S. (2020). Sustainability implementation in restaurants: A comprehensive model ofdrivers, barriers, and competitiveness-mediated effects on firm performance. *International Journal of Hospitality Management, 87*, 1-10.

Chang, X. (2019). Business performance prediction in location-based social commerce. *Expert Systems With Applications, 126*, 112–123.

Chen, L.-F. (2016). Data mining framework based on rough set theory to improve location selection decisions: A case study of a restaurant chain. *Tourism Management, 53*, 197-206.

Chiu, S. (2008). *Data Mining and Market Intelligence for Optimal Marketing Returns*. London: Routledge.

Chu, S.-C. (2019). *Electronic Word of Mouth as a Promotional Technique: New Insights from social media*. New York: Routledge.

David L. Olson. (2008). *Advanced Data Mining Techniques*. Berlin: Springer.

Desai, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). *Clinical eHealth, 4*, 1-11.

Du, K.-L. (2019). *Neural Networks and Statistical Learning*. London: Springer.

Fernandes, E. (2021). A data-driven approach to measure restaurant performance by combining online reviews with historical sales data. *International Journal of Hospitality Management, 94*, 1-10.

Gao, S. (2018). Identifying competitors through comparative relation mining of online reviews in the restaurant industry. *International Journal of Hospitality Management, 71*, 19-32.

Golmohammadi, A. (2020). Negative online reviews and consumers' service consumption. *Journal of Business Research, 116*, 27-36.

Hajli, N. (2017). A social commerce investigation of the role of trust in a social networking site on purchase intentions. *Journal of Business Research, 71*, 133-141.

Hussain, S. (2018). Consumers' online information adoption behavior: Motives and antecedents of electronic word of mouth communications. *Computers in Human Behavior, 80*.

Jia, S. S. (2020). Motivation and satisfaction of Chinese and U.S. tourists in restaurants: A cross-cultural text mining of online reviews. *Tourism Management, 78*, 1-12.

Lee, C. (2018). Investigating the moderating role of education on a structural model of restaurant performance using multi-group PLS-SEM analysis. *Journal of Business Research, 88*, 298-305.

Leonard Gunawan, M. S. A., Louis Wihan, Santy, Gihon Yonathan Lesmana, Sabin Yusuf. (2023). Support vector machine based emotional analysis of restaurant reviews. *Procedia Computer Science, 216*.

Li, H. (2020). Online persuasion of review emotional intensity: A text mining analysis of restaurant reviews. *International Journal of Hospitality Management, 89*, 1-13.

Liao, S.-H. (2021). Investigating online social media users' behaviors for social commerce recommendations. *Technology in Society, 66*, 1-14.

Liu, H. (2021). Social sharing of consumption emotion in electronic word of mouth (eWOM): A cross-media perspective. *Journal of Business Research, 132*, 208-220.

Liu, P. (2021). The effects of social commerce environmental characteristics on customers' purchase intentions: The chain mediating effect of customer-to-customer interaction and customer-perceived value. *Electronic Commerce Research and Applications, 48*, 1-14.

Mehrbakhsh Nilashi, H. A., Goli Arji, Khalaf Okab Alsalem, Sarminah Samad, Fahad Ghabban, Ahmed Omar Alzahrani, Ali Ahani, Ala Abdulsalam Alarood. (2021). Big social data and customer decision making in vegetarian restaurants: A combined machine learning method. *Journal of Retailing and Consumer Services, 62*, 1-18.

Meilatinova, N. (2021). Social commerce: Factors affecting customer repurchase and word-of-mouth intentions. *International Journal of Information Management, 57*, 1-12.

Najwa AlGhamdi, S. K. (2021). Improving Sentiment Prediction using Heterogeneous and Homogeneous Ensemble Methods: A Comparative Study. *Procedia Computer Science, 194*, 60-68.

Namin, A. (2017). Revisiting customers' perception of service quality in fast food restaurants. *Journal of Retailing and Consumer Services, 34*, 70-81.

Nisbet, R. (2009). *Handbook of Statistical Analysis and Data Mining Applications*.

Canada: Elsevier Inc.

oberlo.com. (2023). SOCIAL COMMERCE MARKET SIZE (2020–2026). Retrieved from https://www.oberlo.com/statistics/social-commerce-market-size

Parsa, H. G. (2010). Why do restaurants fail? Part III: An analysis of macro and micro factors. *Emerging Aspects Redefining Tourism and Hospitality*, 16-25.

Pyle, M. A. (2021). In eWOM we trust: Using naïve theories to understand consumer trust in a complex eWOM marketspace. *Journal of Business Research, 122*, 145-158.

Ruiz-Mafe, C. (2018). The role of emotions and conflicting online reviews on consumers' purchase intentions. *Journal of Business Research, 89*, 336-344.

Sebastian Raschka, V. M. (2019). *Python Machine Learning : Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2*: Packt Publishing Ltd.

Sohn, J. W. (2020). Factors that influence purchase intentions in social commerce. *Technology in Society, 63*, 1-11.

Song, H. J. (2021). The influence of board interlocks on firm performance: In the context of geographic diversification in the restaurant industry. *Tourism Management, 83*, 1-9.

Susskind, A. M. (2019). *The Next Frontier of Restaurant Management: Harnessing Data to Improve Guest Service and Enhance the Employee Experience*. New York: Cornell University Press.

Takashi Tanizaki, T. H., Takeshi Shimmura, Takeshi Takenaka. (2019). Demand forecasting in restaurants using machine learning and statistical analysis. *Procedia CIRP, 79*, 679–683.

Turban, E. (2016). *Social Commerce: Marketing, Technology and Management*. Switzerland: Springer.

Wang, L. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Systems With Applications, 42*(2), 855-863.

Wang, X.-Y. (2016). A new SVM-based relevance feedback image retrieval using probabilistic feature and weighted kernel function. *Journal of Visual Communication and Image Representation, 38*, 256-275.

Wang, Y. (2021). Interconnectedness between online review valence, brand, and restaurant performance. *Journal of Hospitality and Tourism Management, 48*, 138-145.

Wu, M. (2021). Roles of locational factors in the rise and fall of restaurants: A case study of Beijing with POI data. *Cities, 113*, 1-14.

Xu, J., Lu, W., Li, J., & Yuan, H. (2022). Dependency maximization forward feature selection algorithms based on normalized cross-covariance operator and its approximated form for high-dimensional data. *Information Sciences, 617*, 416-434.

Yan, Q. (2018). The influences of tourists' emotions on the selection of electronic word of mouth platforms. *Tourism Management, 66*, 348-363.

Yi Luo, X. X. (2021). Comparative study of deep learning models for analyzing online restaurant reviews in the era of the COVID-19 pandemic. *International Journal of Hospitality Management, 91*, 1-8.

Yoni Nazarathy, H. K. (2021). *Statistics with Julia: Fundamentals for Data Science, Machine Learning and Artificial Intelligence*. Gewerbestrasse 11 6330 Cham Switzerland: Springer.

Yuan, H. (2017). Topic sentiment mining for sales performance prediction in e-commerce. *Big Data Analytics in Operations & Supply Chain Management*, 553-576.

Zhang, X. (2020). Hybridization of cognitive computing for food services. *Applied Soft Computing, 89*, 1-8.

Zhao, Y. (2020). Electronic word-of-mouth and consumer purchase intentions in social e-commerce. *Electronic Commerce Research and Applications, 41*, 1-9.

ประวัติผู้เขียน

ชื่อ	นายศุภมิตร บุญทา
วันเกิด	วันที่ 12 เมษายน 2520
สถานที่เกิด	จังหวัดสกลนคร
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 451 ซอยเศรษฐี ต.ธาตุเชิงชุม อ.เมืองสกลนคร จ.สกลนคร
ตำแหน่งหน้าที่การงาน	อาจารย์
สถานที่ทำงานปัจจุบัน	มหาวิทยาลัยราชภัฏสกลนคร
ประวัติการศึกษา	พ.ศ. 2548 วิศวกรรมศาสตรบัณฑิต (วศ.บ.) สาขาวิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีมหานคร พ.ศ. 2553 วิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยเกษตรศาสตร์ พ.ศ. 2566 ปรัชญาดุษฎีบัณฑิต (ปร.ด.) สาขาวิชาบริหารธุรกิจและนวัตกรรมดิจิทัล มหาวิทยาลัยมหาสารคาม
ผลงานวิจัย	Supamit Boonta, & Rujira Chongklaiklang. (2019). Forecasting water levels using data mining techniques: a case study at Nong Han Lake, Sakon Nakhon, Thailand. SNRU Journal of Science and Technology 11(2), 64-70. Supamit Boonta. (2017). Analytics of Students Behavior of a Blended Learning Applying Course Management System to Application Development for Mobile Devices Course Using Data Mining Techniques. The 3rd AsTEN Conference Teaching Competency Development: Issues, Innovation, & Initiatives (p. 60). Bangkok,: Kasetsart University. ศุภมิตร บุญทา, & จีระนันต์ เจริญรัตน์. (2559). การพยากรณ์การเปลี่ยนแปลงระดับน้ำด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กรณีศึกษา ทะเลสาบหนองหาร จังหวัดสกลนคร. การประชุมวิชาการ

ระดับชาติ เครือข่ายวิจัยสถาบันอุดมศึกษาทั่วประเทศ ครั้งที่ 11 (pp. 304-310). นครราชสีมา: มหาวิทยาลัยเทคโนโลยีสุรนารี.