

การพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็น  
บนเครือข่ายสังคมออนไลน์



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต  
สาขาวิชาเทคโนโลยีสารสนเทศ  
มหาวิทยาลัยเทคโนโลยีสุรนารี  
ปีการศึกษา 2566

**THE DEVELOPMENT OF RESTAURANT RECOMMENDATION  
MODEL BASED ON COMMENTS ON ONLINE  
SOCIAL NETWORK**



**A Thesis Submitted in Partial Fulfillment of the Requirements for the  
Degree of Doctor of Philosophy in Information Technology  
Suranaree University of Technology  
Academic Year 2023**

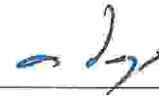
## การพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

มหาวิทยาลัยเทคโนโลยีสุรนารี อนุมัติให้นำวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษา  
ตามหลักสูตรปริญญาปรัชญาดุษฎีบัณฑิต

คณะกรรมการสอบวิทยานิพนธ์



(ผู้ช่วยศาสตราจารย์ ดร.ศุภกฤษฎี นิวัฒนากุล)  
ประธานกรรมการ



(รองศาสตราจารย์ ดร.ชรา อังสกุล)  
กรรมการ (อาจารย์ที่ปรึกษาวิทยานิพนธ์)



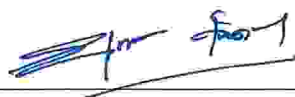
(รองศาสตราจารย์ ดร.ศิริบุษ บุนยกรอง)  
กรรมการ



(ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ จันทวีวัฒน์)  
กรรมการ



(ผู้ช่วยศาสตราจารย์ ดร.หนึ่งหทัย ขอผลกลาง)  
กรรมการ



(รองศาสตราจารย์ ดร.ยุพาพร รักสกุลพิวัฒน์)  
รองอธิการบดีฝ่ายวิชาการและประกันคุณภาพ



(รองศาสตราจารย์ ดร.ชรา อังสกุล)  
คณบดีสำนักวิชาศาสตร์และศิลป์ดิจิทัล

ลดาวัลย์ แยมครวญ : การพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่าย  
สังคมออนไลน์ (THE DEVELOPMENT OF RESTAURANT RECOMMENDATION  
MODEL BASED ON COMMENTS ON ONLINE SOCIAL NETWORK)

อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.ธรา อังสกุล, 86 หน้า

อาหารเป็นหนึ่งในปัจจัยที่มีความสำคัญต่อการดำรงชีวิต คนไทยให้ความสนใจเรื่องอาหาร  
รวมถึงมีร้านอาหารในประเทศไทยเป็นจำนวนมาก โดยในปัจจุบันผู้ใช้มีพฤติกรรมการใช้เวลาบน  
เครือข่ายสังคมออนไลน์มากขึ้น จึงทำให้มีการแบ่งปันข้อมูลเกี่ยวกับอาหารและร้านอาหารผ่าน  
เครือข่ายสังคมออนไลน์ ทั้งในรูปแบบบทวิจารณ์ และภาพเกี่ยวกับอาหารและร้านอาหาร งานวิจัยนี้  
จึงมีวัตถุประสงค์เพื่อพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคม  
ออนไลน์ โดยเก็บรวบรวมข้อมูลจากเครือข่ายสังคมออนไลน์ทวิตเตอร์ (เอ็กซ์) แล้วนำมาวิเคราะห์  
เพื่อจำแนกประเภทบทวิจารณ์และประเภทอาหาร

ในการจำแนกประเภทบทวิจารณ์ ได้จำแนกประเภทบทวิจารณ์ออกเป็น 3 ประเภท ได้แก่  
บทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ ด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายเปรียบเทียบกับ  
ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน ผลการทดลองพบว่า ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีนมี  
ค่าความถูกต้องในการทำนายร้อยละ 82.89 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 82.00 ค่าความ  
ระลึกเฉลี่ยถ่วงน้ำหนัก ร้อยละ 83.00 และค่าวัดประสิทธิภาพ ร้อยละ 82.00 และเมื่อทดสอบกับ  
ข้อมูลใหม่ พบว่า ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้อง ร้อยละ 80.87 ค่าความ  
แม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 80.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก ร้อยละ 81.00 และค่าวัด  
ประสิทธิภาพ ร้อยละ 80.00 ซึ่งให้ผลลัพธ์ที่ดีกว่าขั้นตอนวิธีแบบเบย์อย่างง่าย

ส่วนการจำแนกประเภทอาหารบนภาพใช้วิธีการประมวลผลภาพด้วยโวล เวอร์ชัน 5  
โดยแบ่งประเภทอาหารเป็น 9 ประเภท ได้แก่ อเมริกาน้อยชู ขานมไข่มุก ข้าวปลาแกง ยำไข่แดงต้ม  
หมูยอ แซลมอนเสียบไม้ย่าง สปรักกิ่งโซดา ไก่เกาหลี ผลไม้ดอง และราเมนไอศกรีม ผลการ  
วิเคราะห์พบว่า ให้ค่าความถูกต้องในการทำนายร้อยละ 98.00 ค่าความแม่นยำ ร้อยละ 94.00 ค่า  
ความระลึกเฉลี่ยถ่วงน้ำหนัก ร้อยละ 96.00 และค่าวัดประสิทธิภาพ ร้อยละ 95.00

LADAWAN YAMKUAN : THE DEVELOPMENT OF RESTAURANT  
RECOMMENDATION MODEL BASED ON COMMENTS ON ONLINE  
SOCIAL NETWORK. THESIS ADVISOR : ASSOC. PROF. THARA  
ANGSKUN, Ph. D., 108 PP.

## RESTAURANT RECOMMENDATION/SOCIAL NETWORK

Food is one of the factors that are important for survival. Thai people are interested in food, including many restaurants in Thailand. Currently, users have the habit of spending more time on online social networks. As a result, information about food and restaurants is shared through social networks. Both in the form of reviews and images. This research aims to develop a restaurant recommendation model based on online social networks. This research collects data from the online social network Twitter (X) and then analyzes data to classify reviews and food types.

Reviews are classified into three categories: positive, neutral, and negative reviews using a Naive Bayes compared with a Support Vector Machine (SVM) algorithm. The results indicated that SVM has an accuracy of 82.89 percent, weighted average precision of 82.00 percent, weighted average recall of 83.00 percent, and f-measure of 82.00 percent. And then test with unseen data, The results obtained from SVM with an accuracy of 80.87 percent, weighted average precision of 80.00 percent, weighted average recall of 81.00 percent, and f-measure of 80.00 percent give better results than Naive Bayes.



As for the classification of food on the images, image processing using YOLOv5, with food classified into nine categories, including americano yuzu, bubble tea, crispy fish, egg yolk salad with pork sausage, grilled salmon skewers, sparkling soda, Korean chicken, pickled fruit, and ramen ice cream. The analysis found that the model gives an accuracy of 98.00 percent, a precision of 94.00 percent, a recall of 96.00 percent, and an f-measure of 95.00 percent.



School of Information Technology

Academic Year 2023

Student's Signature Lodawan

Advisor's Signature [Signature]

## กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จลุล่วงได้ด้วยดี ผู้วิจัยขอขอบพระคุณ รองศาสตราจารย์ ดร.ธรา อังสกุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้กรุณาให้คำปรึกษา แนะนำ และให้ความช่วยเหลืออย่างดียิ่ง ทั้งในด้านวิชาการ และด้านการดำเนินการวิจัย

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.สุภกฤษฎ์ นิวัฒนากุล ประธานกรรมการสอบ วิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.สุพจน์ จันทรวิวัฒน์ รองศาสตราจารย์ ดร.ศิริปฐุช บุญครอง และผู้ช่วยศาสตราจารย์ ดร.หนึ่งหทัย ขอผลกลาง กรรมการสอบวิทยานิพนธ์ ที่ได้กรุณาเสียสละ เวลาอันมีค่าในการพิจารณา และให้คำแนะนำในการแก้ไขปรับปรุงวิทยานิพนธ์ ที่เป็นประโยชน์ต่อ วิทยานิพนธ์ฉบับนี้อย่างยิ่ง

ขอขอบพระคุณ บุคลากรประจำสำนักวิชาศาสตร์และศิลป์ดิจิทัล มหาวิทยาลัยเทคโนโลยี สุรนารี ที่ให้ความช่วยเหลือ สนับสนุนและเอื้ออำนวยความสะดวกต่าง ๆ ตลอดระยะเวลาการ ศึกษาวิจัยนี้

ขอขอบคุณเพื่อนนักศึกษามัคคุเทศิกษาทุกท่านที่ให้อำนาจใจและให้การสนับสนุนการวิจัย อันมีส่วนทำให้วิทยานิพนธ์นี้เสร็จสมบูรณ์

สุดท้ายนี้ ขอกราบขอบพระคุณบิดา มารดา ที่ให้การเลี้ยงดูอบรม และส่งเสริมการศึกษา เป็นอย่างดี จนทำให้ผู้วิจัยประสบความสำเร็จตลอดมา

ลดาวัลย์ แยมครวญ

# สารบัญ

หน้า

|                                                  |    |
|--------------------------------------------------|----|
| บทคัดย่อ (ภาษาไทย)                               | ก  |
| บทคัดย่อ (ภาษาอังกฤษ)                            | ข  |
| กิตติกรรมประกาศ                                  | ค  |
| สารบัญ                                           | ง  |
| สารบัญตาราง                                      | ช  |
| สารบัญรูป                                        | ณ  |
| บทที่                                            |    |
| 1 บทนำ                                           | 1  |
| 1.1 ความสำคัญและที่มาของปัญหาการวิจัย            | 1  |
| 1.2 วัตถุประสงค์ของการวิจัย                      | 8  |
| 1.3 สมมุติฐานการวิจัย                            | 8  |
| 1.4 ข้อตกลงเบื้องต้น                             | 9  |
| 1.5 ขอบเขตของการวิจัย                            | 9  |
| 1.6 ประโยชน์ที่คาดว่าจะได้รับ                    | 9  |
| 1.7 คำอธิบายศัพท์                                | 9  |
| 2 ปรัชญารวบรวมและงานวิจัยที่เกี่ยวข้อง           | 10 |
| 2.1 ธุรกิจอาหารและเครื่องดื่ม                    | 10 |
| 2.1.1 ความหมายของธุรกิจอาหารและเครื่องดื่ม       | 10 |
| 2.1.2 ความเป็นมาของธุรกิจร้านอาหารและเครื่องดื่ม | 11 |
| 2.1.3 ประเภทของธุรกิจอาหารและเครื่องดื่ม         | 11 |
| 2.2 สื่อสังคมออนไลน์                             | 12 |
| 2.2.1 ความหมายของสื่อสังคมออนไลน์ (Social Media) | 12 |
| 2.2.2 ประเภทของเครือข่ายสังคมออนไลน์             | 13 |
| 2.3 บทวิจารณ์ออนไลน์                             | 15 |
| 2.3.1 ความหมายของบทวิจารณ์ออนไลน์                | 15 |



## สารบัญ (ต่อ)

### หน้า

|                                                                                                 |    |
|-------------------------------------------------------------------------------------------------|----|
| 2.3.2 บทวิจารณ์ออนไลน์บนเครือข่ายสังคมออนไลน์ .....                                             | 15 |
| 2.4 การประมวลผลทางภาษา (Word Segmentation) .....                                                | 16 |
| 2.4.1 อักษรวิธีของภาษาไทย .....                                                                 | 16 |
| 2.4.2 วิธีการตัดคำ .....                                                                        | 17 |
| 2.4.3 เทคนิคการตัดคำ .....                                                                      | 17 |
| 2.4.4 ขั้นตอนวิธีในการตัดคำ .....                                                               | 18 |
| 2.4.5 การคัดแยกคำตามความสำคัญด้วย (Term Frequency-Inverse Document<br>Frequency: TF-IDF) .....  | 19 |
| 2.5 การประมวลผลภาพดิจิทัล (Digital Image Processing) .....                                      | 20 |
| 2.5.1 ความหมายของการประมวลผลภาพ .....                                                           | 20 |
| 2.5.2 ขั้นตอนการประมวลผลภาพ (Step of Image Processing) .....                                    | 21 |
| 2.5.3 องค์ประกอบในการประมวลผลภาพ (Components of Image Processing) .....                         | 23 |
| 2.6 โยโล (You Only Look Once: YOLO) .....                                                       | 25 |
| 2.6.1 คุณลักษณะของ YOLO .....                                                                   | 25 |
| 2.6.2 โครงสร้างการทำงานของ YOLOv5 .....                                                         | 27 |
| 2.7 การเรียนรู้ของเครื่อง (Machine Learning) .....                                              | 29 |
| 2.7.1 ความหมายของการเรียนรู้ของเครื่อง .....                                                    | 29 |
| 2.7.2 เทคนิคการเรียนรู้ของเครื่อง .....                                                         | 29 |
| 2.8 การเรียนรู้เชิงลึก (Deep Learning) .....                                                    | 34 |
| 2.8.1 ความหมายของการเรียนรู้เชิงลึก (Deep Learning) .....                                       | 34 |
| 2.8.2 ประเภทของการเรียนรู้เชิงลึก (Way to Deep Learning) .....                                  | 34 |
| 2.8.2.1 การเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน (Unsupervised Deep<br>Networks Learning) ..... | 34 |
| 2.8.2.2 การเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน (Supervised Deep<br>Networks Learning) .....      | 35 |

## สารบัญ (ต่อ)

หน้า

|                                                                                                  |           |
|--------------------------------------------------------------------------------------------------|-----------|
| 2.8.2.3 การเรียนรู้โครงข่ายเชิงลึกแบบผสม (Hybrid Deep Networks Learning)                         | 36        |
| 2.9 งานวิจัยที่เกี่ยวข้อง                                                                        | 36        |
| 2.10 กรอบแนวคิดการวิจัย                                                                          | 41        |
| <b>3 วิธีดำเนินการวิจัย</b>                                                                      | <b>42</b> |
| 3.1 วิธีวิจัย                                                                                    | 42        |
| 3.1.1 การศึกษาและวิเคราะห์ปัญหา                                                                  | 42        |
| 3.1.2 การออกแบบแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์                    | 44        |
| 3.1.3 การพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์                        | 45        |
| 3.1.4 การประเมินผลแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์                 | 46        |
| 3.2 เครื่องมือที่ใช้ในการวิจัย                                                                   | 47        |
| 3.2.1 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล                                                      | 47        |
| 3.2.2 เครื่องมือที่ใช้ในการพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ | 47        |
| 3.2.3 เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล                                                       | 48        |
| 3.3 การเก็บรวบรวมข้อมูล                                                                          | 48        |
| 3.4 การวิเคราะห์ข้อมูล                                                                           | 51        |
| 3.4.1 การวิเคราะห์ข้อมูลจากบทวิจารณ์เพื่อจำแนกประเภทบทวิจารณ์                                    | 51        |
| 3.4.2 การวิเคราะห์ข้อมูลจากภาพเพื่อจำแนกประเภทอาหาร                                              | 56        |
| <b>4 ผลการวิจัยและอภิปรายผล</b>                                                                  | <b>58</b> |
| 4.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes                                              | 59        |
| 4.1.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bay เพื่อสร้างแบบจำลอง                           | 59        |
| 4.1.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bay กับข้อมูลใหม่                                | 60        |

## สารบัญ (ต่อ)

หน้า

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| 4.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM .....                        | 61 |
| 4.2.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM เพื่อสร้างแบบจำลอง .....   | 61 |
| 4.2.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่ .....        | 62 |
| 4.3 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) .....                  | 63 |
| 4.3.1 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ .....          | 63 |
| 4.3.2 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของแฮชแท็ก .....     | 64 |
| 4.3.3 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของสัญลักษณ์ .....   | 64 |
| 4.4 ผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพ .....                         | 65 |
| 4.5 อภิปรายผลการวิจัย .....                                              | 71 |
| 4.5.1 อภิปรายผลการวิจัยการสร้างแบบจำลองการจำแนกประเภทบทวิจารณ์ .....     | 71 |
| 4.5.2 อภิปรายผลการวิจัยการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ..... | 72 |
| 4.5.3 อภิปรายผลการวิจัยการสร้างแบบจำลองการจำแนกประเภทอาหารจากภาพ .....   | 72 |
| 5 สรุปผลการวิจัย .....                                                   | 75 |
| 5.1 สรุปผลการวิจัย .....                                                 | 75 |
| 5.2 ข้อจำกัดของงานวิจัย .....                                            | 79 |
| 5.3 การประยุกต์ผลการวิจัย .....                                          | 80 |
| 5.4 ข้อเสนอแนะงานวิจัยครั้งต่อไป .....                                   | 80 |
| รายการอ้างอิง .....                                                      | 81 |
| ประวัติผู้วิจัย .....                                                    | 86 |

## สารบัญตาราง

| ตารางที่                                                                                    | หน้า |
|---------------------------------------------------------------------------------------------|------|
| 2.1 ตัวอย่างการสะกดคำตามอักขรวิธี                                                           | 16   |
| 2.2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง                                                     | 39   |
| 3.1 องค์ประกอบที่มีความเกี่ยวข้องกับการประเมินแบบจำลองการแนะนำร้านอาหาร                     | 47   |
| 3.2 ตัวอย่างคำที่เกี่ยวข้องกับการวิจารณ์อาหาร                                               | 53   |
| 3.3 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40                   | 54   |
| 3.4 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 70:30                   | 55   |
| 3.5 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 80:20                   | 55   |
| 4.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes                                         | 59   |
| 4.2 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี Naive Bayes                | 59   |
| 4.3 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่                           | 60   |
| 4.4 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี Naive Bayesd กับข้อมูลใหม่ | 60   |
| 4.5 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM                                                 | 61   |
| 4.6 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM                        | 61   |
| 4.7 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่                                   | 62   |
| 4.8 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่          | 62   |
| 4.9 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ                                     | 63   |
| 4.10 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของแฮชแท็ก                               | 64   |
| 4.11 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของสัญลักษณ์                             | 65   |
| 4.12 ผลการสร้างแบบจำลองจำแนกประเภทอาหารบนภาพ                                                | 66   |
| 4.13 ตัวอย่างผลการจำแนกประเภทอาหารจากภาพที่นำไปทดสอบ                                        | 69   |
| 5.1 สรุปผลการวิจัยขั้นตอนวิธี SVM และ Naive Bayes                                           | 76   |
| 5.2 สรุปผลการวิจัยขั้นตอนวิธี SVM และ Naive Bayes กับข้อมูลใหม่                             | 77   |

## สารบัญรูป

### รูปที่

### หน้า

|      |                                                                                           |    |
|------|-------------------------------------------------------------------------------------------|----|
| 1.1  | แผนภูมิแสดงจำนวนบัญชีผู้ใช้งานเครือข่ายสังคมออนไลน์ตั้งแต่ปี ค.ศ. 2017 ถึงปี ค.ศ. 2023    | 4  |
| 1.2  | แผนภูมิแสดงการค้นหาข้อมูลข่าวสารออนไลน์จากแหล่งข้อมูลที่แตกต่างกัน                        | 4  |
| 1.3  | วัตถุประสงค์การใช้สื่อสังคมออนไลน์ในการค้นหาทวิจรร้านอาหาร                                | 6  |
| 2.1  | จุดพิกเซลบนภาพดิจิทัล                                                                     | 21 |
| 2.2  | ภาพดิจิทัลขนาด $M \times N$ ในรูปแบบเมทริกซ์                                              | 21 |
| 2.3  | องค์ประกอบในการประมวลผลภาพดิจิทัล                                                         | 24 |
| 2.4  | สถาปัตยกรรมของ YOLO                                                                       | 25 |
| 2.5  | องค์ประกอบในการทำงานของ YOLO                                                              | 26 |
| 2.6  | ขั้นตอนวิธีของ YOLOv5                                                                     | 26 |
| 2.7  | ประสิทธิภาพการทำงานของ YOLOv5                                                             | 27 |
| 2.8  | โครงสร้างการทำงานของ YOLOv5                                                               | 27 |
| 2.9  | ตัวอย่างการตรวจจับวัตถุบนภาพด้วย YOLOv5                                                   | 28 |
| 2.10 | ความกว้างระหว่างเส้นตรงกับซัพพอร์ตเวกเตอร์                                                | 30 |
| 2.11 | ขั้นตอนวิธีค่าเฉลี่ย K                                                                    | 33 |
| 2.12 | ลักษณะการเลื่อนของตัวกรอง                                                                 | 36 |
| 2.13 | กรอบแนวคิดการวิจัย                                                                        | 41 |
| 3.1  | องค์ประกอบและกระบวนการของแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์   | 44 |
| 3.2  | ตัวอย่างการดึงข้อมูลส่วนตัว (Profile) บน Twitter ของผู้ใช้                                | 49 |
| 3.3  | ตัวอย่างการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้แบบไม่มีข้อมูลรูปภาพ                    | 49 |
| 3.4  | ตัวอย่างการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้แบบมีข้อมูลรูปภาพ                       | 50 |
| 3.5  | ตัวอย่างการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้ตามวัตถุ (Object) ที่ Twitter API กำหนด | 50 |
| 3.6  | ตัวอย่างการเรียกข้อมูลด้วยคำสำคัญบน Twitter                                               | 51 |
| 3.7  | สถาปัตยกรรมของแบบจำลองการจำแนกประเภททวิจรร้านอาหาร                                        | 52 |
| 3.8  | ตัวอย่างการจำแนกประเภททวิจรร้านอาหารจากคำที่เกี่ยวข้อง                                    | 53 |

## สารบัญรูป (ต่อ)

| รูปที่                                                                         | หน้า |
|--------------------------------------------------------------------------------|------|
| 3.9 ตัวอย่างการกำหนดคลาสให้กับภาพ .....                                        | 57   |
| 3.10 การดึงข้อมูลจากภาพเพื่อระบุตำแหน่งของร้านอาหาร .....                      | 57   |
| 4.1 จำนวนของอินสแตนซ์ที่ปรากฏบนภาพที่นำไปสร้างแบบจำลอง .....                   | 67   |
| 4.2 Confusion matrix แสดงผลการสร้างแบบจำลองจำแนกประเภทอาหาร .....              | 67   |
| 4.3 ตัวอย่างภาพจากการสร้างแบบจำลองการจำแนกประเภทอาหาร .....                    | 68   |
| 4.4 แบบจำลองสำหรับการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ..... | 70   |
| 4.5 ตัวอย่างการนำเสนอข้อมูลร้านอาหารเพื่อแนะนำร้านอาหารแก่ผู้ใช้ .....         | 74   |





## บทที่ 1

### บทนำ

#### 1.1 ความสำคัญและที่มาของปัญหาการวิจัย

ธุรกิจอาหารและเครื่องดื่มในประเทศไทยมีมูลค่าทางการตลาดสูงถึง 2.3 ล้านล้านบาทต่อปี ซึ่งคาดว่าจะยังคงเติบโตมากขึ้น โดยเฉพาะแนวคิดในการตอบสนองวิถีคนเมืองที่นิยมความรวดเร็ว สะดวกสบาย มีเวลาจำกัด แต่ยังคงต้องการสินค้าและบริการที่ดีมีคุณภาพสำหรับการบริโภค รวมถึงการให้บริการที่สอดคล้องกับพฤติกรรมการใช้เวลาไปกับเครือข่ายสังคมออนไลน์มากขึ้น อาทิ การแจ้งข้อมูลไปยังผู้บริโภคอย่างสม่ำเสมอ เช่น ข้อมูลการปรับปรุงสินค้าและบริการ ข้อมูลโปรโมชั่นผ่านเครือข่ายสังคมออนไลน์ เนื่องจากในปัจจุบันมีจำนวนผู้ใช้เครือข่ายสังคมออนไลน์เพิ่มมากขึ้น ทำให้เกิดกลุ่มผู้บริโภคที่รู้และเข้าใจเทคโนโลยี และมีพฤติกรรมการใช้สื่อสังคมออนไลน์แทบทุกประเภท (ศูนย์วิจัยกสิกรไทย, 2560)

ในประเทศไทยนั้น คนไทยให้ความสนใจเรื่องอาหาร ไม่ว่าจะเป็นในชีวิตประจำวันหรือในขณะท่องเที่ยว รวมถึงมีร้านอาหารในประเทศไทยเป็นจำนวนมาก ผลการวิเคราะห์พฤติกรรมผู้บริโภคออนไลน์พบว่า ร้อยละ 58 ของคนไทยให้ความสนใจเรื่องอาหาร ร้านอาหาร และการทำอาหาร (Global Web Index Q1-Q2, 2015, อ้างถึงใน CleverKid, WWW, 2558) ทั้งนี้ ในปี พ.ศ. 2560 อิทธิพลเรื่องอาหารยังเป็นเรื่องหนึ่งที่เกิดความนิยมในการใช้งานเครือข่ายสังคมออนไลน์ โดยการเขียนบทวิจารณ์เรื่องอาหาร ซึ่งเป็นสิ่งหนึ่งที่ผู้บริโภคให้ความสนใจและแบ่งปันข้อมูลซึ่งกันและกัน (Thoth Zocial Obivoc, 2561) จากการที่ผู้บริโภคให้ความสำคัญในเรื่องอาหาร และมีพฤติกรรมที่เปลี่ยนแปลงไปตามกาลเวลา ทำให้เกิดการแนะนำอาหารผ่านทางสื่อต่าง ๆ ตั้งแต่อดีตจนถึงปัจจุบัน เพื่อตอบสนองความต้องการของผู้บริโภค และแนะนำอาหารให้เป็นทางเลือกแก่ผู้บริโภคในการเลือกร้านอาหาร

ในอดีตมีการแนะนำอาหาร โดยการวิจารณ์อาหารผ่านสื่อสิ่งพิมพ์ เช่น หนังสือพิมพ์ นิตยสาร ไม่ว่าจะเป็นคำแนะนำร้านอาหารโดยการเขียนเป็นคอลัมน์ในหนังสือพิมพ์ ซึ่งมีชื่อเสียงในการแนะนำร้านอาหาร และได้ขยายการตีพิมพ์ไปสู่ นิตยสาร ในเวลาต่อมา (ไทยรัฐ, WWW, 2555) ต่อมาได้มีการเขียนบทวิจารณ์อาหารในรูปแบบคอลัมน์บนหนังสือพิมพ์ ซึ่งมีชื่อเสียงมากกว่า 40 ปี นับตั้งแต่ปี พ.ศ. 2519 (วิญญู อินทรกำแหง และพฤกัทร ทรงเที่ยง, 2560)

การแนะนำอาหารโดยเขียนบทวิจารณ์ลงในสื่อสิ่งพิมพ์ เป็นการแนะนำร้านอาหารที่ต้องอาศัยประสบการณ์และความชำนาญเรื่องอาหารและการเขียนของผู้วิจารณ์ ทำให้คอลัมน์ดังกล่าวมี

ชื่อเสียงเป็นระยะเวลานาน แต่ข้อจำกัดของการเขียนบทวิจารณ์ในสื่อสิ่งพิมพ์คือ บทวิจารณ์จะถูกตีพิมพ์พร้อมกับสื่อสิ่งพิมพ์ ซึ่งมีระยะเวลาในการตีพิมพ์ที่แตกต่างกัน และมีพื้นที่ในการเขียนบทวิจารณ์ที่จำกัด ทำให้ต้องติดตามอ่านคอลัมน์ต่าง ๆ ที่เกี่ยวข้อง เพื่อให้ได้ร้านอาหารที่ผู้บริโภคต้องการ ทำให้ใช้เวลานานในการอ่านบทวิจารณ์ และได้ข้อมูลร้านอาหารที่อาจไม่เป็นปัจจุบัน ซึ่งปัจจุบันได้พัฒนามาเป็นคอลัมน์ในหนังสือพิมพ์ ที่มาเป็นหนังสือพิมพ์ออนไลน์ (มดิชนสุดสัปดาห์, WWW, 2561) เพื่อตอบสนองต่อพฤติกรรมของผู้บริโภคที่นิยมใช้เครือข่ายสังคมออนไลน์ในการค้นหาข้อมูลเกี่ยวกับร้านอาหาร

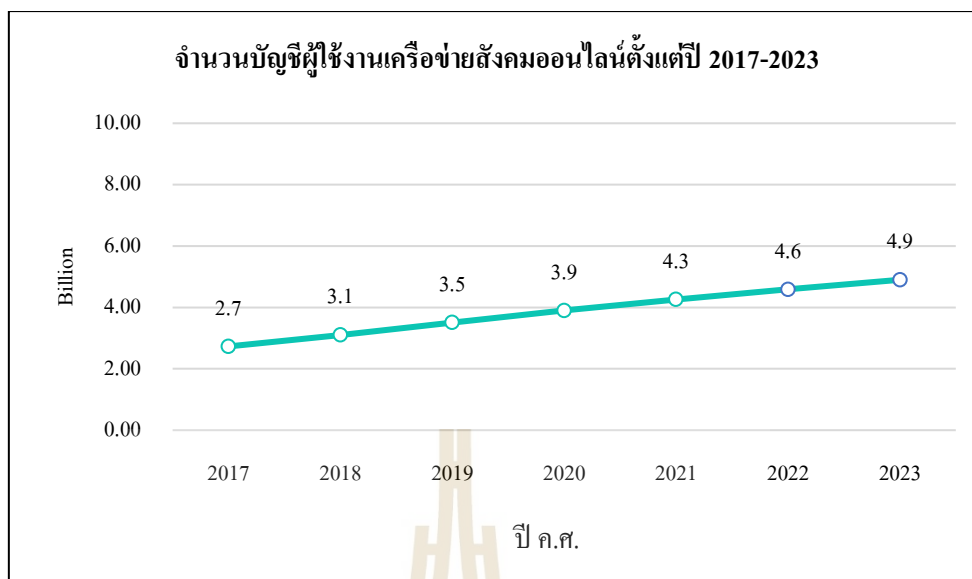
ในยุคต่อมา ได้มีการแนะนำอาหารและวิจารณ์อาหารผ่านทางสื่อวิทยุโทรทัศน์ โดยมีการผลิตรายการที่เกี่ยวกับการแนะนำร้านอาหารอร่อย ซึ่งปัจจุบันได้ปรับเปลี่ยนรายการให้เข้ากับยุคสมัย และตอบสนองความต้องการของผู้บริโภคที่มีพฤติกรรมการใช้เครือข่ายสังคมออนไลน์มากขึ้น นอกจากนี้ ยังมีรายการที่มีชื่อเสียงในด้านการแนะนำร้านอาหาร (Positioningmag, WWW, 2560) จนปัจจุบันมีรายการที่ออกอากาศโดยเชิญเจ้าของร้านอาหารมาสาธิตวิธีการทำอาหารและให้คำวิจารณ์ร้านอาหารดังกล่าว ซึ่งคำวิจารณ์ต่าง ๆ มักเป็นไปในแง่ดี (ไทยรัฐ, WWW, 2560) การวิจารณ์และแนะนำอาหารผ่านสื่อวิทยุโทรทัศน์ได้รับความนิยม และมีการตามรอยรายการต่าง ๆ ไปยังสถานที่จริง แต่อย่างไรก็ตาม การแนะนำร้านอาหารผ่านสื่อวิทยุโทรทัศน์ ยังคงมีข้อจำกัดในเรื่องระยะเวลาในการออกอากาศ เนื่องจากรายการต่าง ๆ มักไม่ใช่รายการที่ออกอากาศสด ต้องมีการเดินทางไปยังร้านอาหาร ซึ่งต้องใช้เวลาในการรอออกอากาศ ข้อมูลที่ออกอากาศจึงไม่เป็นข้อมูลเรียลไทม์ หรือเป็นข้อมูลเก่าที่ทางรายการได้ถ่ายทำผ่านไประยะหนึ่งแล้ว จึงทำให้ผู้บริโภคไม่ได้ข้อมูลที่ตอบสนองความต้องการเกี่ยวกับร้านอาหารอย่างแท้จริง

จากข้อจำกัดของสื่อสิ่งพิมพ์ และสื่อวิทยุโทรทัศน์ในเรื่องตีพิมพ์ และเวลาออกอากาศ รวมถึงเป็นการสื่อสารทางเดียว ผู้ชมไม่สามารถโต้ตอบกับสื่อได้ อีกทั้งในการวิจารณ์ใช้แค่ประสบการณ์ของผู้วิจารณ์เพียงคนเดียว ต่อมาจึงทำให้การเขียนบทวิจารณ์อาหารไม่ได้จำกัดอยู่เพียงแค่สื่อสิ่งพิมพ์ และสื่อวิทยุโทรทัศน์เท่านั้น เมื่อเข้าสู่ยุคของเว็บ 2.0 ทำให้ผู้บริโภคและทีมงานที่ทำหน้าที่ในการเขียนบทวิจารณ์อาหาร สนใจการเขียนบทวิจารณ์อาหารออนไลน์มากขึ้น ซึ่งการเขียนบทวิจารณ์ออนไลน์เกี่ยวกับอาหารในยุคของเว็บ 2.0 สามารถทำได้หลายวิธี และมีช่องทางในการเขียนบทวิจารณ์ออนไลน์ที่หลากหลาย เช่น การเขียนบทวิจารณ์อาหารบนเว็บบล็อก (Web Blog) ซึ่งเว็บบล็อกเป็นแหล่งรวบรวมการวิจารณ์ร้านอาหารของสมาชิกต่าง ๆ ช่วยให้ผู้ใช้งานมีความสะดวกในการค้นหาร้านค้า โดยในการค้นหาร้านค้าต้องค้นหาคำสำคัญ (Keywords) เนื้อหาของบทวิจารณ์มาจากผู้ที่มีประสบการณ์ในการไปรับประทานอาหารร้านนั้น ๆ มีการเชื่อมต่อกับเครือข่ายสังคมต่าง ๆ เช่น เฟสบุ๊ก (Facebook) อินสตาแกรม (Instagram) ทำให้ผู้บริโภคสามารถเข้าถึงบทวิจารณ์ได้หลากหลาย (บุญยาพร วุฒิธรรมคุณ, 2559)

นอกจากนี้ อุตสาหกรรมจำนวนมาก รวมถึงอุตสาหกรรมร้านอาหารต่างก็มองหาวิธีการใหม่ ๆ ที่เป็นนวัตกรรมใหม่เพื่อให้โดดเด่นเหนือคู่แข่ง แม้ว่าความก้าวหน้าทางเทคโนโลยีจะทำให้แบรนด์อาหารและเครื่องดื่มเปิดตัวคุณลักษณะ (Feature) และบริการดิจิทัลในโลกแห่งความเป็นจริงได้ แต่บางแบรนด์ก็กำลังก้าวไปอีกขั้นด้วยการเข้าสู่โลกแห่งความเป็นจริงเสมือนที่เรียกว่า Metaverse (NFTing, WWW, 2022) ซึ่งแบรนด์หรือร้านอาหารบางส่วนของที่เข้าสู่ Metaverse เช่น McDonald's ที่มีแผนการขายอาหารและเครื่องดื่มเสมือนจริง ร้านเบเกอรี่และร้านกาแฟในเครือ Panera Bread ของสหรัฐอเมริกาได้ยื่นคำขอเครื่องหมายการค้าสำหรับ “Paneraverse” เพื่อสร้างพื้นที่ร้านอาหารและร้านกาแฟใน Metaverse หรือแม้กระทั่งร้านเบเกอรี่ที่เพิ่งเปิดตัว Keep It Real Meals ซึ่งเป็นแคมเปญที่ใช้รหัสคิวอาร์ (QR Code) บนกล่องอาหาร เมื่อสแกนแล้วรหัสเหล่านี้จะปลดล็อกของสะสมดิจิทัลและโบนัสที่เป็นโทเคนสำหรับใช้งานบนเครือข่ายดิจิทัล (Non-Fungible Tokens: NFTs) ซึ่งการพัฒนาการบริการดิจิทัลด้านอาหารใน Metaverse ไม่มีกฎหมายตัวว่าแบรนด์อาหารและเครื่องดื่มจะสามารถโต้ตอบกับลูกค้า ผู้ใช้จึงจะได้สัมผัสกับประสบการณ์ที่แปลกใหม่ สำหรับการค้นหาแบรนด์อาหารและเครื่องดื่มได้มากยิ่งขึ้น

อย่างไรก็ตาม เครือข่ายสังคมออนไลน์ที่มีการเขียนบทวิจารณ์ออนไลน์ ผู้เขียนบทวิจารณ์มักเป็นทีมงานเอง ซึ่งเนื้อหาของบทวิจารณ์ส่วนใหญ่มักเป็นไปในด้านดี มีความคิดเห็นที่เป็นด้านลบเกี่ยวกับอาหารเป็นส่วนน้อย เนื่องจากการแนะนำร้านอาหารเพื่อเป็นทางเลือกให้กับผู้บริโภค จึงนำเสนอด้านดีของร้านอาหารนั้น ๆ เป็นส่วนใหญ่ ทำให้ผู้บริโภคได้รับข้อมูลที่เป็นเพียงด้านดีของร้านอาหาร แต่ไม่ทราบว่าร้านใดที่มีข้อมูลที่เป็นด้านลบ เพื่อให้สามารถหลีกเลี่ยงการเลือกหรือรับประทานในร้านอาหารดังกล่าวได้ นอกจากนี้ ในการเขียนบทวิจารณ์ออนไลน์ดังที่กล่าวมา ทำให้ได้ข้อมูลที่ไม่เรียลไทม์ เนื่องจากว่าทีมงานของผู้เขียนบทวิจารณ์จะเดินทางไปยังร้านอาหาร และเขียนบทวิจารณ์ อาจใช้เวลาในการดำเนินการหลายวัน หรือหลายเดือน

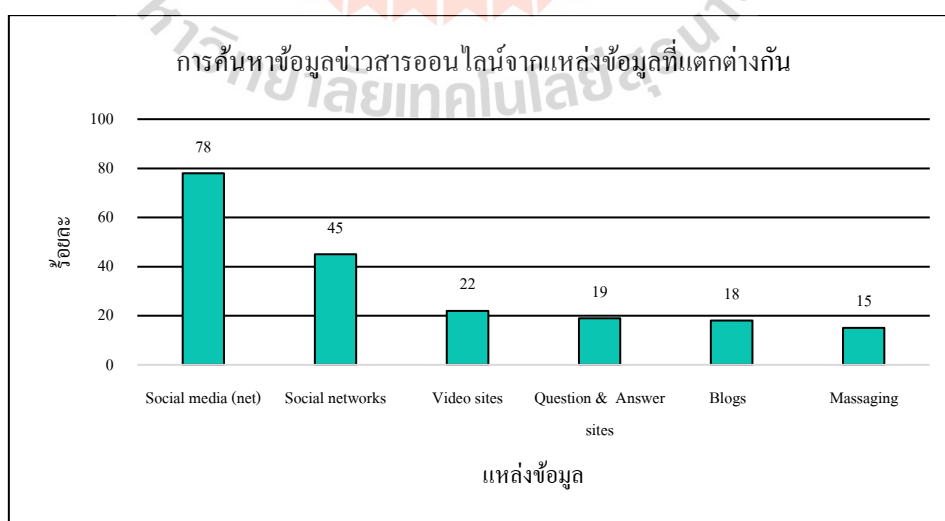
ในปัจจุบัน ผู้บริโภคจำนวนมากนิยมเขียนบทวิจารณ์เกี่ยวกับอาหารด้วยตนเอง เพื่อแบ่งปันประสบการณ์ของตนเองให้กับเพื่อนบนเครือข่ายสังคมออนไลน์ได้รับรู้ และเป็นตัวช่วยในการตัดสินใจเลือกร้านอาหาร ข้อมูลจากสื่อสังคมออนไลน์จึงเป็นปัจจัยสำคัญต่อการเลือกบริโภคของผู้บริโภค พฤติกรรมของผู้บริโภคจึงเปลี่ยนไป คือมีการใช้เครือข่ายสังคมออนไลน์มากขึ้น ทำให้มีบัญชีผู้ใช้เครือข่ายสังคมออนไลน์เพิ่มขึ้น (Global Web Index, 2018) ซึ่งการเพิ่มขึ้นของบัญชีผู้ใช้งานเครือข่ายสังคมออนไลน์ ทำให้เกิดการแบ่งปันข้อมูลข่าวสารที่มากขึ้น ส่งผลให้ผู้บริโภคมีพฤติกรรมเปลี่ยนแปลงไปตามค่านิยม ความต้องการบริโภคสินค้าและบริการ ทำให้ผู้บริโภคใช้เทคโนโลยีเพื่อแลกเปลี่ยนข้อมูล ความคิดเห็นอย่างสม่ำเสมอ ส่งผลให้ค่านิยมของผู้บริโภคมีแนวโน้มที่จะเปลี่ยนแปลงอยู่ตลอดเวลา (Marketeer, WWW, 2560) โดยจำนวนผู้ใช้เครือข่ายสังคมออนไลน์จำนวน 53 เครือข่ายสังคมออนไลน์ ตั้งแต่ปี ค.ศ. 2017 ถึงปี ค.ศ. 2023 มีแนวโน้มเพิ่มขึ้น ดังแสดงในรูปที่ 1.1



ที่มา: Demandsage, WWW, 2023

**รูปที่ 1.1** แผนภูมิแสดงจำนวนบัญชีผู้ใช้งานเครือข่ายสังคมออนไลน์  
ตั้งแต่ปี ค.ศ. 2017 ถึงปี ค.ศ. 2023

จากรูปที่ 1.1 แสดงจำนวนบัญชีผู้ใช้งานเครือข่ายสังคมออนไลน์ตั้งแต่ปี ค.ศ. 2017-2023 มีหน่วยเป็นพันล้าน ทำให้เห็นแนวโน้มว่ามีผู้ใช้งานเครือข่ายสังคมออนไลน์เพิ่มมากขึ้นทุกปี (Demandsage, WWW, 2023) นอกจากนี้ ผู้ใช้ที่เกิดหลังจากปี พ.ศ. 2538-2552 หรือ Generation Z ให้ความสำคัญกับการค้นหาข้อมูลข่าวสารออนไลน์ โดยค้นหาจากแหล่งข้อมูลที่แตกต่างกันในไตรมาสแรกของปี ค.ศ. 2023 (Global Web Index, 2023) แสดงดังรูปที่ 1.2



ที่มา: Global Web Index, 2023

**รูปที่ 1.2** แผนภูมิแสดงการค้นหาข้อมูลข่าวสารออนไลน์จากแหล่งข้อมูลที่แตกต่างกัน

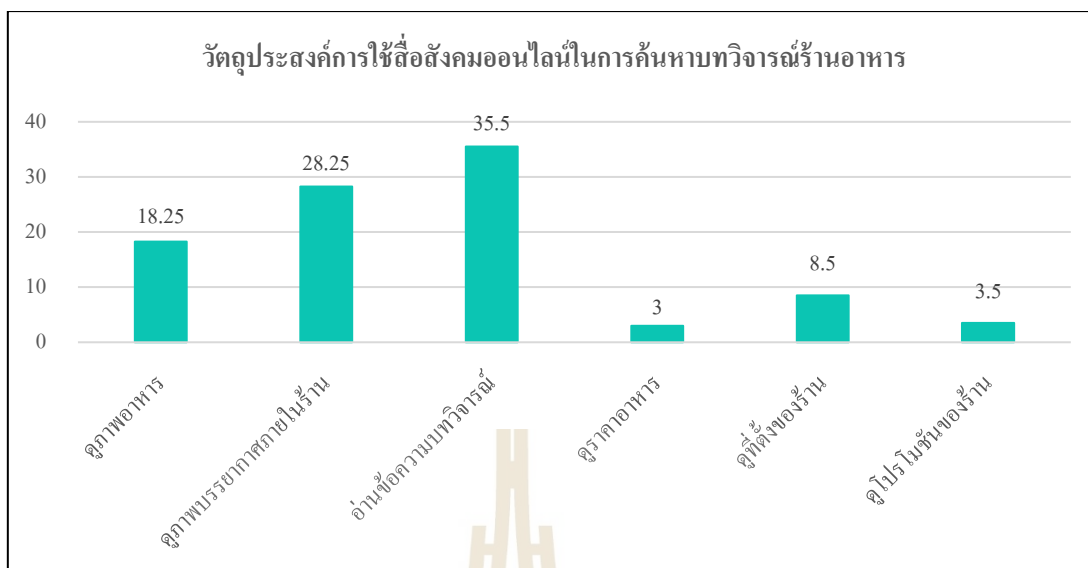
จากรูปที่ 1.2 แสดงการค้นหาข้อมูลข่าวสารออนไลน์จากแหล่งข้อมูลที่แตกต่างกัน ทำให้ทราบว่าผู้ใช้ที่เกิดหลังจากปี พ.ศ. 2538-2552 หรือ Generation Z ใช้แพลตฟอร์มเครือข่ายสังคมออนไลน์ (Social media) ค้นหาข้อมูลข่าวสารออนไลน์ถึงร้อยละ 78 รองลงมาคือ เครือข่ายสังคมออนไลน์ (Social network) ที่มีการใช้ร้อยละ 45 ข้อมูลดังกล่าวแสดงให้เห็นว่า แพลตฟอร์มเครือข่ายสังคมออนไลน์และเครือข่ายสังคมออนไลน์มีบทบาทสำคัญต่อการค้นหาข้อมูลของผู้ใช้ที่เกิดหลังจากปี พ.ศ. 2538-2552 ซึ่งยังคงสามารถใช้งานเครือข่ายสังคมออนไลน์ได้ในอนาคตมากกว่าการค้นหาข้อมูลผ่านเครื่องมือค้นหา ซึ่งในปัจจุบันมีการใช้งานเครือข่ายสังคมออนไลน์เพื่อแบ่งปันข้อมูลระหว่างผู้ใช้อย่างกว้างขวาง โดยผู้ใช้ได้ใช้เครือข่ายสังคมออนไลน์เป็นเครื่องมือในการแบ่งปันประสบการณ์เกี่ยวกับอาหาร และร้านอาหารโดยการเขียนบทวิจารณ์ และแบ่งปันภาพเกี่ยวกับอาหารแก่ผู้ใช้คนอื่น ๆ อีกด้วย

จากการศึกษาปัจจัยที่ส่งผลต่อการเลือกรับประทานอาหารของผู้บริโภคพบว่า ปัจจัยที่มีอิทธิพลต่อการเลือกรับประทานอาหารของผู้บริโภค คือ การแนะนำ การเขียนบทวิจารณ์เกี่ยวกับอาหาร (บุญยาพร วุฒิชรรมคุณ, 2560) และปัจจัยที่มีผลต่อการตัดสินใจเลือกร้านอาหาร ได้แก่ คนรู้จักหรือใกล้ชิดแนะนำร้านอาหาร ที่ตั้งร้านอาหาร และการเดินทาง ซึ่งก่อนและระหว่างมื้ออาหารผู้บริโภคร้อยละ 49 ได้ศึกษาข้อมูลเกี่ยวกับอาหารและร้านอาหารจากเครือข่ายสังคมออนไลน์ (Dancy, WWW, 2012) นอกจากนี้ ผู้บริโภคร้อยละ 25 มักถ่ายภาพอาหารก่อนและขณะรับประทานอาหารและเผยแพร่ผ่านเครือข่ายสังคมออนไลน์เป็นประจำ (Alana, WWW, 2012)

สำหรับการค้นหาข้อมูลจากการเขียนบทวิจารณ์ต่าง ๆ ก็ได้รับความนิยมเช่นเดียวกับการเขียนบทวิจารณ์ โดยจากการสำรวจการค้นหาข้อมูลเพื่ออ่านบทวิจารณ์ของผู้บริโภคในปี ค.ศ. 2016 และ 2017 พบว่า ธุรกิจร้านอาหารหรือคาเฟ่ (Restaurant/Cafe) ได้รับความนิยมในการค้นหาเพื่ออ่านบทวิจารณ์มากที่สุด คิดเป็นร้อยละ 60 (Bright Local, WWW, 2017) จะเห็นได้ว่าผู้บริโภคให้ความสำคัญกับการอ่านบทวิจารณ์เกี่ยวกับอาหารมาเป็นอันดับต้น ๆ และจากการศึกษาปัจจัยที่ส่งผลต่อการอ่านบทวิจารณ์ร้านอาหาร พบว่า เหตุผลในการอ่านบทวิจารณ์ร้านอาหารมากที่สุด คือ เพื่อเปรียบเทียบหรือหาทางเลือกก่อนการตัดสินใจใช้บริการ คิดเป็นร้อยละ 54.3

จากการศึกษาพฤติกรรมในการใช้สื่อสังคมออนไลน์ประเภทต่าง ๆ สำหรับการค้นหามติวิจารณ์เกี่ยวกับร้านอาหาร การใช้เครือข่ายสังคมออนไลน์ เช่น เฟซบุ๊ก (Facebook) อินสตาแกรม (Instagram) เพื่อการค้นหามติวิจารณ์เกี่ยวกับร้านอาหารมีจำนวนการใช้งานสูงที่สุด คิดเป็นร้อยละ 93.25 (บุญยาพร วุฒิชรรมคุณ, 2559) พินเทอเรส (Pinterest) เป็นเครือข่ายสังคมออนไลน์ที่มีผู้ใช้แลกเปลี่ยนข้อมูล แสดงความคิดเห็น เผยแพร่ภาพเกี่ยวกับอาหาร หรือร้านอาหาร และมีอิทธิพลต่อผู้บริโภคในการตัดสินใจเลือกอาหารและเครื่องดื่มมากที่สุด (Thai SME Research, WWW, 2558) ซึ่งผู้บริโภคมีวัตถุประสงค์ในการค้นหามติวิจารณ์ออนไลน์เกี่ยวกับร้านอาหารผ่านเครือข่ายสังคมออนไลน์ที่แตกต่างกัน (บุญยาพร วุฒิชรรมคุณ, 2559) ดังแสดงในรูปที่ 1.3





ที่มา: บุญยาพร วุฒิชัยธรรมคุณ, 2559

### รูปที่ 1.3 วัตถุประสงค์การใช้สื่อสังคมออนไลน์ในการค้นหาทวิจรรย์ร้านอาหาร

จากการศึกษาวัตถุประสงค์ในการค้นหาทวิจรรย์ออนไลน์เกี่ยวกับร้านอาหารผ่านใช้เครือข่ายสังคมพบว่า ผู้บริโภคมีวัตถุประสงค์ในการค้นหาทวิจรรย์เกี่ยวกับร้านอาหารเพื่ออ่านข้อความบทวิจารณ์มากที่สุด รองลงมาคือเพื่อดูภาพบรรยากาศภายในร้านอาหาร และดูภาพอาหารตามลำดับ จึงอธิบายได้ว่า การที่ผู้บริโภคจะตัดสินใจเลือกร้านอาหาร ต้องพิจารณาถึงบทวิจารณ์ก่อนเป็นอันดับแรก รองลงมาคือการดูภาพประกอบทั้งบรรยากาศภายในร้าน และอาหารร่วมด้วย

ผลการศึกษาพฤติกรรมผู้ใช้แอปพลิเคชัน ไอพิก<sup>1</sup> (iPick, WWW, 2015) ซึ่งเป็นโปรแกรมประยุกต์สำหรับการแบ่งปันข้อมูลเกี่ยวกับอาหารพบว่า ผู้บริโภคมีพฤติกรรมในการค้นหาร้านอาหารจากการค้นหาข้อมูลร้านอาหารผ่านเว็บเป็นอันดับแรก ลำดับต่อมาคือการอ่านบทวิจารณ์ร้านอาหารจากเว็บบอร์ด หลังจากนั้นคือการค้นหาที่อยู่ และเบอร์โทรจากเว็บไซต์ร้านอาหาร และใช้โปรแกรมประยุกต์สำหรับการสนทนาเพื่อชักชวนเพื่อน แล้วจึงค้นหาแผนที่บนเว็บไซต์ร้านอาหารอีกครั้งเพื่อเดินทางไปยังร้านอาหาร และสุดท้ายจะจบลงด้วยการแบ่งปันประสบการณ์การรับประทานอาหารผ่านเครือข่ายสังคมออนไลน์ โดยผู้บริโภคจะมีพฤติกรรมเช่นนี้วนไปเรื่อย ๆ แม้แต่การบอกต่อด้วยปากต่อปาก ก็ยังคงมีพฤติกรรมเช่นนี้ในการค้นหาร้านอาหาร นอกจากนี้ ยังมีการเผยแพร่ภาพเกี่ยวกับอาหาร และร้านอาหารร่วมกับบทวิจารณ์อีกด้วย จากที่กล่าวมาข้างต้น จะเห็นว่า เครือข่ายสังคมออนไลน์มีบทบาทสำคัญอย่างยิ่งในการแบ่งปันข้อมูลเกี่ยวกับประสบการณ์ในการรับประทานอาหารของผู้บริโภค จึงทำให้มีการแบ่งปันประสบการณ์ในการรับประทานอาหารแก่ผู้อื่นผ่านเครือข่ายสังคมออนไลน์

<sup>1</sup> <http://www.ipick.com/bangkok/home>

จากการศึกษาปัจจัยที่มีผลต่อความพึงพอใจในการใช้บริการร้านอาหารยังมีความเกี่ยวข้องกับปัจจัยส่วนประสมทางการตลาด (7P's) ประกอบด้วย ปัจจัยด้านผลิตภัณฑ์ (Product) ปัจจัยด้านราคา (Price) ปัจจัยด้านบุคลากร (People) ปัจจัยด้านกระบวนการ (Process) ปัจจัยด้านส่งเสริมการตลาด (Promotion) ปัจจัยด้านสถานที่ (Place) ปัจจัยด้านลักษณะทางกายภาพ (Physical Evidence) โดยปัจจัยที่ส่งผลต่อความพึงพอใจในการใช้บริการร้านอาหารมากที่สุด คือ ปัจจัยปัจจัยด้านบุคลากร ปัจจัยด้านกระบวนการและความสะอาด และปัจจัยด้านราคาตามลำดับ (วิชาทองลัพท์, 2559) ซึ่งปัจจัยดังกล่าวเป็นปัจจัยที่เกิดจากประสบการณ์การให้บริการร้านอาหารของผู้บริโภคที่มีผลต่อความนิยม และการเติบโตของร้านอาหาร โดยเมื่อผู้บริโภคใช้บริการร้านอาหารทำให้เกิดการแบ่งปันประสบการณ์ในการใช้บริการ (รัชนี พันธุ์รุ่งจิตติ และคณะ, 2560)

การที่เครือข่ายสังคมออนไลน์ได้รับความนิยมในการแบ่งปันเกี่ยวกับอาหาร ประกอบกับพฤติกรรมของคนไทยที่นิยมถ่ายรูป โดยมีบอกเล่าประสบการณ์ และวิจารณ์อาหารผ่านเครือข่ายสังคมออนไลน์ ทำให้ข้อมูลเกี่ยวกับอาหาร และร้านอาหารมีจำนวนมากและกระจัดกระจาย (รัชนี พันธุ์รุ่งจิตติ และคณะ, 2560) ซึ่งยากต่อการที่ผู้ใช้จะสามารถค้นหาข้อมูลได้ตรงกับความต้องการได้ โดยในการรับประทานอาหารนอกบ้าน 1 ครั้ง ผู้บริโภคใช้เครือข่ายสังคมออนไลน์ 1-2 ประเภทในการค้นหาทวิจาณ์เกี่ยวกับร้านอาหาร คิดเป็นร้อยละ 61.10 และใช้ระยะเวลาเพื่อค้นหาทวิจาณ์เกี่ยวกับร้านอาหาร 1-2 ชั่วโมง คิดเป็นร้อยละ 57.20 (บุญยาพร วุฒิชรรมคุณ, 2559) เนื่องจากมีจำนวนผู้ใช้เครือข่ายสังคมออนไลน์จำนวนมาก และมีเครือข่ายสังคมออนไลน์หลายประเภท จึงทำให้ผู้บริโภคต้องใช้เวลาาน และต้องค้นหาจากเครือข่ายสังคมออนไลน์หลากหลายประเภท เพื่อให้ได้ข้อมูลที่ตรงกับความต้องการ

จากปัญหาดังกล่าว จึงทำให้มีผู้สนใจศึกษาเกี่ยวกับการสกัดคุณลักษณะของอาหาร และร้านอาหาร เพื่อแนะนำร้านอาหารแก่ผู้บริโภคเป็นจำนวนมาก โดยการสกัดข้อมูลจากทวิจาณ์ออนไลน์และสกัดข้อมูลจากภาพที่เผยแพร่บนเครือข่ายสังคมออนไลน์ การวิเคราะห์ความรู้สึกของผู้ใช้ที่มีต่อร้านอาหารจากการให้คะแนนบนเว็บไซต์ต่าง ๆ (Dixon, N, et.al, 2012; Pugsee and Niyomvanich, 2015; Aye, and Aung, 2017) ด้วยวิธีการที่แตกต่างกัน ทั้งการทำเหมืองข้อมูล (Data Mining) (Tama, 2015) การทำเหมืองข้อความ (Text Mining) (Surjandari, Naffisah and Prawiradinata, 2015) การทำเหมืองความคิดเห็น (Opinion Mining) (Perera, and Caldera, 2017) และการใช้เทคนิคทางด้านการเรียนรู้แบบเชิงลึก (Deep Learning) ในการจำแนกประเภทภาพที่ผู้ใช้เผยแพร่ผ่านเครือข่ายสังคมออนไลน์ (Govindarajan, 2014; Raul, Shah and Devganiya, 2016; Chu and Tsai, 2017) การแนะนำร้านอาหารที่ตรงกับความต้องการของผู้ใช้ โดยพิจารณาจากข้อมูลการแบ่งปัน การแสดงความคิดเห็น การเยี่ยมชมเว็บไซต์ต่าง ๆ ของผู้เข้าร่วมด้วย (Higgs, and Thomas, 2016; Chu, and Tsai, 2017)



อย่างไรก็ตาม การแสดงผลข้อมูลเรื่องอาหาร และร้านอาหารที่ได้รับความนิยม จะทำให้สามารถตอบสนองต่อความต้องการของผู้บริโภค และช่วยให้ผู้บริโภคตัดสินใจเลือกรับประทานอาหาร และร้านอาหารได้รวดเร็วยิ่งขึ้น นอกจากนี้ การที่ผู้บริโภคต้องใช้เครื่องมือในการค้นหาบทวิจารณ์เกี่ยวกับอาหารในแต่ละครั้ง ทำให้เสียเวลา โดยใช้เวลาประมาณ 1-2 ชั่วโมง จึงจะสามารถค้นหาร้านอาหารหรืออาหารที่ตรงกับความต้องการได้ (บุญยาพร วุฒิธรรมคุณ, 2559) อีกทั้ง บทวิจารณ์ออนไลน์ที่เขียนโดยทีมงานที่ทำหน้าที่ในการแนะนำร้านอาหาร นำเสนอเพียงด้านดีของร้านอาหารเป็นส่วนใหญ่ และให้ข้อมูลที่ไม่เรียลไทม์แก่ผู้บริโภคดังที่กล่าวมาข้างต้น ทำให้ผู้บริโภคได้อ่านบทวิจารณ์ที่เป็นเพียงด้านดีของร้านอาหาร ซึ่งผู้บริโภคจำนวนมากในปัจจุบันนิยมเขียนบทวิจารณ์เกี่ยวกับอาหารด้วยตนเอง แต่การอ่านบทวิจารณ์เกี่ยวกับอาหารในแต่ละครั้ง ทำให้ใช้เวลานานในการตัดสินใจเลือกร้านอาหาร

นอกจากนี้ จากงานวิจัยที่ผ่านมาพบข้อจำกัดเกี่ยวกับการแนะนำร้านอาหาร คือ ประสิทธิภาพด้านความถูกต้องในการแนะนำยังอยู่ในระดับที่สามารถเพิ่มประสิทธิภาพได้ หากสามารถอธิบายความสัมพันธ์ระหว่างร้านอาหาร หรือเรียนรู้คุณลักษณะจากภาพที่มีความแตกต่างกันมากขึ้น (Chu and Tsai, 2017) นอกจากนี้ด้านความถูกต้องในการแนะนำแล้ว ยังมีเรื่องการจำแนกประเภทร้านอาหารที่มีการจำแนกร้านอาหารออกเป็นประเภทต่าง ๆ และสรุปว่าร้านอาหารประเภทนั้นมีข้อดีหรือข้อเสียอย่างไร (Yu, Zhou, Zhang, and Cao, 2017) อีกทั้งยังเน้นไปที่กลุ่มนักท่องเที่ยวที่มีการเขียนและอ่านบทวิจารณ์บนเว็บไซต์ แต่ยังไม่มีการแนะนำร้านอาหารที่เจาะจงถึงสถานที่และบ่งบอกคุณลักษณะเด่นและด้อยของร้านอาหารนั้น ๆ ดังนั้น งานวิจัยนี้จึงรวบรวมบทวิจารณ์และภาพเกี่ยวกับอาหาร เพื่อจำแนกประเภทบทวิจารณ์ และประเภทอาหารที่ทำให้ผู้บริโภคตัดสินใจเลือกรับประทานอาหารที่ร้านอาหาร และแนะนำร้านอาหารที่มีคุณลักษณะตามความต้องการของผู้บริโภค

## 1.2 วัตถุประสงค์ของการวิจัย

- 1.2.1 เพื่อพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

## 1.3 สมมุติฐานการวิจัย

- 1.3.1 แบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์มีความถูกต้อง (Accuracy) ในการแนะนำร้อยละ 80 ขึ้นไป

#### 1.4 ข้อตกลงเบื้องต้น

- 1.4.1 ข้อมูลบทวิจารณ์และภาพเรื่องอาหารบนเครือข่ายสังคมออนไลน์ เป็นข้อมูลจากทวิตเตอร์ (เอ็กซ์: X) เท่านั้น
- 1.4.2 การสกัดข้อมูล ใช้ข้อมูลจากบทวิจารณ์ที่เป็นข้อความ และภาพบนเครือข่ายสังคมออนไลน์เท่านั้น

#### 1.5 ขอบเขตของการวิจัย

การพัฒนาแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ พิจารณาข้อมูลการใช้เครือข่ายสังคมออนไลน์ของผู้ใช้ร่วมกับคุณลักษณะของร้านอาหารที่ทำให้ผู้บริโภคตัดสินใจเลือกรับประทานอาหาร และจำแนกประเภทของบทวิจารณ์และภาพที่เกี่ยวข้องกับอาหารและร้านอาหาร และนำเสนอข้อมูลดังกล่าวแก่ผู้ใช้

#### 1.6 ประโยชน์ที่คาดว่าจะได้รับ

ได้แบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ที่สามารถแนะนำร้านอาหาร ตำแหน่งร้านอาหาร และเมนูอาหารได้อย่างถูกต้อง

#### 1.7 คำอธิบายศัพท์

- 1.7.1 แบบจำลองการแนะนำร้านอาหาร หมายถึง แบบจำลองการแนะนำร้านอาหารจากการสกัดบทวิจารณ์และภาพเรื่องอาหารประกอบด้วยบทวิจารณ์ที่เป็นข้อความ และภาพที่ผู้ใช้เผยแพร่ผ่านเครือข่ายสังคมออนไลน์
- 1.7.2 เครือข่ายสังคมออนไลน์ หมายถึง เครือข่ายสังคมออนไลน์ที่งานวิจัยนี้พิจารณาในการสกัดบทวิจารณ์และภาพ คือ ทวิตเตอร์ (เอ็กซ์: X)
- 1.7.3 คุณลักษณะอาหาร หมายถึง ประเภทของอาหารที่ได้จากการสกัดข้อมูลจากเครือข่ายสังคมออนไลน์
- 1.7.4 บทวิจารณ์ หมายถึง บทวิจารณ์เกี่ยวกับอาหารหรือร้านอาหารที่วิจารณ์โดยผู้ใช้งานเครือข่ายสังคมออนไลน์

## บทที่ 2

### ปรัทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

งานวิจัย เรื่อง การพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบน  
เครือข่ายสังคมออนไลน์ ได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ดังนี้

- 2.1 ธุรกิจอาหารและเครื่องดื่ม
- 2.2 เครือข่ายสังคมออนไลน์ (Online Social Media)
- 2.3 บทวิจารณ์ออนไลน์ (Online Reviews)
- 2.4 การประมวลผลทางภาษา (Word segmentation)
- 2.5 การประมวลผลภาพดิจิทัล (Digital Image Processing)
- 2.6 โยโล่ (You Only Look Once: YOLO)
- 2.6 การเรียนรู้ของเครื่อง (Machine Learning)
- 2.7 การเรียนรู้เชิงลึก (Artificial Neural Network)
- 2.8 งานวิจัยที่เกี่ยวข้อง
- 2.9 กรอบแนวคิดการวิจัย

โดยจากการศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง มีรายละเอียดดังนี้

#### 2.1 ธุรกิจอาหารและเครื่องดื่ม

##### 2.1.1 ความหมายของธุรกิจอาหารและเครื่องดื่ม

ธุรกิจอาหารและเครื่องดื่ม หมายถึง กิจการที่ให้บริการอาหารและเครื่องดื่มแก่ทุกคน โดยมี  
สถานที่ให้บริการสำหรับผู้บริโภค หรือมีการบรรจอาหารที่ผู้บริโภคไม่จำเป็นต้องนั่งรับประทาน  
อาหารในสถานที่นั้น ๆ (Tourism Industry DPU, 2553)

### 2.1.2 ความเป็นมาของธุรกิจร้านอาหารและเครื่องดื่ม

ธุรกิจอาหารและเครื่องดื่มเป็นธุรกิจที่มีมาเป็นระยะเวลานานตั้งแต่สมัยโบราณ โดยทวีปยุโรปมีอิทธิพลต่อการพัฒนารูปแบบต่าง ๆ ของธุรกิจดังกล่าว ซึ่งธุรกิจอาหารแบ่งออกเป็น 3 ยุค ได้แก่ ยุคต้น ยุคกลาง และยุคปัจจุบัน (นิศา ชัชกุล, 2550) โดยมีรายละเอียด ดังนี้

1) ยุคต้น ในประเทศอียิปต์ 512 ปีก่อนคริสตศักราช ผู้คนเริ่มรับประทานอาหารนอกบ้าน ซึ่งในขณะนั้นมีร้านอาหารอียิปต์ที่จำหน่ายอาหารเพียงอย่างเดียว ส่วนในประเทศจีนได้มีการจำหน่ายอาหารมาเป็นเวลาสองพันปีมาแล้ว นอกจากนี้ จากการขุดซากเมือง ปอมเปอี ในประเทศอิตาลี ทำให้พบซากร้านอาหารเป็นจำนวนมาก

2) ยุคกลาง ในยุคนี้ มีการบริการอาหารที่มีวิธีการปรุงแบบฝรั่งเศส ที่ผู้บริโภครับประทานอาหารตามที่ต้องการได้โดยไม่ขึ้นอยู่กับเวลา และมีการปรับเปลี่ยนวิธีการให้บริการแก่นักท่องเที่ยว โดยให้บริการอาหารและเครื่องดื่มที่อยู่ในรายการของสถานที่นั้น ๆ ทำให้มีร้านอาหารฝรั่งเศสที่เป็นที่รู้จักแก่ผู้คนเป็นจำนวนมาก นำไปสู่การตั้งชื่อร้านอาหารเป็นภาษาฝรั่งเศสในประเทศอังกฤษ

3) ยุคปัจจุบัน ได้ปรับเปลี่ยนรูปแบบการให้บริการ บรรยากาศ และสถานที่ที่ทันสมัย เพื่อให้สอดคล้องกับรสนิยมหรือดำเนินชีวิตของผู้บริโภคในยุคปัจจุบัน โดยธุรกิจอาหารและเครื่องดื่มที่ได้รับความนิยมในปัจจุบัน ได้แก่ ธุรกิจประเภทผับ (Pub) บาร์ (Bar) ไนต์คลับ (Night Club) สวนอาหาร และร้านอาหาร

### 2.1.3 ประเภทของธุรกิจอาหารและเครื่องดื่ม

ประเภทของธุรกิจร้านอาหารและเครื่องดื่มออกในปัจจุบันออกได้ถูกแบ่งออกเป็น 7 ประเภท โดย นิศา ชัชกุล (2550) ดังนี้

1) ธุรกิจอาหารจานด่วน (Fast-Food Restaurants) เป็นธุรกิจที่ให้บริการแก่ผู้บริโภคที่มีความต้องการในการใช้เวลาและงบประมาณที่จำกัด เร่งรีบ เนื่องจากเป็นธุรกิจที่ให้ความสะดวก รวดเร็ว เปิดบริการอย่างต่อเนื่องทุกคน มีอาหารตามเชื้อชาติที่ให้บริการในรูปแบบอาหารจานเดียว และมีราคาไม่แพง

2) ธุรกิจอาหารสำเร็จรูป (Deli Shops) เป็นอาหารประเภทแช่แข็ง เนย แชนวิช สลัด เป็นต้น ซึ่งส่วนใหญ่ให้บริการในบริเวณที่มีผู้คนพลุกพล่าน มีระยะเวลาในการเปิดและไต่มนั่งที่จำกัด

3) ธุรกิจอาหารบุฟเฟต์ (Buffets) เป็นธุรกิจที่ได้รับความนิยมในสถานที่พักที่เป็นโรงแรม ธุรกิจประเภทนี้ให้ผู้บริโภคสามารถเลือก และตักอาหารได้ตามปริมาณและความต้องการ โดยคิดค่าบริการเป็นจำนวนคน

**4) ธุรกิจประเภทคอฟฟี่ช็อป (Coffee Shops)** เป็นธุรกิจที่มักตั้งอยู่ภายในห้างสรรพสินค้า หรือชุมชนที่ผู้คนผ่านไปผ่านมาเป็นจำนวนมาก ราคาถูก ให้บริการแก่ผู้บริโภคที่ต้องการความรวดเร็วในการรับประทานอาหาร

**2.1.3.5 ธุรกิจอาหารเฉพาะกลุ่มเชื้อชาติ (Ethnic Restaurants)** เป็นธุรกิจอาหารของแต่ละเชื้อชาติ ตั้งอยู่ในท้องถิ่นของชาตินั้น ๆ มีจุดเด่นที่การให้บริการที่เป็นเอกลักษณ์ของชาตินั้น ๆ ทั้งอาหาร บรรยากาศภายในร้าน และพนักงานที่ให้บริการ

**2.1.3.6 ธุรกิจประเภทคาเฟ่ที่เรีย (Cafeterias Self Services)** เป็นธุรกิจที่รองรับผู้บริโภคได้จำนวนมากในเวลาอันสั้น โดยผู้บริโภคต้องบริการตนเอง ไม่มีบริการสำหรับให้บริการแก่ผู้บริโภค

**2.1.3.7 ธุรกิจอาหารгур์เมต์ (Gourmet Restaurants)** เป็นธุรกิจที่ดึงดูดผู้บริโภคด้วยสถานที่ที่มีความหรูหรา มีเป้าหมายเป็นกลุ่มผู้บริโภคที่มีฐานะดี ให้บริการที่ดีกว่าธุรกิจประเภทอื่น ๆ แต่มีราคาแพง

จากความเป็นมาและประเภทของธุรกิจร้านอาหารและเครื่องดื่ม แสดงให้เห็นถึงการพัฒนาด้านธุรกิจร้านอาหารและเครื่องดื่มที่ปรับเปลี่ยนให้สอดคล้องกับพฤติกรรมของผู้บริโภคในแต่ละยุคสมัย สร้างความหลากหลาย และเป็นที่นิยมแก่ผู้บริโภคเป็นจำนวนมาก งานวิจัยนี้จึงพิจารณาธุรกิจอาหารและเครื่องดื่มที่ให้บริการแก่ผู้บริโภคที่ไม่เฉพาะเจาะจง เนื่องจากผู้บริโภคทุกคนสามารถเข้าถึง ได้รับบริการ และสามารถแบ่งปันข้อมูลด้านอาหารหรือร้านอาหารผ่านเครือข่ายสังคมออนไลน์ได้

## 2.2 สื่อสังคมออนไลน์

### 2.2.1 ความหมายของสื่อสังคมออนไลน์ (Social Media)

ราชบัณฑิตยสถาน (2554) ได้บัญญัติคำว่า สื่อสังคม ว่า “สื่อชนิดหนึ่งที่ใช้สามารถมีส่วนร่วม สร้าง และแลกเปลี่ยนความคิดเห็นต่าง ๆ ผ่านระบบอินเทอร์เน็ตได้ ในทางเทคนิค สื่อสังคม จะหมายถึง โปรแกรมกลุ่มหนึ่งที่ทำงานโดยใช้พื้นฐานและเทคโนโลยีของเว็บตั้งแต่รุ่น 2.0 เช่น บีโอบี มายสเปซ เฟซบุ๊ก ทวิตเตอร์ วิกีพีเดีย ไฮไฟฟ์ และบล็อกต่าง ๆ ในทางธุรกิจ สื่อสังคมคือสื่อที่ผู้บริโภคสร้างขึ้น (Consumer-Generated Media หรือ CGM) สำหรับกลุ่มบุคคลที่ติดต่อสื่อสารกันผ่านสื่อสังคม เพื่อแลกเปลี่ยนข้อมูลข่าวสาร และทำกิจกรรมที่สนใจร่วมกัน”

พิชิต วิจิตรบุญยรักษ์ (2554) ได้ให้ความหมายของสื่อสังคมออนไลน์ว่า เป็นสื่อที่ผู้ส่งสารแบ่งปันสารในรูปแบบต่าง ๆ ไปยังผู้รับสารผ่านเครือข่ายสังคมออนไลน์ โดยสามารถโต้ตอบกันระหว่างผู้ส่งสารและผู้รับสาร หรือผู้รับสารด้วยตนเอง

เอลิซาเบธ (Elizabeth, 2012) ได้ให้ความหมายของสื่อสังคมออนไลน์ว่า หมายถึง สื่อดิจิทัล หรือ โปรแกรมที่ทำงานอยู่บนพื้นฐานของเว็บไซต์บนอินเทอร์เน็ต เป็นเครื่องมือในการปฏิบัติการทางสังคม ที่มีผู้สื่อสารจัดทำขึ้น โดยที่ผู้เขียนจัดทำขึ้นเอง และมีการแบ่งปันข้อมูลข่าวสารกับผู้ที่อยู่ภายในสังคมเดียวกันได้อย่างรวดเร็ว

จากความหมายของเครือข่ายสังคมออนไลน์ สามารถสรุปได้ว่า สื่อสังคมออนไลน์ หมายถึง สังคมที่มีการรวมตัวกันเพื่อสร้างความสัมพันธ์บนอินเทอร์เน็ตหรือที่เรียกว่า ชุมชนออนไลน์ (Community Online) มีลักษณะเป็นสังคมเสมือน (Virtual Community) ที่ผู้คนสามารถทำ ความรู้จัก แลกเปลี่ยนความคิด แบ่งปันประสบการณ์ และเชื่อมโยงกันในทุกทิศทางใดทิศทางหนึ่ง ผ่านการติดต่อสื่อสารกันเป็นเครือข่าย โดยงานวิจัยนี้ได้พัฒนาระบบแนะนำร้านอาหารจากบท วิเคราะห์และภาพอาหารบนเครือข่ายสังคมออนไลน์ และพิจารณาข้อมูลการใช้งานเครือข่ายสังคม ออนไลน์ของผู้ใช้ เพื่อให้สามารถเชื่อมโยงข้อมูลของผู้ใช้บนเครือข่ายสังคมออนไลน์และสามารถ แนะนำร้านอาหารที่ตรงกับความต้องการของผู้ใช้ได้

## 2.2.2 ประเภทของเครือข่ายสังคมออนไลน์

เขมณัญญ์ ศิริธรรม (2556) ได้แบ่งประเภทของเครือข่ายสังคมออนไลน์ที่ผู้ใช้ สามารถแสดงความคิดเห็น หรือแลกเปลี่ยนข้อมูลข่าวสารได้ ออกเป็น 5 ประเภท ได้แก่ บล็อก (Blog) หรือเว็บบล็อก ทวิตเตอร์ (Twitter) เครือข่ายสังคม (Social Network) สื่อแบ่งปัน (Media Sharing) และออนไลน์ฟอรัม (Online Forum) โดยประเภทของเครือข่ายสังคมออนไลน์ตามแนวคิด ของ เขมณัญญ์ ศิริธรรม (2556) มีดังนี้

1) บล็อก (Blog) หรือเว็บบล็อก (Web Blog) เป็นระบบจัดการเนื้อหาที่ให้ผู้ ใช้สามารถเขียนเนื้อหาเอง สามารถเผยแพร่แ่งมุมของการตลาด ถูกนำมาใช้ในการสื่อสารการตลาด 2 รูปแบบ คือ สร้างตราสินค้าเพื่อสื่อสารกับผู้บริโภคและบล็อกที่เขียนจากผู้เขียนบล็อก (Blogger) อิสระที่มีผู้ติดตามจำนวนมากจนกลายเป็นผู้มีอิทธิพลทางการตลาด (Marketing Influencer)

2) ทวิตเตอร์ (Twitter) เป็นไมโครบล็อก (Micro Blog) ที่จำกัดตัวอักษรในการ โพสต์ครั้งละ 140 ตัวอักษร ถูกนำไปใช้ทางธุรกิจ ทั้งเพื่อสร้างการบอกต่อ การเพิ่มยอดขาย การ สร้างแบรนด์ หรือเป็นเครื่องมือในการบริหารความสัมพันธ์กับลูกค้า รวมถึงใช้เป็นเครื่องมือในการ ประชาสัมพันธ์

3) เครือข่ายสังคม (Social Network) หมายถึง เครือข่ายสังคมที่ถูกเชื่อมโยงในโลก ออนไลน์ จนกลายเป็นสังคมเสมือน ซึ่งผู้ใช้งานสามารถสร้างตัวตนขึ้นเอง โดยมีข้อมูลส่วนตัว รูป การ จดบันทึก หรือการไลฟ์วิดีโอ เป็นต้น สามารถเพิ่มจำนวนเพื่อนได้ ทำให้นักการตลาดนำเครือข่าย สังคมมาใช้ในการสร้างปฏิสัมพันธ์กับลูกค้า รวมถึงจัดกลุ่มความสนใจของคนในวงสังคมได้

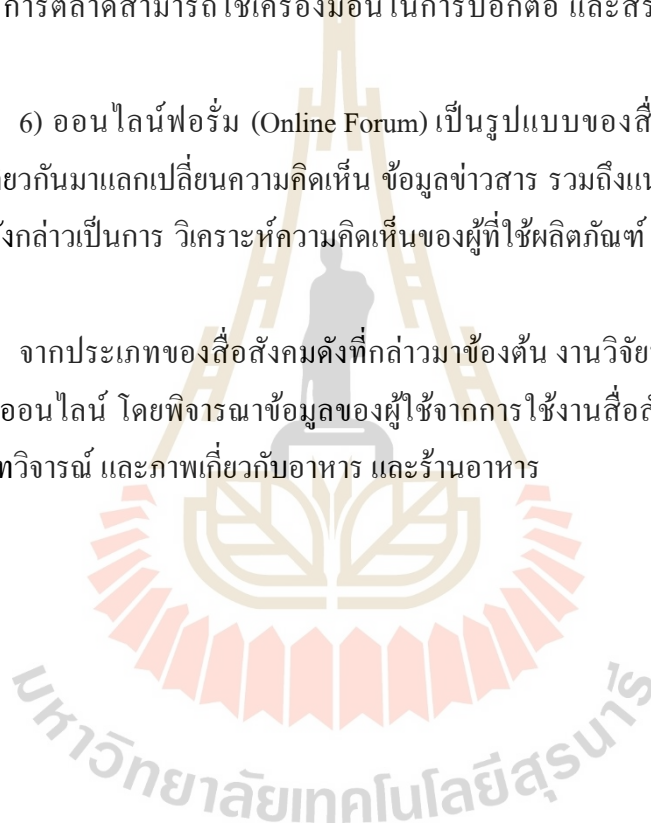


4) สื่อแบ่งปัน (Media Sharing) เป็นเว็บไซต์ที่เปิดโอกาสให้อัปโหลดรูปหรือวิดีโอเพื่อแบ่งปันต่อสาธารณะ และเป็นช่องทางสำหรับนักสื่อสารการตลาดที่สามารถสื่อสารให้ต้นทุนถูกลงและประหยัดงบประมาณในการซื้อสื่อโฆษณา อีกทั้งช่องทางนี้ยังสามารถทำให้เกิดการบอกต่อได้อย่างรวดเร็ว ซึ่งสื่อประเภทแบ่งปันที่ได้รับความนิยมในประเทศไทย เช่น YouTube Multiply เป็นต้น

5) ข่าวสังคมและบุ๊กมาร์ก (Social News and Bookmarking) เป็นเว็บไซต์ที่ประชาชนทั่วไปมีโอกาสให้คะแนนและโหวตเว็บไซต์ที่ชื่นชอบ ซึ่งสามารถเก็บไว้เป็นบุ๊กมาร์กได้ โดยนักสื่อสารการตลาดสามารถใช้เครื่องมือนี้ในการบอกต่อ และสร้างจำนวนคนที่เข้ามายังเว็บไซต์

6) ออนไลน์ฟอรัม (Online Forum) เป็นรูปแบบของสื่อสังคมที่รวมกลุ่มคนที่สนใจในเรื่องเดียวกันมาแลกเปลี่ยนความคิดเห็น ข้อมูลข่าวสาร รวมถึงแนะนำสินค้าต่าง ๆ นักการตลาดจึงใช้สื่อดังกล่าวเป็นการ วิเคราะห์ความคิดเห็นของผู้ที่ใช้ผลิตภัณฑ์ หรือแนะนำกลุ่มต่าง ๆ ที่มีความสนใจ

จากประเภทของสื่อสังคมดังกล่าวมาข้างต้น งานวิจัยนี้ได้ใช้สื่อสังคมประเภทเครือข่ายสังคมออนไลน์ โดยพิจารณาข้อมูลของผู้ใช้จากการใช้งานสื่อสังคมออนไลน์ทวิตเตอร์ ประกอบด้วยบทวิจารณ์ และภาพเกี่ยวกับอาหาร และร้านอาหาร



## 2.3 บทวิจารณ์ออนไลน์

### 2.3.1 ความหมายของบทวิจารณ์ออนไลน์

บทวิจารณ์ออนไลน์ หมายถึง ความคิดเห็นส่วนตัวของผู้เขียนในเรื่องที่ตนมีประสบการณ์พบเจอมาซึ่งอาจแตกต่างจากผู้อื่นได้ตามความชื่นชอบส่วนบุคคล (บุญยาพร วุฒิธรรมคุณ, 2559)

### 2.3.2 บทวิจารณ์ออนไลน์บนเครือข่ายสังคมออนไลน์

ในการเขียนบทวิจารณ์ออนไลน์สามารถแบ่งได้ตามประเภทของเครือข่ายสังคมออนไลน์ ดังนี้ (บุญยาพร วุฒิธรรมคุณ, 2559)

1) บล็อก (Blog) หรือเว็บบล็อก (Web Blog) คือเว็บไซต์ให้พื้นที่ส่วนตัวแก่ผู้ให้บริการ โดยมีฟังก์ชันต่าง ๆ ให้เลือกใช้งาน เช่น พื้นที่สำหรับเขียนประวัติ เขียนบทความ โพสต์รูปหรือวิดีโอ กระดาน ปรับแต่งได้เหมือนเว็บไซต์ สามารถจดทะเบียนเป็นชื่อเจ้าของได้ เชื่อมต่อเครือข่ายสังคมได้ บล็อกที่เกี่ยวข้องกับการวิจารณ์ร้านอาหารและเป็นที่รู้จักกันอย่างกว้างขวางในสังคมออนไลน์ เช่น วงใน (Wongnai)

2) ทวิตเตอร์ (Twitter) คือ เครือข่ายสังคมออนไลน์ประเภทไมโครบล็อก ที่ผู้ใช้ต้องทำการลงทะเบียนสมัครสมาชิกก่อนใช้งานและสามารถส่งข้อความ (Tweet) ยาวไม่เกิน 140 ตัวอักษร

3) เครือข่ายสังคม (Social Network) เป็นอีกหนึ่งช่องทางในการค้นหาบทวิจารณ์ร้านอาหาร และได้รับความนิยมเป็นอย่างมากในปัจจุบัน ผู้ใช้สามารถติดตามอ่านรีวิวร้านอาหารได้ โดยการเป็นเพื่อนหรือกดถูกใจ

4) สื่อแบ่งปัน (Media Sharing) ได้รับความนิยมจากผู้ทั่วโลก ให้บริการทั้งในรูปแบบเว็บไซต์และแอปพลิเคชัน โดยผู้ใช้สามารถอัปโหลดภาพวิดีโอ เปิดดูภาพวิดีโอ และแบ่งภาพวิดีโอให้ผู้อื่นได้โดยมีเสียค่าใช้จ่ายใด ๆ สำหรับผู้ที่ค้นหาการบทวิจารณ์ร้านอาหารนั้นสามารถทำได้โดยการพิมพ์คำสำคัญในช่องค้นหา

5) ข่าวสังคมและบุ๊กมาร์ก (Social News and Bookmarking) เป็นเว็บไซต์ที่รวบรวมและเชื่อมโยงบทความหรือเนื้อหาในอินเทอร์เน็ต และจัดกลุ่มเป็นหมวดหมู่ โดยผู้ใช้เป็นผู้ส่งและมีการให้คะแนนจากผู้ใช้อื่น ๆ เพื่อร่วมกันตัดสินใจว่าเนื้อหาใดเป็นที่น่าสนใจมากที่สุด

6) ออนไลน์ฟอรัม (Online Forum) คือ เว็บไซต์ที่ใช้สำหรับการแลกเปลี่ยนความคิดเห็น บทสนทนา การพูดคุย การอภิปรายในสังคมออนไลน์แบบสองทาง ออนไลน์ฟอรัมที่ได้รับความนิยมอย่างมากในสังคมออนไลน์ของไทย คือ พันทิป (Pantip) ที่ให้บริการแก่สมาชิกในการตั้งกระทู้เพื่อพูดคุยและสอบถามความคิดเห็น โดยแบ่งหัวข้อกระทู้ออกเป็นหมวดหมู่ต่าง ๆ 35 หมวดหมู่

งานวิจัยนี้รวบรวมบทวิจารณ์จากทวิตเตอร์ เนื่องจากเป็นบทวิจารณ์ที่มีความยาวไม่มาก เพราะมีการจำกัดขนาดของข้อความ ทำให้ได้ข้อความที่กระชับ ซึ่งช่วยลดเวลาในการประมวลผลคำ และมียอดประกอบทั้งข้อความและภาพ ซึ่งสามารถรวบรวมบทวิจารณ์ทั้งข้อความและภาพมาใช้ในการวิเคราะห์ข้อมูลได้ รวมถึงในข้อความที่ผู้ใช้เผยแพร่ผ่านทวิตเตอร์ และการแบ่งปันต่อ ๆ กัน ทำให้สามารถนำมาวิเคราะห์ข้อมูลที่เกี่ยวข้อง ได้แก่ การแบ่งประเภทบทวิจารณ์ว่าเป็นด้านบวก ลบ หรือเป็นกลาง

## 2.4 การประมวลผลทางภาษา (Word Segmentation)

การประมวลผลทางภาษา เป็นการตัดคำหรือแบ่งคำเพื่อนำคำที่ถูกตัดไปใช้ในวัตถุประสงค์ต่าง ๆ เช่น การวิเคราะห์ไวยากรณ์ (Syntax analysis) การประมวลผลภาษาธรรมชาติ (Natural language processing) การตรวจสอบการสะกดคำ (Spelling check) การทำดัชนีสำหรับเอกสาร (Document indexing) เป็นต้น ซึ่งในงานวิจัยนี้ใช้การประมวลผลทางภาษาเพื่อวัตถุประสงค์ในการประมวลผลภาษาธรรมชาติ สำหรับวิเคราะห์ความรู้สึกจากคำที่ปรากฏในบทวิจารณ์ด้านอาหารในภาษาไทย ซึ่งการประมวลผลทางภาษามีรายละเอียดของการประมวลผลดังนี้ (กานดา รุณนะพงศา และวปโยธร อุราธรรมกุล, 2548)

### 2.4.1 อักษรวิธีของภาษาไทย

ภาษาไทยมีลักษณะที่แตกต่างจากภาษาอื่น เนื่องจากเป็นภาษาที่เขียนติดกันทั้งประโยค และประกอบไปด้วยพยัญชนะ สระ และวรรณยุกต์ ในบางครั้งมีการันต์ประกอบอยู่ในคำด้วย โดยพยัญชนะบางตัวสามารถเป็นสระ และตัวสะกดได้ ซึ่งพยัญชนะถูกแบ่งออกเป็นอักษรสามหมู่ ได้แก่ อักษรสูง อักษรกลาง และอักษรต่ำ ส่วนสระแบ่งออกเป็นสระเดี่ยว และสระผสม อีกทั้งวรรณยุกต์มี 5 เสียง แต่นับเป็น 4 รูป ซึ่งเมื่อผสมกันจะทำให้เกิดเป็นการออกเสียงและความหมาย โดยตัวอย่างการสะกดคำตามอักษรวิธีแสดงดังตารางที่ 2.1

ตารางที่ 2.1 ตัวอย่างการสะกดคำตามอักษรวิธี

| องค์ประกอบ                                    | ตัวอย่างของคำ    |
|-----------------------------------------------|------------------|
| พยัญชนะ + สระ + วรรณยุกต์                     | หมู กา หมี่      |
| พยัญชนะ + สระ + วรรณยุกต์ + ตัวสะกด           | ลิง ขุน แมว      |
| พยัญชนะ + สระ + วรรณยุกต์ + ตัวสะกด + การันต์ | สิงห์ หงส์ สังกั |

## 2.4.2 วิธีการตัดคำ

ในการประมวลผลทางภาษาในภาษาไทย มีวิธีการตัดคำ 3 วิธี ได้แก่ วิธีการใช้กฎ (Rule-based) การใช้พจนานุกรม (Dictionary) และการใช้คลังคำ (Corpus) ซึ่งแต่ละวิธีมีรายละเอียดดังนี้

- 1) การตัดคำโดยใช้กฎ (Rule-based) เป็นการตัดคำโดยการตรวจสอบกฎเกณฑ์ทางอักขระวิธีที่มีการประสมอักษร การเว้นวรรค และย่อหน้า เพื่อใช้กำหนดขอบเขตของคำ
- 2) การตัดคำโดยใช้พจนานุกรม (Dictionary) วิธีนี้เป็นการอ้างอิงด้วยพจนานุกรมโดยใช้สายอักขระมาเปรียบเทียบกับคำที่อยู่ในพจนานุกรม ซึ่งจะต้องมีการจัดเก็บคำไว้ในพจนานุกรม เป็นวิธีที่ใช้เวลานานในการประมวลผล
- 3) การตัดคำโดยใช้คลังคำ (Corpus) เป็นการตัดคำโดยใช้วิธีทางสถิติในการประมวลผลคำ โดยใช้คลังคำที่เก็บไว้เป็นฐานความรู้ ใช้การนับความถี่ซึ่งประกอบด้วยการนับความถี่โดยใช้ความน่าจะเป็น (Probabilistic Word Segmentation) เป็นวิธีที่ใช้แบบจำลองเอ็นแกรม (N-gram) ในการหารูปแบบในการตัดคำ และลำดับของการตัดคำที่เป็นไปได้มากที่สุด โดยวิธีนี้เป็นการใช้คลังข้อมูลที่มีการตัดคำ และกำหนดหมวดคำไว้ก่อนแล้ว และการตัดคำโดยพิจารณาถึงคุณลักษณะของคำ (Feature-based Word Segmentation) เป็นวิธีในการลดข้อผิดพลาดในการตัดคำ ซึ่งใช้ความน่าจะเป็นของการจำคำหมวดคำเป็นแบบจำลองสำหรับตัดคำ

## 2.4.3 เทคนิคการตัดคำ

ในการตัดคำมีเทคนิคที่ใช้ 4 เทคนิค ได้แก่ เทคนิคการเทียบคำที่ยาวที่สุด (Longest word pattern matching) เทคนิคการเทียบคำที่สั้นที่สุด (Shortest word pattern matching) เทคนิคการใช้ความถี่ของคำ (Word usage frequency) และเทคนิคการย้อนรอยกลับ (Back tracking) ซึ่งแต่ละเทคนิคมีรายละเอียดดังนี้

- 1) เทคนิคการเทียบคำที่ยาวที่สุด (Longest word pattern matching) เป็นเทคนิคตรวจสอบสายอักขระจากซ้ายไปขวา แล้วนำไปเปรียบเทียบกับคำในพจนานุกรม หากพบว่ามีมากกว่า 1 พยางค์ จะเลือกพยางค์ที่ยาวที่สุด แล้วตัดคำต่อไปจนจบสายอักขระ
- 2) เทคนิคการเทียบคำที่สั้นที่สุด (Shortest word pattern matching) เป็นเทคนิคในการเลือกคำที่สั้นที่สุดที่พบก่อน แต่หากมีจำนวนที่มากที่สุดความถูกต้องในการตัดคำจะน้อยกว่าเมื่อเทียบกับการใช้เทคนิคการเทียบคำที่ยาวที่สุด
- 3) เทคนิคการใช้ความถี่ของคำ (Word usage frequency) เทคนิคนี้เป็นเทคนิคที่ช่วยแก้ปัญหาเรื่องคำกำกวมในประโยค โดยใช้การวิเคราะห์ความถี่ของคำที่ใช้ในชีวิตประจำวัน โดยเรียงคำในพจนานุกรมตามความถี่ของคำที่พบในการตัดคำ

4) เทคนิคการย้อนรอยกลับ (Back tracking) เทคนิคนี้ใช้ในกรณีที่เมื่อมีการตัดคำแล้วไม่พบคำที่อยู่ในพจนานุกรม หรืออาจพบคำมากกว่ามากกว่า 1 คำ หากเลือกคำที่ยาวที่สุดจะทำให้ไม่สามารถตัดคำได้ จึงต้องใช้เทคนิคการย้อนรอยกลับย้อนกลับไปอีกคำที่ถูกเลือกแล้วตัดคำใหม่

#### 2.4.4 ขั้นตอนวิธีในการตัดคำ

ในงานวิจัยนี้จะกล่าวถึงขั้นตอนวิธีที่ใช้ในการตัดคำสำหรับการจำแนกประเภทบทวิจารณ์ ซึ่งใช้ขั้นตอนวิธีของ Python ที่ได้พัฒนาขั้นตอนวิธีในการตัดคำเรียกว่า PyThaiNLP

PyThaiNLP คือไลบรารีแพ็คเกจของภาษา Python ใช้สำหรับประมวลผลข้อความและการวิเคราะห์ทางภาษา ซึ่งใช้กับภาษาไทยโดยเฉพาะ มีฟังก์ชันการทำงานที่หลากหลาย เช่น อักษรไทย คำไทย การเรียงคำภาษาไทย คำหยุด (Stop Words) ภาษาไทย การตัดคำภาษาไทย การวิเคราะห์ชนิดของคำทางไวยากรณ์ การตรวจสอบตัวสะกด การแก้คำผิด เป็นต้น

ในการตัดคำภาษาไทย ต้องใช้ API สำหรับเรียกใช้งานไลบรารีของ Python โดยเรียกใช้ขั้นตอนวิธี PyThaiNLP ซึ่งในที่นี้เรียกว่าระบบการตัดคำไทย (Engine) โดยในปัจจุบัน PyThaiNLP ได้พัฒนาระบบการตัดคำไทยที่สามารถใช้งานได้ ซึ่งฟังก์ชันการตัดคำของ PyThaiNLP ถูกออกแบบมาให้รองรับการตัดคำและประโยคหลายขั้นตอนวิธี (วรรณพงษ์ ภัททิย์ไพบุณย์, WWW, 2559) ดังนี้

1) icu เป็นระบบการตัดคำระบบเริ่มแรกของ PyThaiNLP แต่มีความแม่นยำในการตัดอยู่ในระดับต่ำ และเป็นคำเริ่มต้น

2) dict เป็นระบบการตัดคำโดยใช้พจนานุกรมจาก thaiword.txt ในคลังคำ (Corpus) ระบบนี้มีความแม่นยำในระดับปานกลาง ซึ่งจะคืนค่าเป็นเท็จ (False) หากข้อความนั้นไม่สามารถตัดคำได้

3) mm เป็นระบบที่ตัดคำโดยใช้ขั้นตอนวิธีการหาคำที่มีการจับคู่ความสอดคล้องกันมากที่สุด (Maximum Matching algorithm) ในการตัดคำภาษาไทย ซึ่งเป็นการตัดคำที่ใช้กับ API ชุดเก่า

4) newmm เป็นระบบการตัดคำที่ใช้ขั้นตอนวิธีเดียวกับ mm แต่มีการพัฒนาโค้ดชุดใหม่ขึ้นมา โดย Korakot Chaovavanich จากเว็บไซต์ <https://www.facebook.com/groups/408004796247683/permalink/431283740586455/> แล้วนำมาพัฒนาต่อ

5) pylexto เป็นระบบที่ใช้ LexTo ในการตัดคำ

6) deepcut เป็นระบบตัดคำที่ใช้ deepcut จากเว็บไซต์ <https://github.com/rkcosmos/deepcut> ในการตัดคำภาษาไทย

## 2.4.5 การคัดแยกคำตามความสำคัญด้วย (Term Frequency-Inverse Document Frequency: TF-IDF)

การคัดแยกคำตามความสำคัญ หรือ TF-IDF เป็นเทคนิคที่พิจารณาองค์ประกอบของคำภายในประโยค และเอกสาร โดยจะไม่นำลำดับของคำภายในเอกสารมาใช้วิเคราะห์ประกอบด้วย โดยเทคนิคนี้มี 2 องค์ประกอบ คือ ความถี่ของคำ (Term Frequency: TF) และความถี่ของเอกสารผกผัน (Inverse document Frequency: IDF) (Bollacker, et al, 1998) ซึ่งมีรายละเอียดดังนี้

### 2.4.5.1 ความถี่ของคำ (Term Frequency: TF)

ความถี่ของคำ (TF) คือ คำที่ปรากฏในเอกสาร ซึ่งหากคำใดที่ปรากฏอยู่ในเอกสารเป็นจำนวนมาก จะมีความเป็นไปได้สูงว่าคำนั้นมีความเกี่ยวข้องกับใจความสำคัญของเอกสารนั้น ๆ มาก (Spurck, 1972) เช่น หากพิจารณาบทความเกี่ยวกับอาหาร อาจพบคำว่า “อร่อย” ปรากฏหลายจุดภายในบทความนั้น ซึ่งอาจปรากฏคำว่า “อร่อย” จำนวน 3 ครั้ง จากคำทั้งหมดในบทความจำนวน 20 ผลของค่าของ TF เป็นค่าที่บอกความถี่ของคำแต่ละคำที่ปรากฏในเอกสาร โดยคำนวณจากสัดส่วนของจำนวนครั้งที่คำนั้น ๆ ปรากฏในเอกสารกับจำนวนคำทั้งหมดในเอกสาร ในที่นี้ค่า TF ของคำว่าอร่อยจะมีค่าเท่ากับ  $3/20$  หรือ 0.15 ซึ่งสูตรในการคำนวณ แสดงดังสมการที่ (2.1)

$$tf(term, document) = \frac{f(term, document)}{\sum_{term' \in document} f(term', document)} \quad (2.1)$$

### 2.4.5.2 ความถี่ของเอกสารผกผัน (Inverse document Frequency: IDF)

ความถี่ของเอกสารผกผัน (IDF) เป็นการคำนวณค่าน้ำหนัก (weight) ความสำคัญของคำแต่ละคำที่พบเป็นจำนวนมากในหลายเอกสาร (Robertson, 2004) หากค่า IDF ต่ำ หมายความว่าคำนั้นไม่มีความเกี่ยวข้องกับใจความสำคัญของเอกสารนั้น ๆ หากค่า IDF สูง หมายความว่า คำนั้นมีความเกี่ยวข้องกับใจความสำคัญของเอกสารนั้น ๆ ซึ่งคำนวณจากความผกผันของอัตราส่วนของจำนวนเอกสารทั้งหมดต่อจำนวนเอกสารที่มีคำ ๆ นั้นปรากฏอยู่ โดยสูตรในการคำนวณหาค่า IDF แสดงดังสมการที่ (2.2)

$$idf(term, allDocument) = \log \frac{N}{df(t)} \quad (2.2)$$



ทั้งนี้ หากต้องการหาค่าที่เป็นใจความสำคัญของเอกสาร สามารถหาได้จากค่า TF-IDF (Aizawa and Akiko, 2003) ซึ่งหากค่า TF-IDF สูง หมายความว่า คำ ๆ นั้น เป็นคำที่มีความสำคัญสูง และมีแนวโน้มที่จะเป็นใจความสำคัญของเอกสาร โดยการคำนวณหาค่า TF-IDF แสดงดังสมการที่ (2.3)

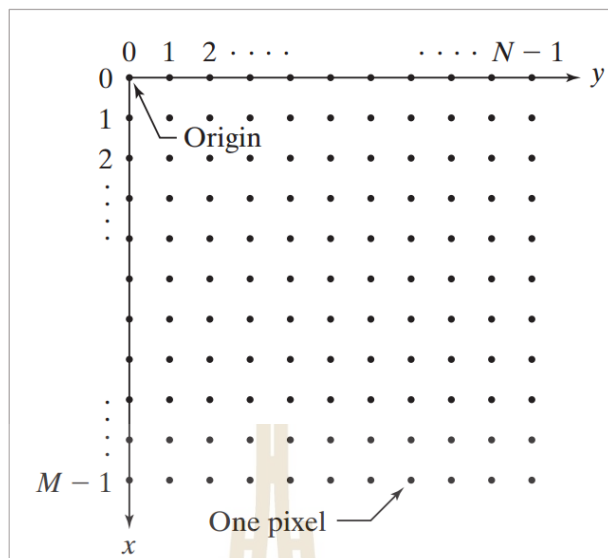
$$TF-IDF = TF \times IDF \quad (2.3)$$

## 2.5 การประมวลผลภาพดิจิทัล (Digital Image Processing)

### 2.5.1 ความหมายของการประมวลผลภาพ

กอนซาเลซ และวู้ด (Gonzalez and Woods, 2002; 2008) ได้ให้ความหมายของการประมวลผลภาพดิจิทัลว่า เป็นการแปลงข้อมูลภาพให้อยู่ในรูปแบบข้อมูลดิจิทัล หรือข้อมูลเชิงตัวเลข โดยวัตถุประสงค์ของการประมวลผลภาพแบ่งเป็นสองประเภท คือการปรับปรุงคุณภาพของภาพเพื่อให้มนุษย์สามารถมองเห็นรายละเอียดได้ชัดเจนมากขึ้น และเพื่อให้คอมพิวเตอร์สามารถแปลความหมายของภาพได้

โดยกอนซาเลซ และคณะ (Gonzalez, et al, 2009) ได้ให้ความหมายของภาพดิจิทัลว่าเป็นภาพที่แสดงเป็นอะเรย์ (array) หลายมิติ หรือเป็นเมทริกซ์ (matrix) เชิงตัวเลข ภาพดิจิทัลจะแสดงในลักษณะสองมิติ มีความกว้างและความสูงของภาพบนแนวแกน X และแกน Y ส่วนจุดใด ๆ ที่อยู่บนระนาบ XY จะเรียกแทนจุดนั้นว่า พิกเซล (pixel) ภาพดิจิทัลที่มี M แถว และ N คอลัมน์ มีจุดกำเนิดคือ  $(x, y) = (0, 0)$  โดยพิกเซลบนภาพดิจิทัล แสดงดังรูปที่ 2.1



ที่มา: Gonzalez, et al (2009)

รูปที่ 2.1 จุดพิกเซลบนภาพดิจิทัล

โดยภาพดิจิทัลขนาด  $M \times N$  ในรูปแบบเมทริกซ์แสดงดังรูปที่ 2.2

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix}$$

ที่มา: Gonzalez, et al (2009)

รูปที่ 2.2 ภาพดิจิทัลขนาด  $M \times N$  ในรูปแบบเมทริกซ์

### 2.5.2 ขั้นตอนการประมวลผลภาพ (Step of Image Processing)

แพรท (Pratt, 2001) ได้จำแนกขั้นตอนในการประมวลผลภาพดิจิทัล โดยมีรายละเอียด ดังนี้

1) การจัดสภาพแวดล้อม (Scene Constraint) ในขั้นตอนนี้มีวัตถุประสงค์เพื่อลดความซับซ้อนในการประมวลผลภาพ เนื่องจากความสามารถในการมองเห็นและรับรู้ของอุปกรณ์ประมวลผลภาพอาจมีข้อจำกัด และยังไม่มีความสามารถเทียบเท่ามนุษย์ โดยมีแนวทางในการจัดสภาพแวดล้อม ดังนี้

- การจัดการกับวัตถุ หากไม่จัดการกับการวางตัวของวัตถุ อุปกรณ์ประมวลผลต้องหาทิศทางของวัตถุเอง
- ระยะระหว่างกล้องหรือเลนส์ถึงวัตถุ ซึ่งจะเป็นตัวกำหนดขนาดของวัตถุที่ระบบมองเห็น
- การจัดการเรื่องแสง โดยแสงเป็นองค์ประกอบที่สำคัญ เนื่องจากการมองเห็นภาพของระบบอัตโนมัติเกิดจากการที่แสงตกกระทบวัตถุ แล้วสะท้อนผ่านเลนส์เข้าตัวเซนเซอร์รับภาพของกล้อง ซึ่งการจัดการเกี่ยวกับแสงจำเป็นต้องพิจารณาเรื่องการเลือกใช้แหล่งกำเนิดแสง

2) การได้มาซึ่งข้อมูลภาพ (Image Acquisition) ในการรับภาพจะมีเซ็นเซอร์รับภาพ (Image Sensor) ประกอบด้วยไดโอดที่มีความไวต่อแสงเป็นจำนวนมาก เมื่อแสงตกกระทบไดโอด เซลล์รับภาพพิกเซลจะจำความเข้มแสงที่ไดโอดอ่านได้ โดยหนึ่งเซลล์จะให้ค่าความเข้มแสงที่เพียงค่าเดียว ค่าที่ได้จากเซลล์รับภาพจะมีค่าระหว่าง 0-255 หากค่าที่ได้มีค่าเท่ากับ 0 หมายถึงเซลล์รับภาพมีความเข้มแสงต่ำสุดหรือเป็นด้านมืด และหากมีค่าเท่ากับ 255 หมายถึงมีความเข้มแสงสูงสุดหรือเป็นด้านสว่าง

3) การประมวลผลภาพเบื้องต้น (Pre-Processing) มีหลายกระบวนการไม่ว่าจะเป็นการลดสัญญาณรบกวนที่ปรากฏขึ้นในภาพ การตรวจจับขอบของวัตถุที่อยู่ในภาพ การแปลงคุณสมบัติทางกายภาพ เช่น การหมุน การเลื่อน การย่อ และการขยาย เป็นต้น การแปลงสี การวิเคราะห์ภาพในเชิงความถี่ การบีบอัดภาพ และอื่น ๆ

4) การแยกบริเวณ (Segmentation) เป็นขั้นตอนการแยกบริเวณภาพที่มีลักษณะร่วมกันออกเป็นส่วน ๆ โดยเน้นไปที่การแยกวัตถุออกจากพื้นหลัง สำหรับการแยกบริเวณมี 2 วิธี ได้แก่ การแยกบริเวณโดยใช้ค่าเทรชโฮลด์ (Threshold) และการแยกบริเวณโดยใช้ขอบของวัตถุ (Edge Based Segmentation)

- การแยกบริเวณโดยใช้ค่าเทรชโฮลด์ (Threshold) โดยการแปลงภาพระดับเทาเป็นภาพไบนารี (Binary image) โดยค่าเทรชโฮลด์มีค่าระหว่าง 0-255 ถ้าค่าความเข้มแสงที่พิกเซลตำแหน่งใดมีค่าต่ำกว่าหรือเท่ากับค่าเทรชโฮลด์ จะกำหนดให้ค่าพิกเซลในตำแหน่งนั้นมีค่าเป็น 0 หรือเป็นด้านมืด และถ้าพิกเซลใดมีค่าสูงกว่าค่าเทรชโฮลด์ ให้พิกเซลนั้นมีค่าเป็น 255 หรือเปลี่ยนเป็นด้านสว่าง

- การแยกบริเวณโดยใช้ขอบของวัตถุ (Edge Based Segmentation) การใช้วิธีนี้ต้องคำนวณหาขอบภาพของวัตถุ ซึ่งขอบในความหมายของการประมวลผลภาพคือพิกเซลที่มีการเปลี่ยนแปลงความเข้มแสงเกินค่าที่กำหนด

5) **การหาคุณลักษณะของวัตถุ (Feature Extraction)** เป็นการหาคุณลักษณะต่าง ๆ ของบริเวณหรือของวัตถุแต่ละชิ้นที่อยู่ในภาพ ผลที่ได้ของขั้นตอนนี้ คือค่าคุณลักษณะต่าง ๆ ของวัตถุแต่ละชิ้นที่อยู่ในภาพ

6) **การจำแนกวัตถุและการแปลความหมาย (Classification and Interpretation)** เป็นขั้นตอนการจัดกลุ่มให้วัตถุว่าเป็นวัตถุที่อยู่ในกลุ่มใด โดยอาศัยข้อมูลที่ได้จากการการคำนวณ ซึ่งเป็นคุณสมบัติของวัตถุนั้น ๆ เปรียบเทียบกับวัตถุตัวอย่างที่อยู่ในแต่ละกลุ่ม ก่อนที่ระบบจะสามารถตัดสินใจได้ ระบบจะต้องมีตัวอย่างของวัตถุในแต่ละกลุ่ม เพื่อใช้ในการเปรียบเทียบและสามารถแปลความหมายที่ถูกต้อง

### 2.5.3 องค์ประกอบในการประมวลผลภาพ (Components of Image Processing)

กอนซาเลซ และวูด (Gonzalez and Woods, 2008) ได้แบ่งองค์ประกอบที่จำเป็นในการประมวลผลภาพดิจิทัลไว้ทั้งสิ้น องค์ประกอบ ดังนี้

1) **ฮาร์ดแวร์การประมวลผลภาพเฉพาะ (Specialized Image Processing Hardware)** ประกอบด้วยฮาร์ดแวร์ที่ใช้งานการดำเนินงานแบบดั้งเดิมทำงานร่วมกับเครื่องมือในการแปลงค่าตัวเลข (Digitizer) มีหน้าที่ในการรับส่งข้อมูลที่ใช้ความเร็วเป็นลักษณะเฉพาะ

2) **คอมพิวเตอร์ (Computer)** ในระบบประมวลผลภาพหมายถึง เครื่องคอมพิวเตอร์ทั่วไปที่สามารถใช้งานได้ไปจนถึงซูเปอร์คอมพิวเตอร์ ในการใช้งานเฉพาะบางครั้งคอมพิวเตอร์ที่กำหนดเองจะใช้เพื่อให้บรรลุระดับของประสิทธิภาพที่ต้องการ ซึ่งระบบแทบทุกระบบมีอุปกรณ์ครบถ้วนเหมาะสำหรับการประมวลผลภาพแบบออฟไลน์

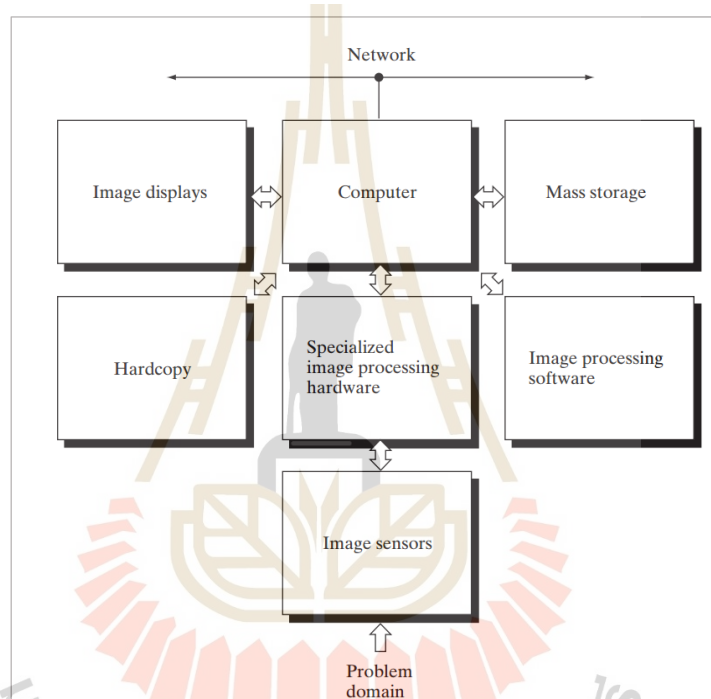
3) **โปรแกรมสำหรับการประมวลผลภาพ (Image Processing Software)** ประกอบด้วยโมดูลเฉพาะที่ดำเนินการเฉพาะ นอกจากนี้ยังมีความสามารถสำหรับผู้ใช้ในการเขียนโปรแกรมที่น้อยที่สุดโดยใช้โมดูลเฉพาะ โปรแกรมที่ซับซ้อนมากขึ้นจะช่วยให้สามารถรวมโมดูลต่าง ๆ เข้าด้วยกัน และใช้คำสั่งในการเขียนโปรแกรมทั่วไปจากภาษาคอมพิวเตอร์อย่างน้อยหนึ่งภาษา

4) **พื้นที่จัดเก็บ (Mass storage)** เป็นสิ่งจำเป็นในการประมวลผลภาพ ภาพขนาดพิกเซลซึ่งความเข้มของแต่ละพิกเซลเป็น 8 บิต หากภาพไม่ถูกบีบอัด ต้องมีพื้นที่จัดเก็บหนึ่งเมกะไบต์ การประมวลผลภาพแบ่งออกเป็น 3 ประเภท คือ (1) การจัดเก็บข้อมูลระยะสั้นในระหว่างการประมวลผล (2) การจัดเก็บออนไลน์เพื่อการเรียกคืนที่รวดเร็วและ (3) การจัดเก็บข้อมูลที่โดยไม่จำกัดการเข้าถึง

5) **จอภาพ (Image Displays)** ภาพที่ใช้ในปัจจุบันส่วนใหญ่เป็นสี จอภาพถูกขับเคลื่อนด้วยผลลัพธ์ของการแสดงภาพและกราฟิก การ์ดที่เป็นส่วนสำคัญของระบบคอมพิวเตอร์ มักไม่มีข้อกำหนดสำหรับแอปพลิเคชันการแสดงภาพที่การ์ดแสดงผล

6) **สิ่งพิมพ์ออก (Hardcopy)** ใช้สำหรับการบันทึกภาพ รวมถึงเครื่องพิมพ์เลเซอร์ กล้องฟิล์ม อุปกรณ์ที่ไวต่ออุณหภูมิหน่วยอิงค์เจ็ทและหน่วยดิจิทัล เช่น ดิสก์ และ CDROM ฟิล์ม ควรมีความละเอียดสูง อุปกรณ์ที่ใช้ควรเป็นอุปกรณ์ที่ได้รับการยอมรับและเป็นมาตรฐานสำหรับการนำเสนอภาพ

7) **เครือข่าย (Networking)** เป็นฟังก์ชันเริ่มต้นในระบบคอมพิวเตอร์ในปัจจุบัน เนื่องจากข้อมูลจำนวนมากที่มีอยู่ในแอปพลิเคชันการประมวลผลภาพ การพิจารณาที่สำคัญในการรับส่งข้อมูลคือแบนด์วิดท์ (Bandwidth)



ที่มา: Gonzalez and Woods (2008)

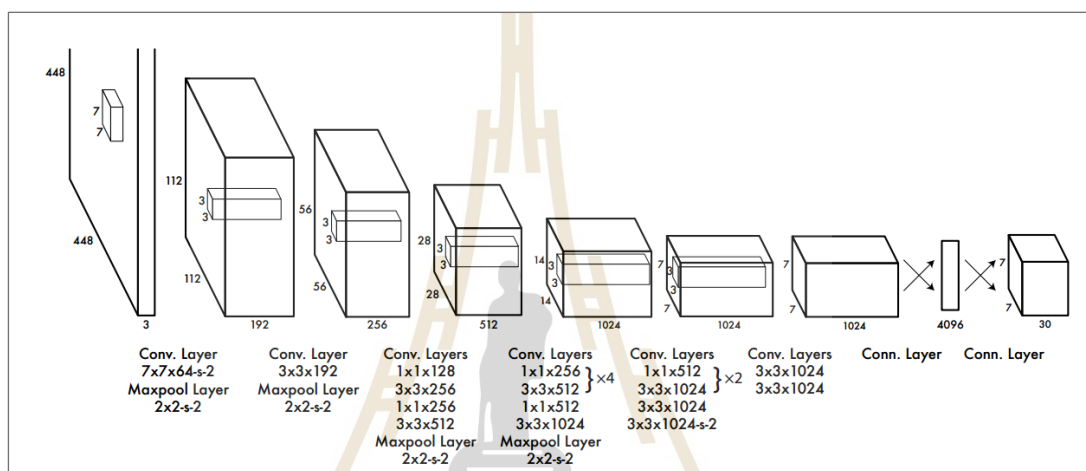
### รูปที่ 2.3 องค์ประกอบในการประมวลผลภาพดิจิทัล

จากการทบทวนวรรณกรรมเกี่ยวกับการประมวลผลภาพดิจิทัล พบว่า ขั้นตอนในการประมวลผลภาพมีหลายขั้นตอน และมีความละเอียดซับซ้อน โดยงานวิจัยนี้ได้นำขั้นตอนการประมวลผลภาพมาใช้ในการสกัดองค์ประกอบภายในภาพจากการรวบรวมภาพที่มาจากสื่อสังคมออนไลน์ เพื่อให้สามารถจำแนกคุณลักษณะที่แตกต่างกันของภาพได้ ซึ่งต้องเรียนรู้ขั้นตอนการประมวลผลภาพ ร่วมกับเทคนิคการเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก เพื่อลดข้อผิดพลาดในการประมวลผลภาพ

## 2.6 โยโล (You Only Look Once: YOLO)

### 2.6.1 คุณลักษณะของ YOLO

โยโล (YOLO) คือเทคโนโลยีการตรวจจับวัตถุแบบเรียลไทม์ ซึ่งเป็นเทคนิคที่ใช้ในการตรวจจับวัตถุจากกล้องวิดีโอ เป็นขั้นตอนวิธีการเรียนรู้เชิงลึก ที่ปรับปรุงประสิทธิภาพการทำงานมาจากโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network) ที่มีประสิทธิภาพการทำงานในการประมวลผลภาพที่รวดเร็วมากที่สุดในปัจจุบัน (Redmon, et al, 2016) โดยสถาปัตยกรรมของ YOLO แสดงดังรูปที่ 2.4



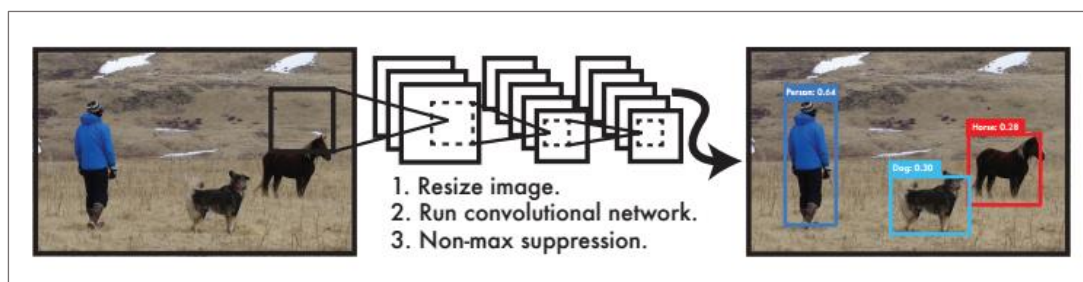
ที่มา: (Redmon, et al, 2016)

รูปที่ 2.4 สถาปัตยกรรมของ YOLO

จากสถาปัตยกรรมของ YOLO ประกอบด้วยชั้นเริ่มต้นที่ทำหน้าที่ในการแยกคุณสมบัติของภาพโดยเชื่อมโยงกับโครงข่ายแบบสังวัตนาการสำหรับสร้างแบบจำลองในการทำนายความน่าจะเป็นของวัตถุ (Object) โดยสถาปัตยกรรมของ YOLO ได้มีการปรับปรุงมาจาก GoogLeNet ซึ่งเป็นแบบจำลองในการจำแนกภาพ ซึ่ง YOLO แบ่งโครงข่ายในการจำแนกภาพออกเป็น 24 ชั้น (Layer) โดยมีการเชื่อมต่อ 2 ชั้นเข้าด้วยกันอย่างสมบูรณ์ และลดขนาดบิตของภาพเป็น 1x1 และ 3x3 ตามลำดับ โดยโครงสร้างของ YOLO มีการทำงานแบบ Pass Through Image โดยในหนึ่งรอบจะมีการวนรอบด้วยโครงข่ายแบบสังวัตนาการ (CNN) และทำนายหาตำแหน่งของวัตถุที่สนใจภายในภาพและประเภทของวัตถุจากคุณลักษณะของภาพ

ในเริ่มแรกกระบวนการตรวจจับภาพของ YOLO มีความเรียบง่ายและรวดเร็ว ประกอบด้วย (1) ข้อมูลนำเข้า ซึ่งเป็นภาพขนาด 448x448 (2) กระบวนการตรวจจับภาพโดยโครงข่ายแบบสังวัตนาการ และ (3) การตรวจสอบความเชื่อมั่นของแบบจำลอง ซึ่งองค์ประกอบในการทำงานของ YOLO แสดงดังรูปที่ 2.5

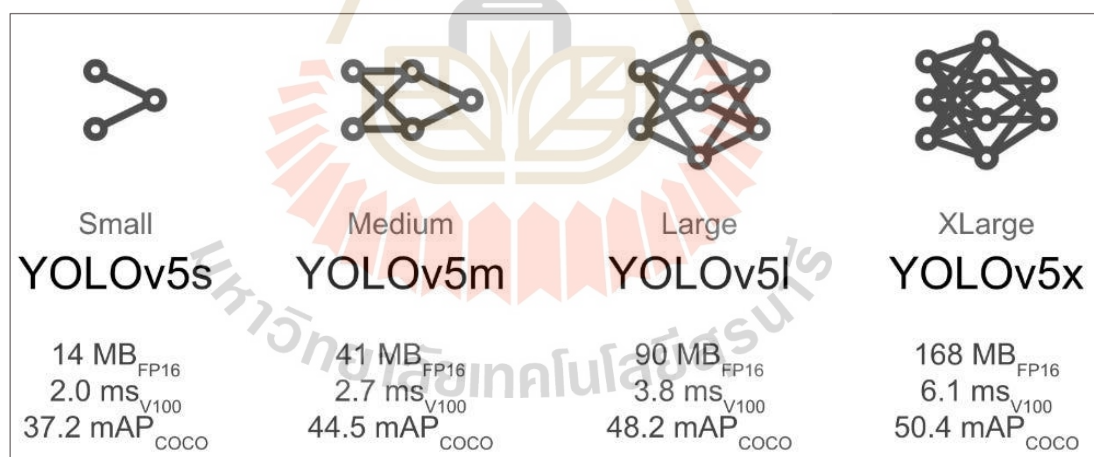




ที่มา: (Redmon, et al, 2016)

### รูปที่ 2.5 องค์ประกอบในการทำงานของ YOLO

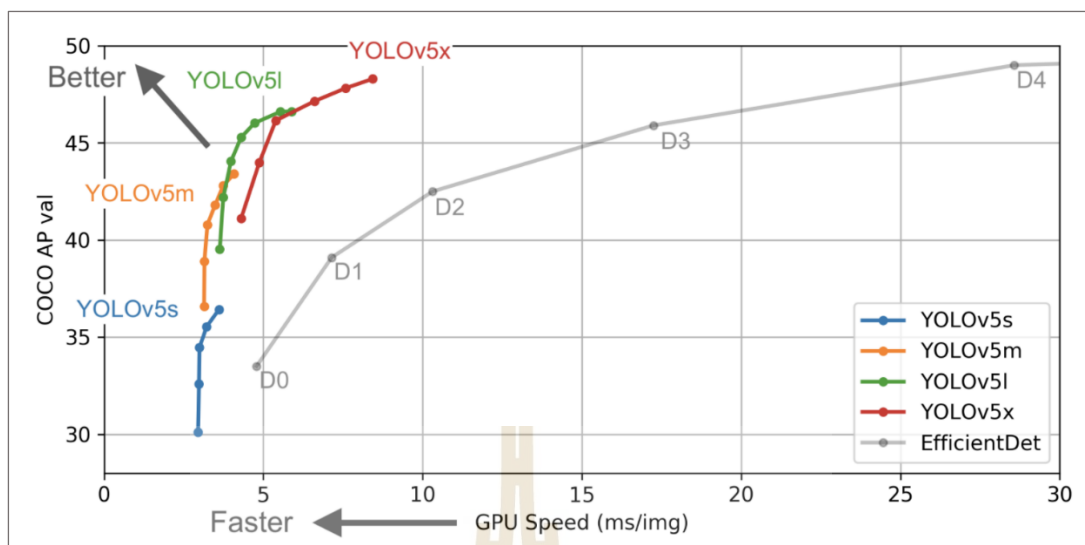
ในปัจจุบันได้มีการพัฒนา YOLO มาหลายเวอร์ชัน โดยเวอร์ชันปัจจุบันคือ YOLOv5 ซึ่งเป็นเวอร์ชันที่งานวิจัยนี้นำมาใช้ในการจำแนกประเภทของอาหารบนภาพ ซึ่ง YOLOv5 ถูกพัฒนามาจาก YOLOv4 ที่มีการปรับปรุงประสิทธิภาพในการทำงานให้ดีขึ้น แต่ยังคงมีความรวดเร็วในการทำงานเช่นเดียวกับเวอร์ชันที่มีมาก่อนหน้านี้ โดย YOLOv5 แบ่งขั้นตอนวิธีในการทำนายผลและสร้างแบบจำลองออกเป็นแบบจำลองขนาดเล็ก ซึ่งใช้รหัสเป็นตัวอักษร s และแบบจำลองขนาดใหญ่ที่ใช้รหัสเป็นตัวอักษร x โดยขั้นตอนวิธีของ YOLOv5 แสดงดังรูปที่ 2.6



ที่มา: (Redmon, et al, 2016)

### รูปที่ 2.6 ขั้นตอนวิธีของ YOLOv5

ทั้งนี้ ประสิทธิภาพการทำงานของ YOLOv5 เมื่อเปรียบเทียบกับกันระหว่างแบบจำลองขนาดเล็ก (YOLOv5s) จนถึงแบบจำลองขนาดใหญ่ (YOLOv5x) แสดงดังรูปที่ 2.7

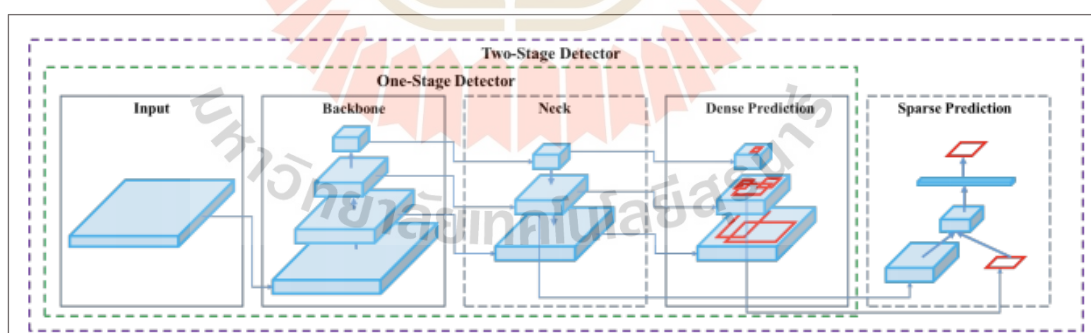


ที่มา: Redmon, et al, 2016

รูปที่ 2.7 ประสิทธิภาพการทำงานของ YOLOv5

## 2.6.2 โครงสร้างการทำงานของ YOLOv5

การทำงานของ YOLOv5 เริ่มต้นด้วยการนำเข้าข้อมูลนำเข้าเข้าสู่แบบจำลองของ YOLOv5 ซึ่งจะมีการแบ่งเซลล์บนภาพเป็นแบบตาราง (Grid cell) ออกเป็นช่องเล็ก ๆ ขนาด  $n \times n$  (Bochkovski, et al, 2020) โดยโครงสร้างการทำงานของ YOLOv5 แสดงดังรูปที่ 2.8



ที่มา: Bochkovski and et al, 2020

รูปที่ 2.8 โครงสร้างการทำงานของ YOLOv5

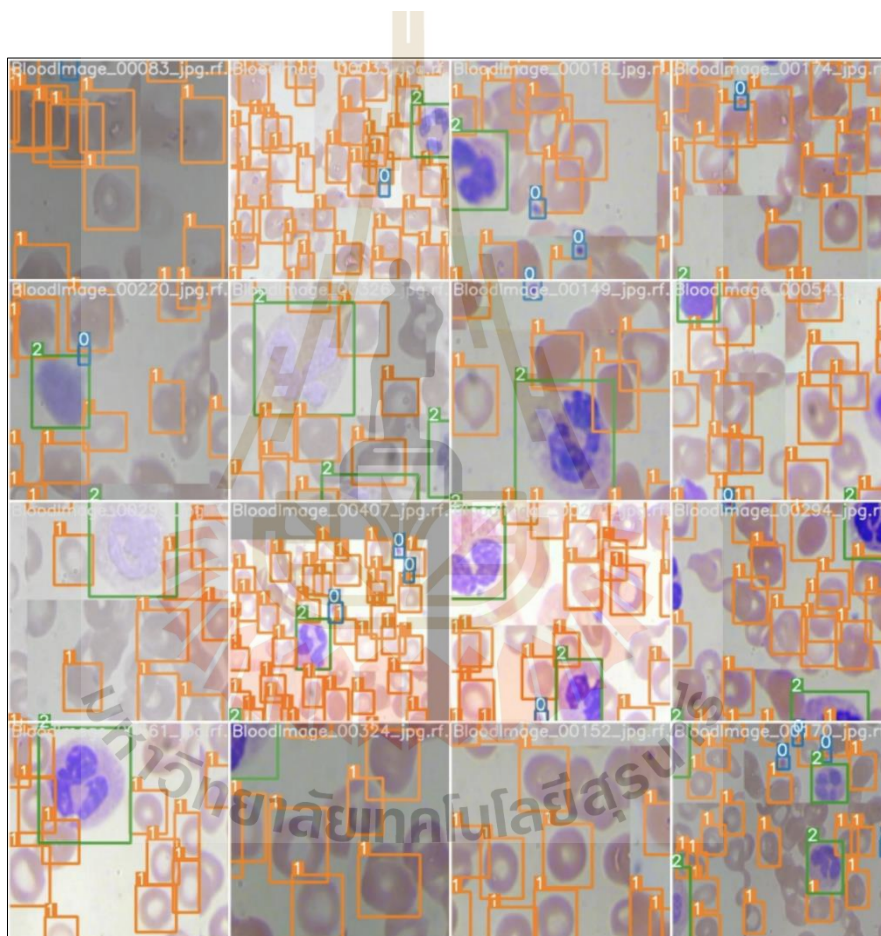
จากโครงสร้างการทำงานของ YOLOv5 ประกอบด้วยส่วนประกอบหลัก 3 ส่วน ดังนี้

1. Backbone คือส่วนที่ทำหน้าที่รับข้อมูลนำเข้าเข้าสู่แบบจำลอง แล้วสกัดคุณลักษณะที่โดดเด่นของภาพออกมา โดยจะส่งต่อไปยัง Neck ซึ่งเป็นส่วนที่รับคุณลักษณะของภาพ

2. Neck คือส่วนที่หน้าที่รับคุณลักษณะของภาพที่ได้จากการสกัดจากส่วน Backbone มาสร้างแผนที่คุณลักษณะ (Feature map) ของภาพ

3. Head คือส่วนที่ทำหน้าที่ตีกรอบตำแหน่งของวัตถุนภาพที่สนใจ แล้วทำนาย คลาส (Class) ของวัตถุนั้น ๆ และส่งต่อไปยังชั้นผลลัพธ์ (Output Layer) เพื่อแสดงผลลัพธ์จากการทำนาย

โดยตัวอย่างภาพที่มีการสร้างแบบจำลองสำหรับการตรวจจับวัตถุนภาพด้วย YOLOv5 แสดงดังรูปที่ 2.9



รูปที่ 2.9 ตัวอย่างการตรวจจับวัตถุนภาพด้วย YOLOv5

## 2.7 การเรียนรู้ของเครื่อง (Machine Learning)

### 2.7.1 ความหมายของการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง เป็นสาขาหนึ่งของปัญญาประดิษฐ์ที่พัฒนามาจากการศึกษา การรู้จำที่เกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยแบบจำลองที่สร้างมาจากชุดข้อมูลนำเข้าเพื่อการทำนายหรือ ตัดสินใจในภายหลัง ซึ่งจะไมทำงานตามลำดับของคำสั่ง โปรแกรมคอมพิวเตอร์ (Han and Kamber, 2006)

### 2.7.2 เทคนิคการเรียนรู้ของเครื่อง

ในการจำแนกการเทคนิคการเรียนรู้ของเครื่องนั้น สามารถจำแนกออกเป็นหมวดหมู่ ดังนี้ (Ian and Eibe, 2005)

1) การจำแนกประเภท (Classification) เป็นการทำนายประเภทของข้อมูล จากการหารูปแบบของชุดข้อมูลที่มีความใกล้เคียงกันหรือเหมือนกันมาก โดยชุดข้อมูลที่ถูกทำนายเกิดจากการเรียนรู้จากชุดข้อมูลที่มีอยู่แล้ว แบบจำลองที่เกิดจากการทำนาย มีดังนี้

#### 1.1) ต้นไม้ตัดสินใจ (Decision Tree)

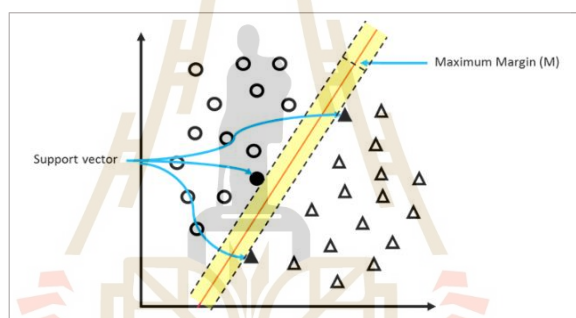
ต้นไม้ตัดสินใจที่นิยมใช้ในปัจจุบันได้แก่ ID3 โดยการสร้างต้นไม้การตัดสินใจจากบนลงล่างด้วยการระบุลักษณะที่ ควรเป็นรากของต้นไม้ (Root node) และทำแบบเดิมไปเรื่อย ๆ เพื่อระบุลักษณะอื่นให้ได้เป็นต้นไม้ทั้งต้นที่มีความสมบูรณ์ โดยในการเลือกว่าลักษณะใดดีที่สุดนั้นดูจากค่าของลักษณะ เรียกว่า เกนความรู้ (Information gain) (Mitchell, 1997) โดยค่านี้คำนวณได้จากสมการที่ (2.4)

$$Gain(S,A) = E(S) - \sum_{v \in \text{value}(A)} \frac{|S_v|}{|S|} E(S_v) \quad (2.4)$$

|           |                                                      |
|-----------|------------------------------------------------------|
| โดยที่ S  | คือ ตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปรตาม |
| E         | คือ เอนโทรปีของตัวอย่าง                              |
| A         | คือ ตัวแปรต้นที่พิจารณา                              |
| value (A) | คือ เซตของค่าของ A ที่เป็นไปได้                      |
| $S_v$     | คือ ตัวอย่างที่ A มีค่า v ทั้งหมด                    |

## 1.2) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

ซัพพอร์ตเวกเตอร์แมชชีนเป็นหนึ่งในวิธีการเรียนรู้แบบมีผู้สอน ใช้เพื่อการแบ่งประเภทข้อมูลและการวิเคราะห์การถดถอย โดยจำแนกเชิงเส้น (Linear Classifier) แบบ 2 คลาส เมื่อมีข้อมูลฝึกมาให้และแต่ละข้อมูลถูกจัดอยู่ในประเภทใดประเภทหนึ่งจากสองประเภท โดยจะสร้างแบบจำลองที่สามารถพยากรณ์ได้ว่าตัวอย่างใหม่นี้จะตกอยู่ในกลุ่มใด ข้อดีของขั้นตอนวิธีนี้คือมีประสิทธิภาพในการจำแนกข้อมูลที่มีมิติจำนวนมากได้ นอกจากนี้ การใช้ฟังก์ชันเคอร์เนล (Kernel Function) เพื่อแปลงข้อมูลไปยังมิติที่สูงขึ้นในปริภูมิคุณลักษณะ (Feature Space) สามารถจำแนกข้อมูลที่มีความคลุมเครือได้อย่างมีประสิทธิภาพ หลักการของ SVM คือการหาเส้นตรงที่มีมาร์จินที่โตที่สุด (Maximum Margin) ที่สามารถแบ่งข้อมูลออกเป็น 2 คลาส ได้แก่ + และคลาส - เป็นข้อมูลขนาด 2 มิติ โดยเส้นตรงที่ใช้แบ่งข้อมูลมีมาร์จินเท่ากับ  $M = 2w$  ซึ่ง เป็นความกว้างระหว่างเส้นตรงกับซัพพอร์ตเวกเตอร์ (Support vector) ดังรูปที่ 2.10 (Mitchell, 1997)



ที่มา: ศาสตรา วังศันวสุ (2560)

รูปที่ 2.10 ความกว้างระหว่างเส้นตรงกับซัพพอร์ตเวกเตอร์

คู่ของเส้นประที่กว้างที่สุดเท่าที่จะเป็นไปได้ มีสองแบบ Hard margin classification คือคู่เส้นประที่ห้ามไม่ให้มีจุดข้อมูลอยู่ในพื้นที่ระหว่างเส้นประ Soft margin classification คืออนุญาตให้มีข้อมูลอยู่ในพื้นที่ระหว่างเส้นประได้บ้าง ซึ่ง SVM ถูกออกแบบมาสำหรับ Binary classification แต่สามารถนำไปประยุกต์ใช้กับ Multiclass classification และ Linear regression ได้ โดย SVM ใช้ฟังก์ชันสมมติฐาน (Hypothesis function) แบบเส้นตรง เหมือนกับ Linear regression (Cortes and et al., 1995) ซึ่งฟังก์ชันสมมติฐานแสดงดังสมการที่ (2.5)

$$h_0(x) = w_1x_1 + w_2x_2 + \dots + w_nx_n + b = w^T x + b \quad (2.5)$$

หากผลลัพธ์ที่ได้เป็นบวก จะทำนาย Class  $\hat{y}$  ว่าเป็น หากเป็นลบ ทำนายว่าเป็น 0 โดยวิธีการตัดสินใจตามเงื่อนไขแสดงดังสมการที่ (2.6)



$$\hat{y} = \begin{cases} 0 & \text{if } w^T x + b < 0 \\ 1 & \text{if } w^T x + b \geq 0 \end{cases} \quad (2.6)$$

เมื่อนิยามเส้นแบ่งการตัดสินใจแล้ว (เส้นทึบ) ต้องกำหนดเส้นประทั้งสองด้านของเส้นทึบ โดยเส้นประแต่ละด้านคือตำแหน่งที่  $h_0(x)$  เท่ากับ -1 และ 1 ความชันของฟังก์ชันการตัดสินใจ เท่ากับ Norm ของ Vector ที่ค่าน้ำหนัก  $w$  ดังสมการที่ (2.7)

$$\frac{\partial}{\partial x} h_0(x) = \sum_{i=1}^m |w_i| \quad (2.7)$$

ดังนั้น จะได้เป้าหมายที่ต้องลด  $\|w\|$  เพื่อให้ได้ขอบเขตเส้นแบ่งที่กว้างที่สุดเท่าที่จะเป็นไปได้ อย่างไรก็ตาม หากไม่ต้องการให้ขอบเขตเส้นแบ่งนั้นกว้างเกินไปจนกระทั่งครอบคลุมจุดข้อมูล ต้องทำให้ฟังก์ชันการตัดสินใจนั้นมีค่ามากกว่า 1 ซึ่งเป้าหมายของการกำหนดค่าแสดงดังสมการที่ (2.8)

$$\begin{aligned} \text{minimise}_{w,b} &= \frac{1}{2} w^T w \\ \text{subject to } t^{(i)}(w^T x^i + b) &\geq 1 \end{aligned} \quad (2.8)$$

หากเป็น Soft margin ที่อนุญาตให้พื้นที่เส้นขอบเขตการตัดสินใจนั้นกินบริเวณที่มีจุดข้อมูลอยู่ด้วยได้ ต้องเพิ่มตัวแปรที่เรียกว่า Slack โดยระดับของ Slack จะถูกกำหนดโดย Hyperparameter  $C$  เป้าหมายของการกำหนดค่าแสดงดังสมการที่ (2.9)

$$\begin{aligned} \text{minimise}_{w,b} &= \frac{1}{2} w^T w + C \sum_{i=1}^m \zeta^{(i)} \\ \text{subject to } t^{(i)}(w^T x^i + b) &\geq 1 - \zeta^{(i)} \end{aligned} \quad (2.9)$$

อย่างไรก็ตาม SVM มีข้อจำกัดคือสามารถสร้างเส้นแบ่งขอบเขตการตัดสินใจแบบเส้นตรงเท่านั้น ซึ่งอาจทำงานได้ไม่ดีหากความสัมพันธ์ของข้อมูลมีความซับซ้อน และทำให้ไม่สามารถแบ่งด้วยเส้นตรงไม่ได้ จึงใช้วิธีแก้ปัญหาที่เรียกว่าฟังก์ชันเคอร์เนลเคอร์เนล (Kernel Function)



Kernel คือวิธีการคณิตศาสตร์ที่ทำให้ขั้นตอนวิธี SVM สามารถหาค่าตัวแปรแบบพหุนาม (Polynomial) ได้ โดยไม่ต้องไปเปลี่ยนรูปแบบและความสัมพันธ์ของคุณลักษณะตั้งต้น ซึ่งต้องคิดแปลงเป้าหมายการกำหนดค่าให้อยู่ในรูปแบบที่ Kernel ทำงานได้ โดยปรับเปลี่ยนจากรูปแบบในสมการที่ (2.9) ให้เป็นดังสมการที่ (2.10)

$$\begin{aligned} \text{minimise}_{w,b} = & \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \alpha^{(i)} \alpha^{(j)} t^{(i)} t^{(j)} x^{(i)T} x^{(j)} - \sum_{i=1}^m \alpha^{(i)} \\ & \text{subject to } \alpha^{(i)} \geq 0 \end{aligned} \quad (2.10)$$

### 1.3) นาอ์ฟเบย์ (Naive Bayes)

นาอ์ฟเบย์หรือเบย์อย่างง่าย เป็นขั้นตอนวิธีสำหรับการจำแนกที่ใช้หลักการของความน่าจะเป็น (Probability) ในการทำนาย ซึ่งมีพื้นฐานมาจากทฤษฎีของเบย์ (Bayes theorem) ซึ่งเป็นทฤษฎีที่คำนวณหาความน่าจะเป็นในโอกาสของการเกิดเหตุการณ์ต่าง ๆ โดยมีเงื่อนไขในการคำนวณที่เรียกว่า การคำนวณความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) (Dietrich, et al., 2015) ซึ่งสูตรในการคำนวณแสดงดังสมการที่ (2.11)

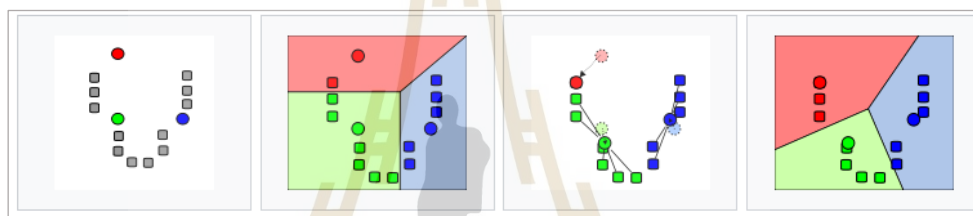
$$P(h|D) = \frac{P(D|h) \times P(h)}{P(D)} \quad (2.11)$$

โดยที่  $P(h)$  คือ ค่าความน่าจะเป็นของสมมติฐานที่คลาสเป็น  $h$   
 $P(D)$  คือ ค่าความน่าจะเป็นของสมมติฐานที่ข้อมูลเป็น  $D$   
 $P(h|D)$  คือ ค่าความน่าจะเป็นของสมมติฐานที่ข้อมูลเป็น  $D$  จะมีคลาสเป็น  $h$   
 $P(D|h)$  คือ ค่าความน่าจะเป็นของสมมติฐานที่คลาส  $h$  จะมีข้อมูลเป็น  $D$

ในกรณีที่มีตัวแปรต้นหรือข้อมูลที่ต้องพิจารณามากกว่า 1 ค่า ความน่าจะเป็นสามารถคำนวณได้จากผลคูณของความน่าจะเป็นของแต่ละข้อมูลที่มีคลาส  $h$  ดังสมการที่ (2.12)

$$P(h|D) = P(D_1|h) \times P(D_2|h) \times \dots \times P(D_n|h) \times P(h) \quad (2.12)$$

2) **การจัดกลุ่ม (Clustering)** เป็นการวิเคราะห์เพื่อจัดกลุ่มข้อมูลที่มีความคล้ายคลึงกัน โดยข้อมูลที่ถูกจัดกลุ่มภายในกลุ่มเดียวกันจะมีความคล้ายคลึงกันมาก ส่วนข้อมูลที่อยู่ต่างกลุ่มจะมีความแตกต่างกันมาก ตัวอย่างเทคนิคการจัดกลุ่ม เช่น ขั้นตอนวิธีค่าเฉลี่ย K (K-mean algorithm) โดยในการจัดกลุ่มด้วยวิธีค่าเฉลี่ย K มีขั้นตอนจากการกำหนดค่าแต่ละข้อมูลการสังเกตไปยังคลัสเตอร์ โดยเลือกคลัสเตอร์ที่จะทำให้ค่าผลบวกกำลังสองภายในคลัสเตอร์นั้น ๆ มีค่าน้อยที่สุด ซึ่งจะได้ผลที่เป็นค่าเฉลี่ยที่ใกล้ที่สุด และปรับค่าของจุดกึ่งกลาง โดยคำนวณหาค่าเฉลี่ยค่าใหม่เพื่อเป็นจุดกึ่งกลาง (Centroid) ของข้อมูลการสังเกตในแต่ละคลัสเตอร์ที่สร้างขึ้นใหม่ โดยจะทำให้ได้ค่าเฉลี่ยใหม่ที่มีค่าน้อยกว่าเดิม เพื่อให้ได้จุดของคลัสเตอร์ที่ใกล้กันมากที่สุด โดยขั้นตอนวิธีค่าเฉลี่ย K แสดงดังรูปที่ 2.11



ที่มา: MacQueen (1967)

รูปที่ 2.11 ขั้นตอนวิธีค่าเฉลี่ย K

3) **กฎความสัมพันธ์ (Association Rules)** เป็นการวิเคราะห์ความสัมพันธ์ที่มีผลลัพธ์แสดงในรูปแบบของกฎ สำหรับแสดงความสัมพันธ์ของค่าความสัมพันธ์และมีเงื่อนไขตรงกับข้อกำหนดและลักษณะของข้อมูลที่ทำนาย ในการวิเคราะห์ความสัมพันธ์ของกฎต้องมีการกำหนดค่าสนับสนุน (Support) และค่าความเชื่อมั่น (Confidence) เพื่อกำหนดระดับความสัมพันธ์และความเกี่ยวข้องกันของกฎ เช่น ขั้นตอนวิธีแบบเอพริออริ (The Apriori Algorithm) การวิเคราะห์แบบตะกร้าสินค้า (Market Basket Analysis) กฎความสัมพันธ์แบบหลายระดับ (Multilevel Association Rules)

จากเทคนิคการเรียนรู้ของเครื่องดังที่กล่าวมาข้างต้น ทำให้ทราบว่า การเลือกใช้งานเทคนิคต่าง ๆ จำเป็นต้องพิจารณาถึงข้อมูล และวัตถุประสงค์ในการใช้งาน ซึ่งหากใช้งานได้ไม่ตรงกับวัตถุประสงค์ หรือไม่สอดคล้องกับข้อมูล อาจส่งผลให้ผลลัพธ์ที่ได้ไม่มีประสิทธิภาพ ซึ่งงานวิจัยนี้เป็นการแนะนำร้านอาหารแก่ผู้ใช้ โดยใช้ข้อมูลบทวิจารณ์เกี่ยวกับอาหารบนสื่อสังคมออนไลน์ ซึ่งก่อนการแนะนำร้านอาหารแก่ผู้ใช้ จำเป็นต้องมีการจำแนกอาหารและร้านอาหารก่อน จึงจะสามารถแนะนำร้านอาหารแก่ผู้ใช้ได้อย่างถูกต้อง งานวิจัยนี้จึงใช้เทคนิคการเรียนรู้ของเครื่องแบบการจำแนกประเภท ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน

## 2.8 การเรียนรู้เชิงลึก (Deep Learning)

### 2.8.1 ความหมายของการเรียนรู้เชิงลึก (Deep Learning)

เต็ง และยู (Deng and Yu, 2014) ได้ให้แนวคิดเกี่ยวกับการเรียนรู้เชิงลึกว่า เป็นสาขาของการเรียนรู้ของเครื่อง พื้นฐานของการเรียนรู้เชิงลึกคือ ขั้นตอนวิธีที่พยายามจะสร้างแบบจำลองเพื่อแทนความหมายของข้อมูลในระดับสูง โดยการสร้างสถาปัตยกรรมข้อมูลที่ประกอบด้วยโครงสร้างย่อย ๆ หลายโครงสร้าง ซึ่งแต่ละโครงสร้างได้มาจากการแปลงที่ไม่เป็นเชิงเส้น

### 2.8.2 ประเภทของการเรียนรู้เชิงลึก (Way to Deep Learning)

การเรียนรู้เชิงลึกเป็นเทคนิคหนึ่งของการเรียนรู้ของเครื่อง ซึ่งมีการแบ่งออกเป็นประเภทต่าง ๆ เพื่อให้เหมาะสมกับการนำไปใช้งาน โดยเต็ง และยู (Deng and Yu, 2014) ได้แบ่งประเภทของการเรียนรู้เชิงลึกออกเป็น 3 ประเภท ดังนี้

#### 2.8.2.1 การเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน (Unsupervised Deep Networks Learning)

การเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน มีวัตถุประสงค์เพื่อสร้างหรือทำนายความสัมพันธ์ของข้อมูลที่สังเกตหรือมองเห็นได้สำหรับการวิเคราะห์รูปแบบหรือการสังเคราะห์ โดยที่ไม่มีข้อมูลหรือการกำหนดคลาสเป้าหมาย ซึ่งการเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน สามารถแบ่งจำแนกออกเป็นเทคนิคต่าง ๆ เช่น โครงข่ายประสาทเทียมแบบลึก (Deep Artificial Neural Network) โครงข่ายประสาทเทียมแบบเพอร์เซพตรอนหลายชั้น (Multi-Layer Perceptron Network)

1) โครงข่ายประสาทเทียมแบบลึก (Deep Artificial Neural Network) เป็นโครงข่ายที่มีโครงสร้างลึก มีแนวโน้มที่จะสามารถคำนวณอะไรได้ซับซ้อน ซึ่งการเพิ่มจำนวนชั้นเป็นจำนวนมาก จะทำให้สามารถคำนวณหรือสร้างรูปแบบที่ซับซ้อนได้ แต่ก็ยังมีปัจจัยอื่น ๆ ที่ต้องคำนึงถึงจึงจะทำให้โครงข่ายประเภทนี้สามารถทำงานได้อย่างมีประสิทธิภาพ

2) โครงข่ายประสาทเทียมแบบเพอร์เซพตรอนหลายชั้น (Multi-Layer Perceptron Network) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลาย ๆ ชั้น ใช้สำหรับงานที่มีความซับซ้อนได้ผลเป็นอย่างดี ใช้ขั้นตอนการส่งค่าย้อนกลับ (Backpropagation) สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับ ประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า (Forward Pass) การส่งผ่านย้อนกลับ (Backward Pass) สำหรับการส่งผ่านไปข้างหน้า ข้อมูลจะผ่านเข้าโครงข่ายประสาทเทียมที่ชั้นข้อมูลเข้า และจะส่งผ่านจากอีกชั้นหนึ่งไปสู่อีกชั้นหนึ่งจนกระทั่งถึงชั้นข้อมูลออก ส่วนการส่งผ่านย้อนกลับค่าน้ำหนักการเชื่อมต่อจะถูกปรับเปลี่ยนให้สอดคล้องกับกฎการแก้ข้อผิดพลาด (Error-Correction) คือผลต่างของผลตอบที่แท้จริง (Actual Response) กับผลตอบเป้าหมาย (Target Response) เกิดเป็นสัญญาณผิดพลาด (Error Signal) ซึ่งสัญญาณผิดพลาดนี้จะถูกส่งย้อนกลับเข้าสู่โครงข่ายประสาทเทียมในทิศทางตรงกันข้ามกับการเชื่อมต่อ และค่าน้ำหนักของการเชื่อมต่อจะถูกปรับจนกระทั่งผลตอบที่แท้จริงเข้าใกล้ผลตอบเป้าหมาย

### 2.8.2.2 การเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน (Supervised Deep Networks

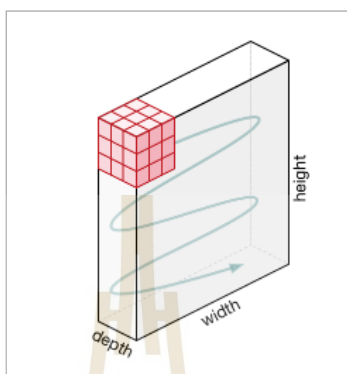
#### Learning)

การเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน มีวัตถุประสงค์เพื่อจำแนกหรือแยกแยะรูปแบบโดยตรง ซึ่งมักอธิบายลักษณะของข้อมูล ที่สามารถมองเห็นได้ มีการกำหนดคลาสเป้าหมายทั้งในรูปแบบทางตรงหรือทางอ้อมโดยเรียกอีกอย่างหนึ่งว่า เครือข่ายเชิงลึกแบบจำแนก โดยการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอนมีเทคนิคที่ถูกนำไปใช้เพื่อทำนายหรือจำแนกคลาสที่เป็นเป้าหมาย เช่น โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network)

1) โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) เป็นโครงข่ายประสาทเทียมชนิดที่มีการคำนวณซ้ำ นิยมใช้ในการวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลา

2) โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network) เป็นรูปแบบหนึ่งของการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน ซึ่งแต่ละโมดูลประกอบด้วยชั้นสังวัตนาการ และชั้นรวม โมดูลเหล่านี้มักจะซ้อนกันอยู่กับส่วนบนของอีกส่วนหนึ่งหรือมีโครงข่ายประสาทเทียมแบบลึก อยู่ด้านบนเพื่อสร้างแบบจำลองเชิงลึก การแบ่งปันน้ำหนักในชั้นรวมกับรูปแบบการรวมร่วมกันที่เหมาะสมทำให้โครงข่ายประสาทเทียมแบบสังวัตนาการมีคุณสมบัติบางอย่างที่เรียกว่า ความแปรปรวน (Invariance) ซึ่งโครงข่ายประสาทเทียมแบบสังวัตนาการ มีประสิทธิภาพและเป็นที่นิยมในการรับรู้ภาพและการประมวลผลเกี่ยวกับภาพ และการรู้จำเสียงพูด (Speech Recognition) เริ่มจากการกำหนดค่าใน ตัวกรอง (filter) หรือ เคอร์เนล (kernel) ที่

ช่วยดึงคุณลักษณะที่ใช้ในการรู้จำวัตถุออก โดยปกติตัวกรองอันหนึ่งจะดึงคุณลักษณะที่สนใจออกมาได้หนึ่งอย่าง จึงจำเป็นต้องมีตัวกรองหลายตัว เพื่อหาคุณลักษณะทางพื้นที่ต่าง ๆ ประกอบกัน ซึ่งในการหาคุณลักษณะตัวกรองจะเลื่อนมีลักษณะการเลื่อนของตัวกรองแสดงดังรูปที่ 2.12



ที่มา: Saha (2018)

รูปที่ 2.12 ลักษณะการเลื่อนของตัวกรอง

### 2.8.2.3 การเรียนรู้โครงข่ายเชิงลึกแบบผสม (Hybrid Deep Networks Learning)

การเรียนรู้โครงข่ายเชิงลึกแบบผสม มีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพของการเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน หรือการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน โดยสามารถใช้ทั้งการเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน และการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอนร่วมกันได้

## 2.9 งานวิจัยที่เกี่ยวข้อง

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการพัฒนาระบบแนะนำร้านอาหารพบว่า มีงานวิจัยที่ศึกษาเกี่ยวกับการพัฒนาแบบจำลองการแนะนำร้านอาหารจำนวนมาก โดยในการศึกษา งานวิจัยต่าง ๆ ผู้วิจัยได้จำแนกวัตถุประสงค์ในการพัฒนาระบบแนะนำร้านอาหารจากงานวิจัยที่ผ่านมา ซึ่งมีวัตถุประสงค์ที่แตกต่างกัน ได้แก่ การพัฒนาระบบแนะนำร้านอาหารเพื่อทำนายร้านอาหารที่ผู้ใช้ชอบหรือเลือกใช้บริการ การพัฒนาระบบแนะนำร้านอาหารเพื่อจำแนกประเภทร้านอาหาร การพัฒนาระบบแนะนำร้านอาหารเพื่อแนะนำร้านอาหารแก่ผู้ใช้โดยจำแนกตามลักษณะบุคคล และการพัฒนาระบบแนะนำร้านอาหารเพื่อทำนายรายได้ของร้านอาหาร โดยจากการศึกษางานวิจัยตามวัตถุประสงค์ต่าง ๆ มีรายละเอียด ดังนี้

งานวิจัยที่พัฒนาระบบแนะนำร้านอาหารเพื่อทำนายร้านอาหารที่ผู้ใช้ชอบหรือเลือกใช้บริการนั้น จากการศึกษาพบว่า มีการใช้วิธีในการแนะนำร้านอาหาร และการจำแนกประเภทร้านอาหารที่แตกต่างกัน โดยในการพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และการให้คะแนนจากผู้ใช้ ด้วยการจำแนกบทวิจารณ์และคะแนนที่ได้รับจากผู้ให้ โดยเปรียบเทียบระหว่างวิธีการถดถอยเชิงเส้น (Linear Regression) และวิธีการถดถอยเชิงตรรกะ (Logistic Regression) ผลที่ได้ พบว่าวิธีการถดถอยเชิงเส้นใช้เวลาในการแนะนำน้อยกว่าวิธีการถดถอยเชิงตรรกะ แต่วิธีการถดถอยเชิงตรรกะให้ค่าความถูกต้องในการแนะนำมากกว่าวิธีการถดถอยเชิงเส้น โดยอยู่ที่ร้อยละ 52.3 ในขณะที่วิธีการถดถอยเชิงตรรกะ ให้ค่าความถูกต้องในการแนะนำอยู่ที่ร้อยละ 50.0 (Gao, et al, 2015)

หลังจากนั้นได้มีงานวิจัยที่พัฒนาระบบแนะนำร้านอาหาร ซึ่งพิจารณาข้อมูลความสัมพันธ์ของผู้ใช้จากเครือข่ายสังคมออนไลน์ร่วมกับบทวิจารณ์บนเว็บไซต์ ผลการวิจัยที่ได้คือ ระบบแนะนำดังกล่าวสามารถให้ความถูกต้องในการแนะนำร้านอาหารได้ถูกต้องกว่าระบบของงานวิจัยอื่น ๆ ถึงร้อยละ 18.6 (Ma and Che, 2016) หลังจากนั้น จึงมีการพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และการให้คะแนนจากผู้ให้ ซึ่งจำแนกประเภทของบทวิจารณ์โดยใช้วิธีการจำแนกแบบเบย์อย่างง่าย (Naive Bayes) และวิเคราะห์ความรู้สึกด้านบวกและลบของผู้ใช้จากบทวิจารณ์และคะแนนที่ได้รับจากผู้ให้ (Bhojne, et al, 2017) โดยข้อเสนอแนะของงานวิจัยนี้ ได้นำเสนอเกี่ยวกับแนวทางในการเพิ่มความถูกต้องให้กับระบบแนะนำร้านอาหาร ด้วยวิธีแบบผสม ซึ่งจะช่วยให้ความถูกต้องในการแนะนำร้านอาหารสูงขึ้น

ต่อมาจึงมีงานวิจัยที่พัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพ โดยพิจารณาจากความชอบส่วนตัวของผู้ใช้ (Chu and Tsai, 2017) จำแนกประเภทของภาพที่ผู้ใช้อัปโหลดลงบนเว็บไซต์ โดยแบ่งออกเป็น 4 ประเภท ได้แก่ ภาพอาหาร ภาพเครื่องดื่ม ภาพร้านอาหารกลางแจ้ง และภาพร้านอาหารในร่ม ด้วยวิธีการจำแนกแบบโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks) และจำแนกบทวิจารณ์ประเภทข้อความด้วยวิธีการจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ผลที่ได้พบว่า ความถูกต้องในการแนะนำร้านอาหารมีความถูกต้องร้อยละ 70.78 จากงานวิจัยดังกล่าว จะเห็นว่าความถูกต้องในการแนะนำยังอยู่ในระดับที่สามารถเพิ่มประสิทธิภาพด้านความถูกต้องได้ โดยหากมีการหาความสัมพันธ์ระหว่างร้านอาหาร และผู้ใช้ที่มีความละเอียดมากขึ้น อาจทำให้ความถูกต้องในการแนะนำร้านอาหารแก่ผู้ใช้สูงขึ้นได้



ในส่วนองงำนวิจัยที่พัฒนาระบบแนะนำร้านอาหารเพื่อจำแนกประเภทร้านอาหาร จากบทวิจารณ์อาหารบนเว็บไซต์ ด้วยวิธีการจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน และวิเคราะห์ความรู้สึกด้านบวกและด้านลบของผู้ใช้ (Sentimental Analysis) (Yu, et al, 2017) ผลการวิจัยพบว่าระบบให้ความถูกต้องในการจำแนกความรู้สึกด้านบวกและด้านลบร้อยละ 88.90 และมีความถูกต้องในการจำแนกร้านอาหารร้อยละ 80 ส่วนงานวิจัยที่พัฒนาระบบแนะนำร้านอาหารเพื่อแนะนำร้านอาหารจากคุณลักษณะของร้านอาหาร ที่ตั้งร้านอาหาร และตำแหน่งของผู้ใช้ เพื่อแนะนำร้านอาหารแก่ผู้ใช้ตามตำแหน่งของผู้ใช้และร้านอาหาร ผลที่ได้พบว่า ระบบการแนะนำร้านอาหารมีประสิทธิภาพด้านความถูกต้องในการแนะนำแก่ผู้ใช่มากยิ่งขึ้น (Zeng, et al, 2016) นอกจากนี้ยังมีงานวิจัยที่พัฒนาระบบแนะนำร้านอาหารเพื่อทำนายรายได้ของร้านอาหาร ด้วยวิธีการจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีน โดยพิจารณาคุณลักษณะของร้านอาหาร 43 คุณลักษณะ เช่น ประเภทร้านอาหาร ที่ตั้งร้านอาหาร ข้อมูลทางการเงิน วันเปิดร้านอาหาร เป็นต้น ผลที่ได้พบว่า ระบบแนะนำดังกล่าวสามารถนำมาใช้ในการทำนายรายได้ของร้านอาหารได้ โดยวิธีการจำแนกแบบซัพพอร์ตเวกเตอร์แมชชีนมีความถูกต้องในการทำนายมากกว่าวิธีแบบเบย์อย่างง่าย (Raul, et al, 2016)

นอกจากนี้ งานวิจัยที่มีการประมวลผลเกี่ยวกับภาพ ได้มีการพัฒนาระบบตรวจสอบองค์ประกอบของฮาร์ดดิสก์ไดรฟ์ (อนุภาค ปัญญาสมบัติ, 2564) โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ และโวลูเมอร์ชัน 5 ผลการศึกษาที่ได้พบว่า ขั้นตอนวิธีดังกล่าวเป็นขั้นตอนวิธีที่มีประสิทธิภาพและสามารถตรวจสอบได้อย่างรวดเร็ว โดยความถูกต้องในการตรวจสอบองค์ประกอบของฮาร์ดดิสก์ไดรฟ์เป็นร้อยละ 99.49 ส่วนงานวิจัยที่ใช้โวลูเมอร์ชันสำหรับการประมวลผลภาพนั้น เป็นการพัฒนาระบบจำลองในการตรวจจับหนังสือแบบเรียลไทม์ (Kalinina, M., 2020) โดยทดสอบขั้นตอนวิธีของโวลูเมอร์ชัน 3 ทั้งหมด 3 ขั้นตอนวิธี คือแบบ 3 ผลลัพธ์แบบ 2 ผลลัพธ์ และแบบขนาดเล็ก (Tiny) ผลการศึกษาพบว่า ค่าความถูกต้องในการตรวจจับหนังสือมีค่าเป็นร้อยละ 94.43 92.35 และ 88.75 ตามลำดับ อีกทั้งมีงานวิจัยที่ตรวจจับศัตรูพืชที่ส่งผลกระทบต่อผลผลิตทางการเกษตรโดยใช้ขั้นตอนวิธีโวลูเมอร์ชัน (Luo, Y., et al, 2023) ผลการวิจัยที่ได้พบว่าสามารถระบุชนิดของศัตรูพืชได้ด้วยค่าความแม่นยำร้อยละ 96.8

โดยตารางการเปรียบเทียบงานวิจัยที่กล่าวมาข้างต้น แสดงดังตารางที่ 2.2

ตารางที่ 2.2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

| งานวิจัย    |                                      | Luo, Y.,<br>et al<br>(2023) | Punyathansombat,<br>A. (2021) | Kalinina,<br>M.<br>(2020) | Bhojne,<br>et al<br>(2017) | Chu and<br>Tsai<br>(2017) | Yu, et<br>al<br>(2017) | Zeng, et al<br>(2016) | Ma and<br>Che<br>(2016) | Raul, et<br>al<br>(2016) | Gao, et<br>al<br>(2015) | งานวิจัยนี้ |
|-------------|--------------------------------------|-----------------------------|-------------------------------|---------------------------|----------------------------|---------------------------|------------------------|-----------------------|-------------------------|--------------------------|-------------------------|-------------|
| Objective   | Personalized recommendation          | -                           | -                             | -                         | -                          | -                         | -                      | ✓                     | -                       | -                        | -                       | -           |
|             | Predicting favorite restaurants      | -                           | -                             | -                         | ✓                          | ✓                         | -                      | -                     | ✓                       | -                        | ✓                       | ✓           |
|             | Classify feature of restaurant types | -                           | -                             | -                         | -                          | -                         | ✓                      | -                     | -                       | -                        | -                       | -           |
|             | Classify feature of food             | -                           | -                             | -                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | ✓           |
|             | Predict the annual revenue           | -                           | -                             | -                         | -                          | -                         | -                      | -                     | -                       | ✓                        | -                       | -           |
| Data        | Text                                 | -                           | -                             | -                         | ✓                          | ✓                         | ✓                      | -                     | ✓                       | ✓                        | ✓                       | ✓           |
|             | Images                               | ✓                           | ✓                             | ✓                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | ✓           |
|             | Rating                               | -                           | -                             | -                         | ✓                          | -                         | -                      | -                     | -                       | -                        | ✓                       | -           |
|             | Demographic                          | -                           | -                             | -                         | -                          | -                         | -                      | -                     | -                       | ✓                        | -                       | -           |
|             | User preference                      | -                           | -                             | -                         | -                          | ✓                         | -                      | ✓                     | -                       | -                        | -                       | -           |
|             | User and restaurant Location         | -                           | -                             | -                         | -                          | -                         | -                      | ✓                     | -                       | -                        | -                       | ✓           |
| Data Source | Social Network                       | -                           | -                             | -                         | -                          | -                         | -                      | -                     | ✓                       | -                        | -                       | ✓           |
|             | Weblog/Website                       | -                           | -                             | -                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | -           |
|             | Offline Resource                     | ✓                           | ✓                             | ✓                         |                            |                           |                        |                       |                         |                          |                         |             |
| Environment | Mobile                               | -                           | -                             | -                         | -                          | -                         | -                      | ✓                     | -                       | -                        | -                       | -           |
|             | Website/System/Application           | ✓                           | -                             | ✓                         | ✓                          | ✓                         | ✓                      | -                     | ✓                       | ✓                        | ✓                       | ✓           |

ตารางที่ 2.2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง (ต่อ)

| งานวิจัย |                         | Luo, Y.,<br>et al<br>(2023) | Punyathansombat,<br>A. (2021) | Kalinina,<br>M.<br>(2020) | Bhojne,<br>et al<br>(2017) | Chu and<br>Tsai<br>(2017) | Yu, et<br>al<br>(2017) | Zeng, et al<br>(2016) | Ma and<br>Che<br>(2016) | Raul, et<br>al<br>(2016) | Gao, et<br>al<br>(2015) | งานวิจัยนี้ |
|----------|-------------------------|-----------------------------|-------------------------------|---------------------------|----------------------------|---------------------------|------------------------|-----------------------|-------------------------|--------------------------|-------------------------|-------------|
| Method   | Text Mining             | -                           | -                             | -                         | ✓                          | -                         | -                      | -                     | -                       | -                        | -                       | -           |
|          | SVM                     | -                           | -                             | -                         | -                          | -                         | ✓                      | -                     | -                       | ✓                        | -                       | ✓           |
|          | Regression              | -                           | -                             | -                         | -                          | -                         | -                      | -                     | -                       | -                        | ✓                       | -           |
|          | Content Based           | -                           | -                             | -                         | -                          | ✓                         | -                      | -                     | -                       | -                        | -                       | -           |
|          | Collaborative Filtering | -                           | -                             | -                         | ✓                          | ✓                         | -                      | -                     | ✓                       | -                        | -                       | -           |
|          | Sentiment Analysis      | -                           | -                             | -                         | ✓                          | -                         | ✓                      | -                     | -                       | -                        | -                       | ✓           |
|          | YOLO                    | ✓                           | ✓                             | ✓                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | ✓           |
|          | CNN                     | -                           | ✓                             | -                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | -           |
|          | Fass-RNN                | -                           | ✓                             | -                         | -                          | -                         | -                      | -                     | -                       | -                        | -                       | -           |



จากการทบทวนวรรณกรรม และงานวิจัยที่เกี่ยวข้อง จะเห็นว่างานวิจัยที่เกี่ยวข้องกับการพัฒนาระบบแนะนำร้านอาหาร มีวัตถุประสงค์ในการวิจัย และวิธีการแนะนำที่หลากหลาย ที่มีข้อดีและข้อจำกัดที่แตกต่างกัน งานวิจัยนี้ จึงเห็นความสำคัญของการแนะนำร้านอาหารแก่ผู้ใช้ และเพื่อให้ระบบแนะนำร้านอาหารสามารถแนะนำร้านอาหารที่มีความถูกต้องและตรงกับความต้องการของผู้ใช้ที่มีการใช้เครือข่ายสังคมออนไลน์เป็นจำนวนมาก งานวิจัยนี้จึงได้นำวิธีการแนะนำแบบต่าง ๆ มาใช้ โดยใช้ข้อมูลของผู้ใช้ และแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ เนื่องจากการแนะนำโดยใช้ข้อมูลจากบทวิจารณ์หรือภาพเพียงอย่างเดียว อาจไม่เพียงพอให้สามารถแนะนำร้านอาหารแก่ผู้ใช้ได้อย่างถูกต้อง อีกทั้งการใช้ข้อมูลของผู้ใช้จากเครือข่ายสังคมออนไลน์ จะช่วยให้สามารถแนะนำร้านอาหารที่ตรงกับความต้องการของผู้ใช้ได้มากขึ้น โดยงานวิจัยนี้มีกรอบแนวคิดการวิจัย ซึ่งจะกล่าวถึงในหัวข้อถัดไป

## 2.10 กรอบแนวคิดการวิจัย

งานวิจัย เรื่อง การพัฒนาระบบแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ เป็นงานวิจัยที่ใช้ข้อมูลบนสื่อสังคมออนไลน์ ประกอบด้วย บทวิจารณ์และภาพอาหารบนสื่อสังคมออนไลน์ โดยนำข้อมูลดังกล่าวมาจำแนกประเภทบทวิจารณ์และภาพ โดยใช้เทคนิคแบบเบย์อย่างง่ายและซัพพอร์ตเวกเตอร์แมชชีน ในการจำแนกประเภทของบทวิจารณ์ และใช้เทคนิคการประมวลผลภาพโวล์ เวอร์ชัน 5 ในการจำแนกภาพ ซึ่งจะนำมาสู่การจำแนกประเภทอาหาร โดยจะสามารถนำไปพัฒนาเป็นแบบจำลองในการแนะนำร้านอาหารแก่ผู้ใช้ ซึ่งกรอบแนวคิดงานวิจัยแสดงดังรูปที่ 2.13



รูปที่ 2.13 กรอบแนวคิดการวิจัย

## บทที่ 3

### วิธีดำเนินการวิจัย

การวิจัยเรื่อง การพัฒนาระบบแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ประกอบด้วยบทวิจารณ์ และภาพอาหาร โดยได้ออกแบบและพัฒนาแบบจำลองการแนะนำร้านอาหารโดยพิจารณาจากข้อมูลบนเครือข่ายสังคมออนไลน์ ซึ่งในบทนี้กล่าวถึงวิธีดำเนินการวิจัย ประกอบด้วยหัวข้อตามลำดับ ดังนี้

- 3.1 วิธีวิจัย
  - 3.2 เครื่องมือที่ใช้ในการวิจัย
  - 3.3 การเก็บรวบรวมข้อมูล
  - 3.4 การวิเคราะห์ข้อมูล
- โดยมีรายละเอียดดังนี้

#### 3.1 วิธีวิจัย

การวิจัยเรื่อง การพัฒนาระบบแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ เป็นการวิจัยเชิงประยุกต์ (Applied Research) ในหัวข้อนี้ได้กล่าวถึงวิธีวิจัยโดยใช้วงจรพัฒนาระบบ (System Development Life Cycle: SDLC) ซึ่งประกอบด้วย 4 ขั้นตอน ได้แก่ การศึกษาและวิเคราะห์ปัญหา การออกแบบ การพัฒนา และการประเมินผลแบบจำลองการจำแนกประเภทบทวิจารณ์ และจำแนกประเภทอาหารจากภาพอาหารบนเครือข่ายสังคมออนไลน์ ซึ่งอธิบายได้ดังนี้

##### 3.1.1 การศึกษาและวิเคราะห์ปัญหา

ในการศึกษาและวิเคราะห์ปัญหา เป็นการศึกษาข้อมูลเบื้องต้นจากเอกสารงานวิจัยที่เกี่ยวข้อง เพื่อนำข้อมูลดังกล่าวมาใช้เป็นแนวทางในการออกแบบและพัฒนาแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ซึ่งผู้วิจัยได้ศึกษาถึงปัญหาของการค้นหาร้านอาหาร และความเป็นไปได้ในการออกแบบและพัฒนาแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ โดยพบว่าปัจจัยที่มีอิทธิพลต่อการเลือกรับประทานอาหารของผู้บริโภค คือ การแนะนำ การเขียนบทวิจารณ์เกี่ยวกับอาหาร (บุญยาพร วุฒิธรรมคุณ, 2560) ซึ่งก่อนและระหว่างมื้ออาหาร ผู้บริโภคได้ศึกษาข้อมูลเกี่ยวกับอาหารและร้านอาหารจาก

เครือข่ายสังคมออนไลน์ และมักถ่ายภาพอาหารก่อนและขณะรับประทานอาหารและเผยแพร่ผ่านเครือข่ายสังคมออนไลน์เป็นประจำ จึงทำให้เกิดการเขียนบทวิจารณ์บนเครือข่ายสังคมออนไลน์ทั้งจากผู้ใช้อินเทอร์เน็ต และจากทีมงานที่ทำหน้าที่ในการแนะนำร้านอาหารผ่านช่องทางต่าง ๆ บนเครือข่ายสังคมออนไลน์ อย่างไรก็ตาม บทวิจารณ์ออนไลน์ที่เขียนโดยทีมงานที่ทำหน้าที่ในการแนะนำร้านอาหาร นำเสนอเพียงด้านดีของร้านอาหาร และให้ข้อมูลที่ไม่เรียลไทม์แก่ผู้บริโภค และการอ่านบทวิจารณ์อาหารในแต่ละครั้งใช้เวลานานในการตัดสินใจเลือกร้านอาหาร

จากการที่เครือข่ายสังคมออนไลน์ได้รับความนิยมในการแบ่งปันข้อมูลเกี่ยวกับอาหาร ประกอบกับพฤติกรรมของคนไทยที่นิยมถ่ายรูป โดยมักบอกเล่าประสบการณ์ และวิจารณ์อาหารผ่านเครือข่ายสังคมออนไลน์ ทำให้ข้อมูลเกี่ยวกับอาหาร และร้านอาหารมีจำนวนมากและกระจัดกระจาย (รัชณี พันธุ์จิตติ และคณะ, 2560) ซึ่งยากต่อการที่ผู้ใช้งานจะสามารถค้นหาข้อมูลได้ตรงกับความต้องการได้ ซึ่งจากการศึกษาปัญหาในการค้นหาร้านอาหารของผู้บริโภคพบว่า ในการรับประทานอาหารนอกบ้าน 1 ครั้ง ผู้บริโภคใช้เครือข่ายสังคมออนไลน์ 1-2 ประเภทในการค้นหาบทวิจารณ์เกี่ยวกับร้านอาหาร และใช้เวลาในการค้นหาบทวิจารณ์เกี่ยวกับร้านอาหาร 1-2 ชั่วโมง (บุญยาพร วุฒิชรรมคุณ, 2559) เนื่องจากมีจำนวนผู้ใช้อินเทอร์เน็ตจำนวนมาก และมีเครือข่ายสังคมออนไลน์หลายประเภท จึงทำให้ผู้บริโภคต้องใช้เวลา และต้องค้นหาจากเครือข่ายสังคมออนไลน์หลากหลายประเภท เพื่อให้ได้ข้อมูลที่ตรงกับความต้องการ

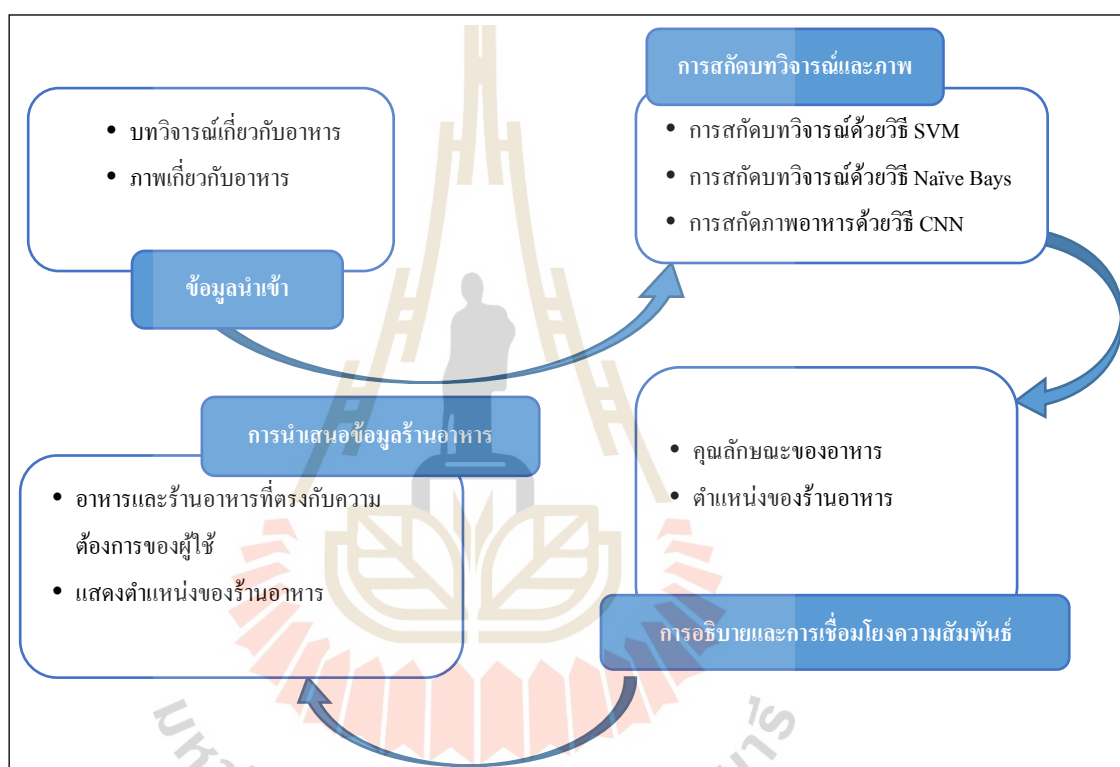
นอกจากนี้ จากการศึกษาความเป็นไปได้ในการออกแบบและพัฒนาแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์นั้น ยังพบว่า จากงานวิจัยที่ผ่านมา ประสิทธิภาพด้านความถูกต้องในการแนะนำยังอยู่ในระดับที่สามารถเพิ่มประสิทธิภาพได้ ซึ่งหากสามารถอธิบายความสัมพันธ์ระหว่างร้านอาหาร หรือเรียนรู้คุณลักษณะจากภาพที่มีความแตกต่างกัน จะทำให้สามารถแนะนำร้านอาหารที่ตรงกับความต้องการของผู้บริโภคมากขึ้น (Chu and Tsai, 2017) อีกทั้ง งานวิจัยที่ผ่านมาพบข้อจำกัดเกี่ยวกับการแนะนำร้านอาหาร คือการจำแนกประเภทร้านอาหารออกเป็นประเภทต่าง ๆ เพื่อสรุปว่าร้านอาหารประเภทนั้นมีข้อดีหรือข้อเสียอย่างไร ซึ่งเน้นไปที่กลุ่มนักท่องเที่ยวที่มีการเขียนและอ่านบทวิจารณ์บนเว็บไซต์ (Yu, Zhou, Zhang, and Cao, 2017) แต่ยังไม่มีการแนะนำร้านอาหารที่เจาะจงถึงสถานที่ และบ่งบอกคุณลักษณะเด่นและด้อยของร้านอาหารนั้น ๆ ผ่านการวิเคราะห์ข้อมูลจากเครือข่ายสังคมออนไลน์

จากการศึกษาและวิเคราะห์ปัญหาดังที่กล่าวมาข้างต้น งานวิจัยนี้จึงรวบรวมบทวิจารณ์และภาพเกี่ยวกับอาหาร เพื่อสกัดคุณลักษณะ และแนะนำร้านอาหารที่เฉพาะเจาะจงสำหรับผู้บริโภคแต่ละคน โดยพิจารณาจากความชอบของผู้บริโภคผ่านการใช้เครือข่ายสังคมออนไลน์ ร่วมกับคุณลักษณะของร้านอาหารที่ และแนะนำร้านอาหารที่มีคุณลักษณะตามความต้องการของผู้บริโภค



### 3.1.2 การออกแบบแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

ในการออกแบบแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ผู้วิจัยได้ใช้การออกแบบแบบจำลอง โดยพิจารณาข้อมูลจากบทวิจารณ์ และภาพบนเครือข่ายสังคมออนไลน์ โดยองค์ประกอบและกระบวนการของแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์แสดงดังรูปที่ 3.1



รูปที่ 3.1 องค์ประกอบและกระบวนการของแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

จากองค์ประกอบและกระบวนการของแบบจำลองการแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ สามารถแบ่งออกเป็นส่วนต่าง ๆ ดังนี้

#### ส่วนที่ 1 ข้อมูลนำเข้า

ในส่วนของข้อมูลนำเข้า งานวิจัยนี้ได้ใช้ข้อมูลนำเข้าจากเครือข่ายสังคมออนไลน์ ประกอบด้วย ข้อมูลของผู้ใช้ บทวิจารณ์ และภาพอาหารหรือร้านอาหารที่มีผู้ใช้ได้แบ่งปันผ่านเครือข่ายสังคมออนไลน์ เพื่อนำไปใช้ในการสกัดและจำแนกคุณลักษณะของอาหารและร้านอาหาร และนำไปแนะนำให้กับผู้ใช้ต่อไป

## ส่วนที่ 2 การสกัดบทวิจารณ์และภาพอาหาร

ในการสกัดบทวิจารณ์ออนไลน์นั้น ผู้วิจัยได้สกัดบทวิจารณ์ที่อยู่บนเครือข่ายสังคมออนไลน์ โดยในการสกัดบทวิจารณ์เกี่ยวกับอาหารนั้น ผู้วิจัยได้ใช้วิธีการดึงข้อมูลจากเครือข่ายสังคมออนไลน์ด้วย API ของเครือข่ายสังคมออนไลน์ และใช้การค้นหาด้วยคำสำคัญเกี่ยวกับอาหารหรือร้านอาหาร เพื่อนำบทวิจารณ์ที่เกี่ยวข้องกับอาหารหรือร้านอาหารมาสกัดและจำแนกคุณลักษณะของอาหาร และร้านอาหาร โดยคุณลักษณะของอาหารที่ได้จากการจำแนก เช่นรสชาติอาหาร ประเภทอาหาร ส่วนคุณลักษณะของร้านอาหาร เช่น ประเภทร้านอาหาร

ในส่วนของการสกัดภาพอาหารหรือร้านอาหาร ผู้วิจัยได้ใช้เทคนิคการเรียนรู้เชิงลึก (Deep Learning) ด้วยวิธี CNN สำหรับสกัดข้อมูลและจำแนกประเภทของรูปภาพ โดยใช้การจำแนกคุณลักษณะของภาพเป็นคุณลักษณะของอาหาร เช่น ชื่ออาหาร ประเภทเครื่องดื่ม อาหารคาว อาหารหวาน

## ส่วนที่ 3 การอธิบายและการเชื่อมโยงความสัมพันธ์

ในการอธิบายและเชื่อมโยงความสัมพันธ์นั้น เป็นการอธิบายความสัมพันธ์ระหว่างบทวิจารณ์กับภาพอาหาร และอาหารกับร้านอาหาร ซึ่งในการเชื่อมโยงความสัมพันธ์นั้น ผู้วิจัยได้ใช้ข้อมูลการบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์มาพิจารณาร่วมกัน เพื่อให้สามารถเชื่อมโยงความสัมพันธ์ระหว่างองค์ประกอบต่าง ๆ ได้อย่างถูกต้อง และตรงกับความต้องการของผู้ใช้

## ส่วนที่ 4 การนำเสนอข้อมูลร้านอาหาร

ในการนำเสนอข้อมูลร้านอาหารแก่ผู้ใช้นั้น งานวิจัยนี้ได้ใช้บทวิจารณ์และภาพอาหารหรือร้านอาหารบนเครือข่ายสังคมออนไลน์ ซึ่งการนำเสนอข้อมูลเกี่ยวกับร้านอาหารที่ผู้ใช้งานต้องการ จะแสดงผลในรูปแบบของลำดับการแนะนำ โดยพิจารณาจากประเภทของบทวิจารณ์ และประเภทอาหาร

### 3.1.3 การพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

ในส่วนของการพัฒนาแบบจำลองแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ งานวิจัยนี้ได้พัฒนาแบบจำลองตามกรอบการทำงานของแบบจำลองดังที่ได้อธิบายไว้ในหัวข้อ 3.1.2 ซึ่งเมื่อมีการพัฒนาแบบจำลองเสร็จสิ้น ได้มีการนำเสนอข้อมูลที่ได้จากการวิเคราะห์ในรูปแบบลำดับการแนะนำ ประเภทของบทวิจารณ์ และประเภทของอาหาร

### 3.1.4 การประเมินผลแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

ในการประเมินผลแบบจำลองการจำแนกประเภทบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ ประกอบด้วย การประเมินประสิทธิภาพของแบบจำลองการจำแนกประเภทบทวิจารณ์ และแบบจำลองการจำแนกประเภทอาหารจากภาพ ซึ่งในการประเมินประสิทธิภาพของแบบจำลองการจำแนกประเภทบทวิจารณ์ ได้ใช้การประเมินความถูกต้องของแบบจำลองด้วยค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก (Weighted Average<sub>p</sub>) ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก (Weighted Average<sub>r</sub>) ค่าวัดประสิทธิภาพ (F-Measure) และค่าความถูกต้อง (Accuracy) โดยข้อมูลที่นำมาทดสอบเป็นข้อมูลคนละชุดกับข้อมูลที่นำมาสร้างแบบจำลอง โดยได้ทดสอบแบบจำลองกับข้อมูลใหม่ (Unseen data)

ส่วนการประเมินประสิทธิภาพของแบบจำลองการจำแนกประเภทอาหาร ได้ประเมินประสิทธิภาพของแบบจำลองโดยคำนวณหาค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) ค่าวัดประสิทธิภาพ (F-Measure) และค่าความถูกต้อง (Accuracy) (Davis and Goadrich, 2006) ซึ่งในการประเมินผลแบบจำลอง มีสูตรดังสมการที่ (3.1) – (3.6)

$$Precision = \frac{TP}{(TP+FP)} \times 100\% \quad (3.1)$$

$$Recall = \frac{TP}{(TP+FN)} \times 100\% \quad (3.2)$$

$$Weighted\ Average_p = \frac{\sum (P_i \times N_i)}{N} \quad (3.3)$$

$$Weighted\ Average_r = \frac{\sum (R_i \times N_i)}{N} \quad (3.4)$$

$$F-Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.5)$$

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \times 100\% \quad (3.6)$$

ในการประเมินแบบจำลองการจำแนกประเภททวิภาคจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ แบ่งองค์ประกอบที่มีความเกี่ยวข้องออกเป็น 4 กลุ่ม ได้แก่ ข้อมูลที่สามารถทำนายได้และมีความเกี่ยวข้อง (True Positive: TP) ข้อมูลที่สามารถทำนายได้แต่ไม่มีความเกี่ยวข้อง (False Positive: FP) ข้อมูลที่ไม่สามารถทำนายได้และไม่มีความเกี่ยวข้อง (True Negative: TN) และข้อมูลที่ไม่สามารถทำนายได้แต่มีความเกี่ยวข้อง (False Negative: FN) (Jesse Davis Miao et al., 2009) ดังแสดงในตารางที่ 3.1

ตารางที่ 3.1 องค์ประกอบที่มีความเกี่ยวข้องกับการประเมินแบบจำลองการแนะนำร้านอาหาร

|                          | ข้อมูลที่มีความเกี่ยวข้อง (+) | ข้อมูลที่ไม่มีความเกี่ยวข้อง (-) |
|--------------------------|-------------------------------|----------------------------------|
| ข้อมูลที่ทำนายได้ (+)    | TP                            | FP                               |
| ข้อมูลที่ทำนายไม่ได้ (-) | FN                            | TN                               |

## 3.2 เครื่องมือที่ใช้ในการวิจัย

### 3.2.1 เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูล

- โปรแกรม Jupyter Notebook เวอร์ชัน 3.6 โดยใช้ภาษา Python
- Twitter API
- Roboflow
- YOLOv5

### 3.2.2 เครื่องมือที่ใช้ในการพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์

#### 1) ด้านฮาร์ดแวร์

- Computer Notebook ASUS Zenbook UX331UN

Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80 GHz

RAM 8 GB System type 64-bit

- Computer PC DESKTOP-P023GBJ

Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz 3.70 GHz

RAM 32 GB System type 64-bit

- อุปกรณ์อื่น ๆ เช่น เมาส์ คีย์บอร์ด

#### 2) ด้านซอฟต์แวร์

- Windows 11 64 bit

- โปรแกรม Jupyter Notebook เวอร์ชัน 3.6 โดยใช้ภาษา Python
- โปรแกรม Microsoft Excel 365
- Google Sheet
- Google AppScript
- Google Map

### 3.2.3 เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูล

เครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลประกอบด้วย

#### 1) เครื่องมือที่ใช้ในการจำแนกประเภททวิจาร์

- โปรแกรม Jupyter Notebook เวอร์ชัน 3.6 โดยใช้ภาษา Python
- โปรแกรม Microsoft Excel 365

#### 2) เครื่องมือที่ใช้ในการจำแนกประเภทอาหารบนภาพ

- Roboflow
- YOLO5
- โปรแกรม Microsoft Excel 365

#### 3) เครื่องมือที่ใช้ในการนำเสนอข้อมูลที่ได้จากแบบจำลอง

- Google Sheet
- Google AppScript
- Google Map

## 3.3 การเก็บรวบรวมข้อมูล

ในการออกแบบและพัฒนาแบบจำลองการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ผู้วิจัยเก็บรวบรวมข้อมูลจากเครือข่ายสังคมออนไลน์ ประกอบด้วย ข้อมูลผู้ใช้ บทวิจารณ์ที่เป็นข้อความและรูปภาพเกี่ยวกับอาหาร เพื่อใช้ในการวิเคราะห์ จำแนกประเภททวิจาร์ และจำแนกประเภทอาหาร และนำเสนอข้อมูลร้านอาหารให้แก่ผู้ใช้ โดยในการเก็บรวบรวมข้อมูล ผู้วิจัยได้รวบรวมข้อมูลการใช้งานเครือข่ายสังคมออนไลน์จากผู้ใช้งานโดยการดึงข้อมูลผู้ใช้จาก API ของ Twitter ประกอบด้วย ข้อมูลประวัติ (Profile) ของผู้ใช้ ข้อมูลการแบ่งปันเกี่ยวกับอาหารหรือร้านอาหารของผู้ใช้ ทั้งในรูปแบบข้อความและภาพ โดยรวบรวมข้อมูลทั้งหมดจำนวน 1,490 บทวิจารณ์ แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80 (1,192 บทวิจารณ์) และ 20 (298 บทวิจารณ์) ข้อมูลที่ใช้ทดสอบแบ่งเป็นด้านบวก เป็นกลาง และด้านลบ จำนวน 212 50 และ 36 บทวิจารณ์ตามลำดับ และวิเคราะห์ข้อมูลแบบ 10-folds cross validation





```

Status(_api=<tweepy.api.API object at 0x00000187ABBD3FD0>, _json=
  {'created_at': 'Mon Jun 11 21:44:33 +0000 2018',
   'id': 1006291118282047490,
   'id_str': '1006291118282047490',
   'text': 'I'm backkkkk https://t.co/56yNa5AxEg',
   'truncated': False,
   'entities':
     {'hashtags': [],
      'symbols': [],
      'user_mentions': [],
      'urls': [],
      'media': [{ 'id': 1006291116293963776,
                    'id_str': '1006291116293963776',
                    'indices': [13, 36],
                    'media_url': 'http://pbs.twimg.com/media/DfcQcE9VQAAzgK_.jpg',
                    'media_url_https': 'https://pbs.twimg.com/media/DfcQcE9VQAAzgK_.jpg',
                    'url': 'https://t.co/56yNa5AxEg',
                    'display_url': 'pic.twitter.com/56yNa5AxEg',
                    'expanded_url':

```

รูปที่ 3.4 ตัวอย่างการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้แบบมีข้อมูลรูปภาพ

ในการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้นั้น สามารถเรียกดูข้อมูลการ Tweet ได้ตามวัตถุ (Object) ที่ Twitter API กำหนด เช่น การเรียกดูข้อมูลการ Tweet ของผู้ใช้ ในรูปแบบข้อความที่ Tweet (text) หรือเรียกดูข้อมูลการ Tweet ของผู้ใช้ในรูปแบบวันและเวลา Tweet ข้อความนั้น ๆ (create\_at) เป็นต้น ซึ่งตัวอย่างการเรียกข้อมูลการ Tweet บน Twitter ของผู้ใช้ตามวัตถุ (Object) ที่ Twitter API กำหนด แสดงดังรูปที่ 3.5

```

In [55]: others_statuses = API.user_timeline('Khunnie0624', count=150)
for status in others_statuses:
    print(status.text)
    print(status.created_at)
    print('By ' + status.user.screen_name)
    if status.place: print('At ' + status.place.full_name)
    print('-----')

Thank you Beijing!! I had a blast!! See you again in Nanjing!
我的HOTTEST最棒!!! 🤔🤔🤔 https://t.co/cy0IDpxURb
2018-10-13 14:15:14
By Khunnie0624
-----

Let's take a walk. = ] https://t.co/QzKKENQCdE
2018-10-12 11:51:41
By Khunnie0624
-----

RT @UNICEF_Thailand: Art for the future แคมเปญนี้จะใช้งานศิลปะและซีรีส์ของอุณ
ร่วมเรียนรู้เกี่ยวกับงานศิลปะและศิลปะ
เพื่ออนาคตของโลก ๆ ธีรรัตน์ i...
2018-10-11 02:54:27
By Khunnie0624
-----

RT @follow_2PMJP: 「NICHKHUN (From 2PM) Premium Solo Concert 2018 "HOME"」

Hottest Japan第2次超最速先行受付とHottest Japan Mobile超最速先行受付がスタートしました! 受付は1...
2018-10-06 16:53:31

```

รูปที่ 3.5 ตัวอย่างการดึงข้อมูลการ Tweet บน Twitter ของผู้ใช้ตามวัตถุ (Object) ที่ Twitter API กำหนด

นอกจากนี้ ยังสามารถเรียกดูข้อมูลจากคำสำคัญ (Keywords) เพื่อค้นหาและเข้าถึงข้อมูลที่อยู่บน Twitter ได้ตามคำสำคัญที่ต้องการ ซึ่งในการค้นหาคำสำคัญดังกล่าว ได้ถูกนำมาใช้ในการค้นหาบทวิจารณ์และภาพที่เกี่ยวข้องกับอาหารหรือร้านอาหาร เช่น การค้นหาด้วยคำสำคัญ “bingsu” จะแสดงผล Tweet ที่ปรากฏคำว่า “bingsu” อยู่ในบทวิจารณ์ เป็นต้น ตัวอย่างการค้นหาด้วยคำสำคัญจากการ Tweet ของผู้ใช้ แสดงดังรูปที่ 3.6

|    | A                   | B                                  | C     | D              | E         | F       | G                               |
|----|---------------------|------------------------------------|-------|----------------|-----------|---------|---------------------------------|
| 1  | created             | text                               | retwc | hashtag        | followers | friends | username                        |
| 2  | 2018-10-16 18:43:13 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 239       | 265     | jjy//imy btob tt tt             |
| 3  | 2018-10-16 18:20:24 | RT @kwamfhunn: 🍳Salty egg yol      | 119   |                | 90        | 430     | วไย🍳                            |
| 4  | 2018-10-16 18:15:59 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 11        | 252     | 🍳人己古🍳                           |
| 5  | 2018-10-16 18:12:22 | [I Ate] Loaded bingsu (Korean shav | 0     |                | 699       | 116     | Argemiro Jaramillo              |
| 6  | 2018-10-16 17:46:51 | MUKA MACAM MAMAT MESYO ADI         | 0     |                | 511       | 198     | ur mader                        |
| 7  | 2018-10-16 17:30:14 | RT @Bear03_04: น้องหมีฝากบอกเป็    | 15    |                | 399       | 235     | shabu • bingsu💕                 |
| 8  | 2018-10-16 17:29:39 | RT @Bear03_04: @shabu_bingsu       | 1     | #น้องหมีฝากบอก | 399       | 235     | shabu • bingsu💕                 |
| 9  | 2018-10-16 17:28:28 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 60        | 88      | Sasa #Thumbs Up 🍳❤️             |
| 10 | 2018-10-16 17:18:20 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 260       | 832     | gyrating half a mm. away from t |
| 11 | 2018-10-16 17:10:57 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 562       | 242     | Hanna 🍳   semi hiatus           |
| 12 | 2018-10-16 17:09:28 | RT @peartunwarat: บางทีก็น้อยใจ    | 1125  |                | 399       | 235     | shabu • bingsu💕                 |
| 13 | 2018-10-16 17:08:35 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 376       | 216     | reg                             |
| 14 | 2018-10-16 17:07:32 | @shabu_bingsuเป็นคนใจร้าย ๗        | 1     | #น้องหมีฝากบอก | 44        | 0       | 🍳น้องหมีฝากบอก🍳                 |
| 15 | 2018-10-16 17:06:10 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 1235      | 935     | Chokiee🍳                        |
| 16 | 2018-10-16 17:02:22 | RT @Meltagon_10: Hongseok: ...+S   | 20    |                | 565       | 391     | haera 🍳🍳                        |
| 17 | 2018-10-16 16:59:24 | 【CAFE de BINGSU】@東京 新大             | 0     |                | 162       | 155     | 究極のグルメ情報🍳                       |
| 18 | 2018-10-16 16:53:49 | RT @foodyoucaneatfc: 🍳🍳🍳           | 3100  |                | 156       | 594     | N                               |
| 19 | 2018-10-16 16:45:21 | RT @ifernweh: You'll be okay soon  | 2032  |                | 399       | 235     | shabu • bingsu💕                 |
| 20 | 2018-10-16 16:45:01 | @minnanoseoyoo 오 ㅋㅋㅋ대학             | 0     |                | 36        | 90      | 달깨롱                             |
| 21 | 2018-10-16 16:44:25 | @Cooool_Bingsu 오 좀 각인데             | 0     |                | 16        | 302     | 합격시켜주세요                         |
| 22 | 2018-10-16 16:43:48 | @minnanoseoyoo 이과랑 이과워             | 0     |                | 36        | 90      | 달깨롱                             |
| 23 | 2018-10-16 16:43:28 | @minnanoseoyoo 너도 나중에 나            | 0     |                | 36        | 90      | 달깨롱                             |
| 24 | 2018-10-16 16:43:06 | @minnanoseoyoo 아니면                 | 0     |                | 36        | 90      | 달깨롱                             |
| 25 | 2018-10-16 16:42:53 | @Cooool_Bingsu 애틋 캐 작 파 고          | 0     |                | 16        | 302     | 합격시켜주세요                         |

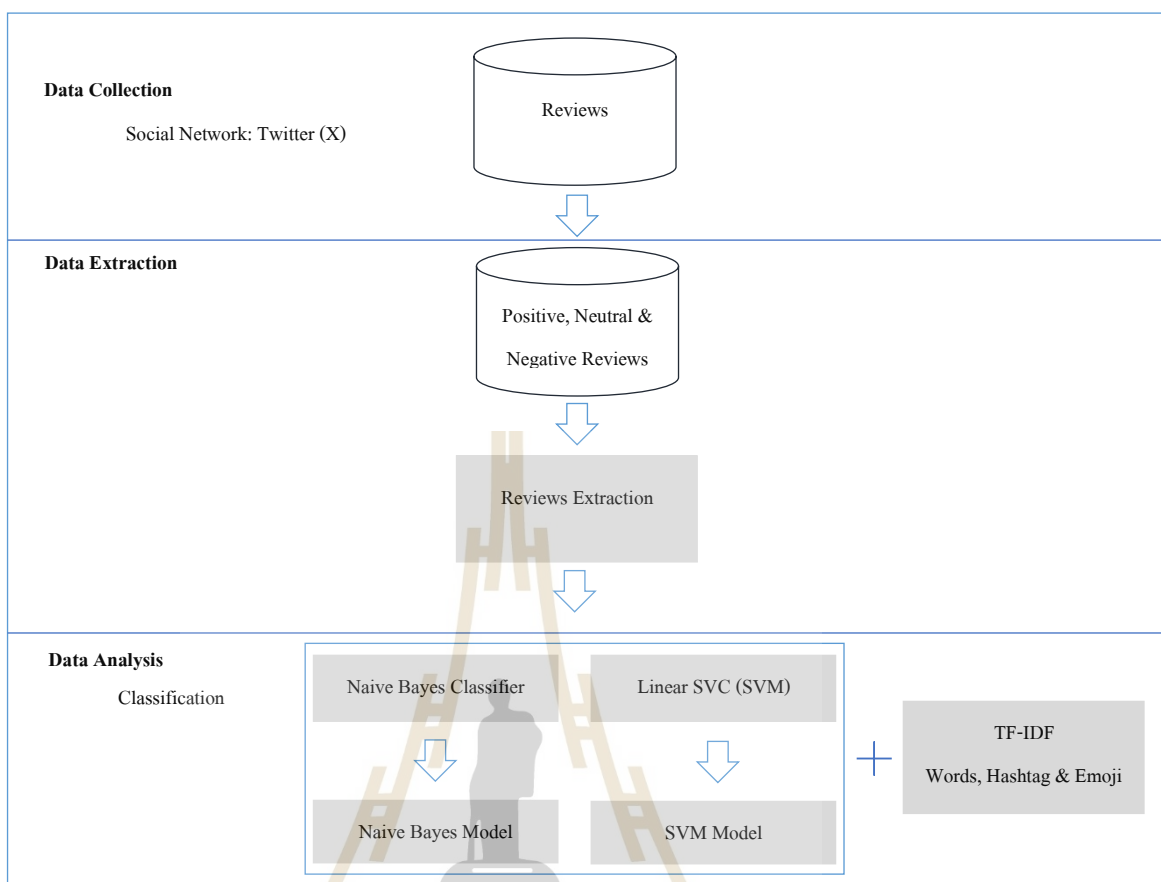
รูปที่ 3.6 ตัวอย่างการเรียกข้อมูลด้วยคำสำคัญบน Twitter

### 3.4 การวิเคราะห์ข้อมูล

ในการวิเคราะห์ข้อมูล ผู้วิจัยได้วิเคราะห์ข้อมูลจากเครือข่ายสังคมออนไลน์ด้วยวิธีการเขียนโปรแกรมดึงข้อมูลจากเครือข่ายสังคมออนไลน์ และวิเคราะห์การจำแนกประเภทบทวิจารณ์ และประเภทอาหารบนภาพบนเครือข่ายสังคมออนไลน์ โดยมีรายละเอียด ดังนี้

#### 3.4.1 การวิเคราะห์ข้อมูลจากบทวิจารณ์เพื่อจำแนกประเภทบทวิจารณ์

ในการวิเคราะห์ข้อมูลจากบทวิจารณ์เพื่อจำแนกประเภทบทวิจารณ์ งานวิจัยนี้ได้วิเคราะห์ข้อมูลจากบทวิจารณ์ออนไลน์ โดยจำแนกบทวิจารณ์ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เปรียบเทียบกับวิธีแบบเบย์อย่างง่าย (Naive Bays) ในการสร้างแบบจำลองการทำนายบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบของบทวิจารณ์ งานวิจัยนี้มีสถาปัตยกรรมของแบบจำลองแสดงดังรูปที่ 3.7



รูปที่ 3.7 สถาปัตยกรรมของแบบจำลองการจำแนกประเภททวิจารณ์

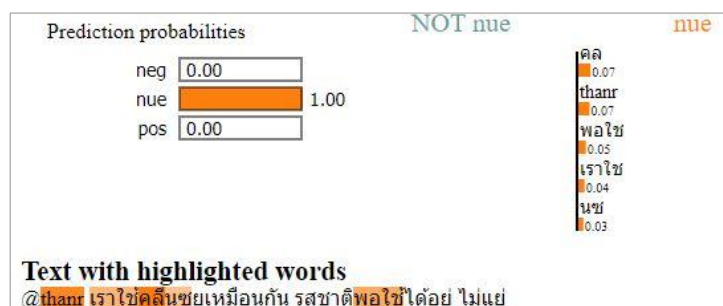
จากสถาปัตยกรรมของแบบจำลองการจำแนกประเภททวิจารณ์ ประกอบด้วยขั้นตอนสำหรับสร้างแบบจำลองซึ่งมีรายละเอียดดังนี้

#### 3.4.1.1 การรวบรวมข้อมูล (Data Collection)

งานวิจัยนี้รวบรวมบทวิจารณ์จากทวิตเตอร์ เป็นข้อมูลที่รวบรวมในเดือนสิงหาคม-ตุลาคม ปี 2562 โดยรวบรวมข้อมูลแบบออนไลน์ แล้วนำมาวิเคราะห์ข้อมูลแบบออฟไลน์ และพัฒนาวิธีการรวบรวมบทวิจารณ์ ด้วยภาษาไพธอน

#### 3.4.1.2 การสกัดข้อมูล (Data Extraction)

การจำแนกประเภททวิจารณ์ ได้จำแนกบทวิจารณ์ออกเป็น 3 ประเภท ได้แก่บทวิจารณ์ด้านบวก เป็นกลาง และลบ ซึ่งมีการจำแนกคำที่เกี่ยวข้องกับความเห็นด้านบวก เป็นกลาง และลบ จากความถี่ที่ปรากฏในบทวิจารณ์ เพื่อใช้ในการกำหนดคุณลักษณะเป้าหมาย (Label Attribute) ในการตัดคำในบทวิจารณ์ งานวิจัยนี้ใช้ PyThaiNLP ซึ่งเป็นไลบรารีสำหรับการตัดคำในภาษาไพธอน โดยตัวอย่างของคำที่เกี่ยวข้องกับการกำหนดคุณลักษณะเป้าหมาย เพื่อจำแนกประเภททวิจารณ์แสดงดังรูปที่ 3.8



รูปที่ 3.8 ตัวอย่างการจำแนกประเภทวิจารณ์จากคำที่เกี่ยวข้อง

จากตัวอย่างของการตัดคำที่เกี่ยวข้องกับบทวิจารณ์ ผลการวิเคราะห์ข้อความที่ปรากฏในบทวิจารณ์ จะใช้วิธีการเปรียบเทียบกับข้อความที่ปรากฏอยู่ในพจนานุกรมที่งานวิจัยนี้ได้รวบรวมคำที่เกี่ยวข้องกับการวิจารณ์อาหาร ในที่นี้เมื่อเจอคำว่า “พอใช้” จะถูกกำหนดให้บทวิจารณ์นี้เป็นบทวิจารณ์ด้านเป็นกลาง โดยตัวอย่างคำที่เกี่ยวข้องกับการวิจารณ์อาหารที่ใช้ในการจำแนกประเภทบทวิจารณ์ มาจากการใช้ขั้นตอนวิธีในการตัดคำจากบทวิจารณ์ที่มีการกำหนดคลาสไว้ก่อนแล้ว โดยคลาสที่กำหนดคือ บทวิจารณ์ด้านบวก เป็นกลาง และด้านลบ และนำคำที่ได้จากการตัดคำมาให้น้ำหนักด้วยวิธีการหาความถี่ของคำ (TF-IDF) ซึ่งคำที่ได้จากการคำนวณความถี่ของคำได้อธิบายไว้ในหัวข้อ การวิเคราะห์ข้อมูล (Data Analysis) โดยตัวอย่างของคำที่ใช้ในการจำแนกประเภทบทวิจารณ์แสดงดังตารางที่ 3.2

ตารางที่ 3.2 ตัวอย่างคำที่เกี่ยวข้องกับการวิจารณ์อาหาร

| บทวิจารณ์ด้านบวก | บทวิจารณ์ด้านเป็นกลาง | บทวิจารณ์ด้านลบ |
|------------------|-----------------------|-----------------|
| อร่อย            | พอได้                 | ไม่อร่อย        |
| อร่อยมาก         | พอใช้                 | แย่             |
| รสชาติดี         | พอใช้ได้              | แย่มาก          |
| อร่อยสุด         | โอเคอยู่              | ใช้ไม่ได้       |

### 3.4.1.3 การวิเคราะห์ข้อมูล (Data Analysis)

ในการวิเคราะห์ข้อมูลใช้จำแนกข้อมูล (Classification) ด้วยขั้นตอนวิธี Naive Bayes ซึ่งเป็นขั้นตอนวิธีที่จำแนกข้อมูลตามหลักความน่าจะเป็น เปรียบเทียบกับขั้นตอนวิธี Support Vector Machine (SVM) ซึ่งเป็นขั้นตอนวิธีที่มีการจำแนกข้อมูลในรูปแบบเชิงเส้น ซึ่งการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทพบทวิการณได้แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40 70:30 และ 80:20 โดยพบว่า การแบ่งแบบ 80:20 มีค่าความถูกต้องในการวิเคราะห์มากกว่า ซึ่งในการวิเคราะห์ข้อมูลโดยแบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40 แสดงดังตารางที่ 3.3

ตารางที่ 3.3 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40

| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40               |                           |                        |                           |          |
|-------------------------------------------------------------------------------------|---------------------------|------------------------|---------------------------|----------|
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.79                      | 0.72                   | 0.73                      | 0.72     |
| SVM                                                                                 | 0.82                      | 0.83                   | 0.82                      | 0.83     |
| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40 กับข้อมูลใหม่ |                           |                        |                           |          |
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.69                      | 0.71                   | 0.68                      | 0.71     |
| SVM                                                                                 | 0.80                      | 0.80                   | 0.80                      | 0.80     |

ส่วนผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 70:30 แสดงดังตารางที่ 3.4

ตารางที่ 3.4 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 70:30

| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 70:30               |                           |                        |                           |          |
|-------------------------------------------------------------------------------------|---------------------------|------------------------|---------------------------|----------|
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.69                      | 0.70                   | 0.68                      | 0.72     |
| SVM                                                                                 | 0.83                      | 0.82                   | 0.82                      | 0.82     |
| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 70:30 กับข้อมูลใหม่ |                           |                        |                           |          |
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.77                      | 0.79                   | 0.76                      | 0.79     |
| SVM                                                                                 | 0.78                      | 0.79                   | 0.76                      | 0.79     |

ส่วนผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 80:20 แสดงดังตารางที่ 3.5

ตารางที่ 3.5 ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 80:20

| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 80:20               |                           |                        |                           |          |
|-------------------------------------------------------------------------------------|---------------------------|------------------------|---------------------------|----------|
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.78                      | 0.80                   | 0.78                      | 0.80     |
| SVM                                                                                 | 0.82                      | 0.83                   | 0.82                      | 0.82     |
| ผลการวิเคราะห์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 80:20 กับข้อมูลใหม่ |                           |                        |                           |          |
| ขั้นตอนวิธี/ผลการทำนาย                                                              | Weighted Avg<br>Precision | Weighted Avg<br>Recall | Weighted Avg<br>F-Measure | Accuracy |
| Naive Bayes                                                                         | 0.77                      | 0.79                   | 0.76                      | 0.79     |
| SVM                                                                                 | 0.80                      | 0.81                   | 0.80                      | 0.81     |

จากผลการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทบทวิจารณ์ที่แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40 และ 70:30 มีค่าความถูกต้องจากการวิเคราะห์น้อยกว่า เนื่องจากการแบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80:20 นั้น ผลที่ได้ของทั้งขั้นตอนวิธี SVM และขั้นตอนวิธี Naive Bayes ให้ค่าความถูกต้องที่ใกล้เคียงกับสมมติฐานการวิจัย คือมีค่าความถูกต้องในการทำนายตั้งแต่ร้อยละ 80 ขึ้นไป จึงใช้การแบ่งจำนวนข้อมูลเพื่อสร้างและทดสอบ



แบบจำลองเป็นร้อยละ 80:20 จากข้อมูลทั้งหมดจำนวน 1,490 บทวิจารณ์ ข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80 (1,192 บทวิจารณ์) และ 20 (298 บทวิจารณ์) และได้้นำแบบจำลองที่ได้ไปทดสอบกับข้อมูลใหม่ (Unseen data) ข้อมูลที่ใช้ทดสอบแบ่งเป็นด้านบวก เป็นกลาง และด้านลบ จำนวน 212 50 และ 36 บทวิจารณ์ตามลำดับ วิเคราะห์ข้อมูลแบบ 10-folds cross validation และเปรียบเทียบระยะเวลาที่ใช้ในการทำนายของทั้งสองขั้นตอนวิธี ซึ่งรายละเอียดของผลการวิเคราะห์ข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80:20 ได้อธิบายไว้ในบทที่ 4 ผลการวิจัยและการอภิปรายผล

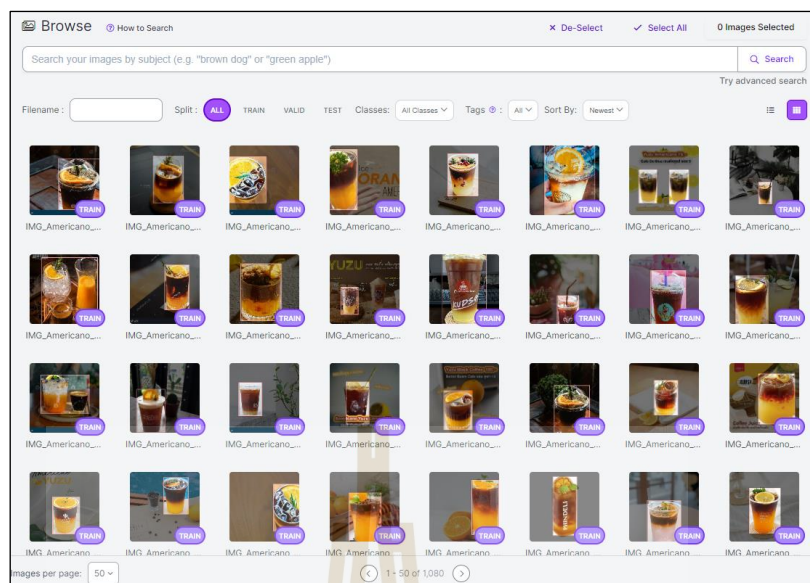
ทั้งนี้ จำนวนของบทวิจารณ์ที่รวบรวมมานั้นมีบทวิจารณ์ด้านบวกเป็นจำนวนมาก ซึ่งมีจำนวนแตกต่างกับบทวิจารณ์ด้านเป็นกลาง และด้านลบ ซึ่งอาจทำให้เกิดความไม่สมดุลกันของข้อมูลที่นำมาสร้างและทดสอบแบบจำลอง ในกรณีนี้สามารถใช้วิธีการสุ่มตัวอย่าง (Over-sampling) โดยสุ่มเพิ่มจำนวนข้อมูลกลุ่มน้อยให้มีปริมาณใกล้เคียงกับข้อมูลกลุ่มหลักสำหรับการจัดการให้ข้อมูลให้มีความสมดุลกัน

นอกจากนี้ ยังใช้ข้อมูลอื่น ๆ ที่ปรากฏอยู่ในบทวิจารณ์ ได้แก่ คำ แชชแท็ก และสัญลักษณ์ต่าง ๆ เพื่อใช้ในการวิเคราะห์ความถี่ (TF-IDF) ที่ปรากฏในบทวิจารณ์ที่เป็นข้อมูลใหม่ (Unseen Data) เพื่อใช้ระบุประเภทของบทวิจารณ์ โดยในการคำนวณหาความถี่ของคำ แชชแท็ก และสัญลักษณ์ต่าง ๆ ที่ปรากฏในบทวิจารณ์นั้นคำนวณจากบทวิจารณ์ที่เป็นข้อมูลใหม่ (Unseen Data) จำนวน 298 บทวิจารณ์ โดยได้หาความถี่หลังจากมีการตัดคำที่ปรากฏในบทวิจารณ์แล้ว ซึ่งตัวอย่างในการหาความถี่ของคำ แชชแท็ก และสัญลักษณ์ต่าง ๆ แสดงได้ดังนี้

หากในจำนวนบทวิจารณ์ทั้งหมด 298 บทวิจารณ์ พบคำว่า “อร่อย” จำนวน 120 คำ จะสามารถคำนวณด้วย จำนวนคำที่พบในบทวิจารณ์/จำนวนของคำในบทวิจารณ์ทั้งหมด ซึ่งคือ  $120/298 = 0.40$  ดังนั้นความถี่ของคำว่าอร่อยที่พบในบทวิจารณ์มีความถี่ 0.40

### 3.4.2 การวิเคราะห์ข้อมูลจากภาพเพื่อจำแนกประเภทอาหาร

ในการจำแนกประเภทอาหารจากภาพ งานวิจัยนี้สกัดภาพด้วยโบลี เวอร์ชัน 5 (YOLOv5) โดยใช้ Roboflow ซึ่งเป็นเว็บไซต์สำหรับการกำหนดคุณลักษณะ (Attributes) ของวัตถุบนภาพ เพื่อนำไปสร้างแบบจำลองการจำแนกประเภทอาหาร ซึ่งตัวอย่างการกำหนดคลาสให้กับภาพแสดงดังรูปที่ 3.9



รูปที่ 3.9 ตัวอย่างการกำหนดคลาสให้กับภาพ

งานวิจัยนี้ได้แบ่งประเภทของอาหารออกเป็น 9 คลาส โดยอ้างอิงจากอาหารที่ได้รับความนิยมในปี ค.ศ. 2022 (wongnai, 2565) ได้แก่ อเมริกาโน่ ชานมไข่มุก ข้าวปลาแกง ยำ ไข่แดงเต็มหมวยอ แซลมอนเสียบไม้ย่าง สปรักลิ่งโซดา ไก่เกาหลี ผลไม้ดอง และราเมนไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นคลาสละ 120 ภาพ และใช้ YOLO5 ในการสร้างแบบจำลองจำแนกประเภทอาหาร และทดสอบกับภาพที่เป็นข้อมูลทดสอบจำนวน 270 ภาพ แบ่งเป็นคลาสละ 30 ภาพ สำหรับทดสอบแบบจำลอง และนำเสนอผลการจำแนกคุณลักษณะของอาหารและร้านอาหารเพื่อใช้ในการแนะนำแก่ผู้ใช้ต่อไป ซึ่งในการนำเสนอข้อมูลแก่ผู้ใช้ ได้มีการดึงตำแหน่งของภาพ โดยรายละเอียดของภาพ (EXIF) ที่ได้ งานวิจัยนี้นำไปใช้เพื่อระบุตำแหน่งของร้านอาหารคือละติจูดและลองจิจูด เพื่อนำทางผู้ใช้ให้ทราบตำแหน่งของร้านอาหาร โดยตัวอย่างของการดึงข้อมูลจากภาพเพื่อระบุตำแหน่งของร้านอาหาร แสดงดังรูปที่ 3.10

```
{'imageTakenTime': '2023:03:23 13:01:47',
'geolocation_lat': 14.639513888888889,
'geolocation_lng': 101.46035833333333}
```

รูปที่ 3.10 การดึงข้อมูลจากภาพเพื่อระบุตำแหน่งของร้านอาหาร

## บทที่ 4

### ผลการวิจัยและการอภิปรายผล

งานวิจัยนี้รวบรวมบทวิจารณ์จากทวิตเตอร์ เป็นข้อมูลที่รวบรวมในเดือนสิงหาคม-ตุลาคม ปี พ.ศ. 2562 โดยรวบรวมข้อมูลแบบออนไลน์ แล้วนำมาวิเคราะห์ข้อมูลแบบออฟไลน์ และพัฒนาวิธีการรวบรวมบทวิจารณ์และภาพด้วยภาษาไพธอน โดยงานวิจัยนี้ได้แบ่งการวิเคราะห์ออกเป็น 2 ส่วน ได้แก่

ส่วนที่ 1 คือการวิเคราะห์ข้อมูลจากบทวิจารณ์เพื่อจำแนกประเภทบทวิจารณ์ เพื่อสร้างแบบจำลองการทำนายบทวิจารณ์ โดยจำแนกบทวิจารณ์เป็น 3 ด้าน ได้แก่ ด้านบวก ด้านเป็นกลาง และด้านลบของบทวิจารณ์ และเปรียบเทียบผลการจำแนกบทวิจารณ์ระหว่างขั้นตอนวิธี แบบเบย์อย่างง่าย (Naive Bayes) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) โดยรวบรวมข้อมูลทั้งหมดจำนวน 1,490 บทวิจารณ์ แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80 (1,192 บทวิจารณ์) และ 20 (298 บทวิจารณ์) ข้อมูลที่ใช้ทดสอบแบ่งเป็นด้านบวก (pos) ด้านเป็นกลาง (neu) และด้านลบ (neg) จำนวน 212 50 และ 36 บทวิจารณ์ตามลำดับ และวิเคราะห์ข้อมูลแบบ 10-folds cross validation ทั้งนี้ จากการแบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็น 60:40 70:30 และ 80:20 พบว่า การแบ่งแบบ 80:20 มีค่าความถูกต้องในการวิเคราะห์มากกว่า และได้้นำแบบจำลองที่ได้ไปทดสอบกับข้อมูลใหม่ (Unseen data) จำนวน 298 บทวิจารณ์ แบ่งเป็นบทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ จำนวน 209 55 และ 34 บทวิจารณ์ และเปรียบเทียบระยะเวลาที่ใช้ในการทำนายของทั้งสองขั้นตอนวิธี นอกจากนี้ ยังใช้ข้อมูลอื่น ๆ ที่ปรากฏอยู่ในบทวิจารณ์ ได้แก่ คำ แสขแท้ก และสัญลักษณ์ต่าง ๆ เพื่อใช้ในการวิเคราะห์ความถี่ (TF-IDF) ที่ปรากฏในบทวิจารณ์

ส่วนที่ 2 คือการวิเคราะห์ข้อมูลที่เป็นภาพเพื่อจำแนกประเภทอาหาร โดยแบ่งประเภทอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกาน้อยชู ชานมไข่มุก ข้าวปลาแกะ ยาไข่แดงเค็มหมูยอ แซลมอนเสียบไม้ย่าง สปรักกิ่งโซดา ไก่เกาหลี ผลไม้ดอง และราเมนไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นประเภทละ 120 ภาพ และใช้ YOLO5 ในการสร้างแบบจำลองจำแนกประเภทอาหาร และใช้ภาพที่เป็นข้อมูลสำหรับทดสอบแบบจำลอง จำนวน 270 ภาพ แบ่งเป็นประเภทละ 30 ภาพ และนำเสนอผลการจำแนกคุณลักษณะของอาหารและร้านอาหาร โดยผลการวิจัยมีดังนี้

#### 4.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes

ในการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes แบ่งผลการวิเคราะห์เป็น ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes เพื่อทำนายประเภทของบทวิจารณ์ว่าเป็นด้านบวก ด้านลบ หรือด้านเป็นกลาง และผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่ เพื่อทดสอบข้อมูลใหม่กับแบบจำลองที่มีการทำนายประเภทของบทวิจารณ์ โดยในหัวข้อนี้แบ่งผลการวิจัยดังนี้

##### 4.1.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bay เพื่อสร้างแบบจำลอง

ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes แสดงดังตารางที่ 4.1

ตารางที่ 4.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes

| ค่าที่ทำนาย/ค่าจริง | Precision | Recall | F-Measure |
|---------------------|-----------|--------|-----------|
| ด้านบวก             | 0.82      | 0.96   | 0.89      |
| เป็นกลาง            | 0.58      | 0.28   | 0.38      |
| ด้านลบ              | 0.81      | 0.58   | 0.68      |
| Weighted Avg        | 0.78      | 0.80   | 0.78      |

จากผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes พบว่า มีค่าความถูกต้องในการทำนายร้อยละ 80.20 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนักร้อยละ 78.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนักร้อยละ 80.00 และค่าวัดประสิทธิภาพร้อยละ 78.00 โดยผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบ แสดงดังตารางที่ 4.2

ตารางที่ 4.2 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี Naive Bayes

| ค่าที่ทำนาย/ค่าจริง | ด้านบวก | เป็นกลาง | ด้านลบ |
|---------------------|---------|----------|--------|
| ด้านบวก             | 204     | 33       | 11     |
| เป็นกลาง            | 6       | 14       | 4      |
| ด้านลบ              | 2       | 3        | 21     |

จากตารางที่ 4.2 ผลการจำแนกบทวิจารณ์ด้วยขั้นตอนวิธี Naive Bayes พบว่า แบบจำลองทำนายถูกต้องว่าเป็นบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบจำนวน 204 14 และ 21 บทวิจารณ์ตามลำดับ โดยระยะเวลาในการประมวลผลคือ 1.51 วินาที

#### 4.1.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bay กับข้อมูลใหม่

เมื่อนำแบบจำลองที่ได้ไปทดสอบกับข้อมูลใหม่ ได้ผลลัพธ์ดังตารางที่ 4.3

ตารางที่ 4.3 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่

| ค่าที่ทำนาย/ค่าจริง | Precision | Recall | F-Measure |
|---------------------|-----------|--------|-----------|
| ด้านบวก             | 0.81      | 0.96   | 0.88      |
| เป็นกลาง            | 0.55      | 0.29   | 0.38      |
| ด้านลบ              | 0.83      | 0.56   | 0.67      |
| Weighted Avg        | 0.77      | 0.79   | 0.76      |

จากผลการวิเคราะห์ด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่ พบค่าความถูกต้องในการทำนายร้อยละ 78.86 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนักร้อยละ 77.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนักร้อยละ 79.00 และค่าวัดประสิทธิภาพร้อยละ 76.00 ซึ่งทำนายบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบกับข้อมูลใหม่แสดงดังตารางที่ 4.4

ตารางที่ 4.4 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่

| ค่าที่ทำนาย/ค่าจริง | ด้านบวก | เป็นกลาง | ด้านลบ |
|---------------------|---------|----------|--------|
| ด้านบวก             | 200     | 35       | 11     |
| เป็นกลาง            | 9       | 16       | 4      |
| ด้านลบ              | 0       | 4        | 19     |

ในการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี Naive Bayes กับข้อมูลใหม่ ผลที่ได้พบว่า ผลการทำนายบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบมีความถูกต้องจำนวน 200 16 และ 19 บทวิจารณ์ตามลำดับ มีระยะเวลาในการประมวลผล 1.54 วินาที

## 4.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM

ในการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM แบ่งผลการวิเคราะห์เป็น ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM เพื่อทำนายประเภทของบทวิจารณ์ว่าเป็นด้านบวก ด้านลบ หรือด้านเป็นกลาง และผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่ เพื่อทดสอบข้อมูลใหม่กับแบบจำลองที่มีทำนายประเภทของแบบจำลอง ซึ่งผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM มีดังนี้

### 4.2.1 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM เพื่อสร้างแบบจำลอง

ในการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM เพื่อสร้างแบบจำลอง แสดงดังตารางที่ 4.5

ตารางที่ 4.5 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM

| ค่าที่ทำนาย/ค่าจริง | Precision | Recall | F-Measure |
|---------------------|-----------|--------|-----------|
| ด้านบวก             | 0.86      | 0.95   | 0.91      |
| เป็นกลาง            | 0.62      | 0.48   | 0.54      |
| ด้านลบ              | 0.84      | 0.58   | 0.69      |
| Weighted Avg        | 0.82      | 0.83   | 0.82      |

จากผลการวิเคราะห์ด้วย SVM มีค่าความถูกต้องในการทำนายร้อยละ 82.89 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 82.00 ค่าความระลึกละเอียดถ่วงน้ำหนัก ร้อยละ 83.00 และค่าวัดประสิทธิภาพ ร้อยละ 82.00 ซึ่งผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM แสดงดังตารางที่ 4.6

ตารางที่ 4.6 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM

| ค่าที่ทำนาย/ค่าจริง | ด้านบวก | เป็นกลาง | ด้านลบ |
|---------------------|---------|----------|--------|
| ด้านบวก             | 202     | 23       | 9      |
| เป็นกลาง            | 9       | 24       | 6      |
| ด้านลบ              | 1       | 3        | 21     |

จากผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM พบว่าผลการทำนายบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบมีความถูกต้องจำนวน 202 24 และ 21 บทวิจารณ์ตามลำดับ โดยระยะเวลาในการประมวลผลคือ 1.38 วินาที



#### 4.2.2 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่

ในการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM เพื่อทดสอบแบบจำลองกับข้อมูลใหม่ ผลที่ได้แสดงดังตารางที่ 4.7

ตารางที่ 4.7 ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่

| ค่าที่ทำนาย/ค่าจริง | Precision | Recall | F-Measure |
|---------------------|-----------|--------|-----------|
| ด้านบวก             | 0.85      | 0.91   | 0.88      |
| เป็นกลาง            | 0.58      | 0.51   | 0.54      |
| ด้านลบ              | 0.85      | 0.65   | 0.73      |
| Weighted Avg        | 0.80      | 0.81   | 0.80      |

จากผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM พบค่าความถูกต้องในการทำนายร้อยละ 80.87 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 80.00 ค่าความระลึกละเอียดถ่วงน้ำหนัก ร้อยละ 81.00 และค่าวัดประสิทธิภาพ ร้อยละ 80.00 ซึ่งผลการทำนายบทวิจารณ์ด้านบวก ด้านเป็นกลางกลาง และด้านลบด้วยขั้นตอนวิธี SVM แสดงดังตารางที่ 4.8

ตารางที่ 4.8 ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่

| ค่าที่ทำนาย/ค่าจริง | ด้านบวก | เป็นกลาง | ด้านลบ |
|---------------------|---------|----------|--------|
| ด้านบวก             | 191     | 24       | 9      |
| เป็นกลาง            | 17      | 28       | 3      |
| ด้านลบ              | 1       | 3        | 22     |

ผลการจำแนกบทวิจารณ์ด้านบวก เป็นกลาง และด้านลบด้วยขั้นตอนวิธี SVM กับข้อมูลใหม่พบว่าผลทำนายบทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบถูกต้องจำนวน 191 28 และ 22 บทวิจารณ์ตามลำดับ โดยระยะเวลาในการประมวลผลคือ 1.43 วินาที

### 4.3 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF)

นอกจากการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทบทวิจารณ์ด้วยขั้นตอนวิธีแบบ Naive Bayes และ SVM แล้ว งานวิจัยนี้ได้วิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ แสขเท็ก และสัญลักษณ์ต่าง ๆ จากบทวิจารณ์ที่เป็นข้อมูลใหม่ (Unseen Data) โดยจัดอันดับ 3 อันดับแรกเพื่อนำไปใช้ในการระบุประเภทของบทวิจารณ์ โดยผลที่ได้จากการนับความถี่คำ แสขเท็ก และสัญลักษณ์ ประกอบด้วยผลการวิเคราะห์ดังรายละเอียดต่อไปนี้

#### 4.3.1 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ

ในการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ เพื่อนับจำนวนคำที่แสดงในบทวิจารณ์ที่บ่งบอกถึงความรู้สึกด้านบวก ด้านลบ และเป็นกลาง ผลการวิเคราะห์แสดงดังตารางที่ 4.9

ตารางที่ 4.9 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ

| ลำดับ | ด้านบวก  | ความถี่  |
|-------|----------|----------|
| 1     | อร่อย    | 0.018524 |
| 2     | อร่อยมาก | 0.013394 |
| 3     | ดี       | 0.010429 |
| ลำดับ | เป็นกลาง | ความถี่  |
| 1     | กินได้   | 0.017096 |
| 2     | อยากกิน  | 0.073787 |
| 3     | พอได้    | 0.009799 |
| ลำดับ | ด้านลบ   | ความถี่  |
| 1     | เหม็น    | 0.036148 |
| 2     | ชาติหมา  | 0.025803 |
| 3     | อย่าแฉก  | 0.024391 |

จากผลการวิเคราะห์ความถี่คำพบว่า อันดับแรกของคำในบทวิจารณ์ด้านบวก คือ อร่อย มีความถี่ร้อยละ 1.85 ส่วนบทวิจารณ์ที่เป็นเป็นกลาง คำที่ปรากฏเป็นอันดับแรกคือ กินได้ มีความถี่ร้อยละ 1.71 และคำในบทวิจารณ์ด้านลบเป็นอันดับแรกคือ เหม็น ซึ่งเป็นคำสบถ มีความถี่ร้อยละ 3.61

#### 4.3.2 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของแฮชแท็ก

ในการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของแฮชแท็ก เพื่อนับจำนวนแฮชแท็กที่แสดงในบทวิจารณ์ ผลการวิเคราะห์แสดงดังตารางที่ 4.10

ตารางที่ 4.10 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของแฮชแท็ก

| ลำดับ | ด้านบวก          | ความถี่  |
|-------|------------------|----------|
| 1     | #อร่อยไปแดก      | 0.125069 |
| 2     | #อร่อยบอกต่อ     | 0.087849 |
| 3     | #รีวิวคาเฟ่      | 0.059043 |
| ลำดับ | เป็นกลาง         | ความถี่  |
| 1     | #อร่อยไปแดก      | 0.028296 |
| 2     | #รีวิวเซเว่น     | 0.023350 |
| 3     | #รีวิวเกาหลี     | 0.022144 |
| ลำดับ | ด้านลบ           | ความถี่  |
| 1     | #ไม่อร่อยอย่าแดก | 0.021010 |
| 2     | #รีวิวเซเว่น     | 0.019869 |
| 3     | #ชิมช้อปใช้      | 0.016575 |

จากผลการวิเคราะห์ความถี่ของแฮชแท็กในบทวิจารณ์ด้านบวกและบทวิจารณ์ที่เป็นกลางมีความถี่เป็นอันดับแรกคือ #อร่อยไปแดก มีความถี่ร้อยละ 12.51 และ 2.83 ส่วนแฮชแท็กที่ปรากฏในบทวิจารณ์ที่ด้านลบที่มีความถี่เป็นอันดับแรกคือ #ไม่อร่อยอย่าแดก มีความถี่ร้อยละ 2.10

#### 4.3.3 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของสัญลักษณ์

ในการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของสัญลักษณ์ เพื่อนับจำนวนสัญลักษณ์ที่แสดงในบทวิจารณ์ ผลการวิเคราะห์แสดงดังตารางที่ 4.11

ตารางที่ 4.11 ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของสัญลักษณ์

| ลำดับ | ด้านบวก  | ความถี่  |
|-------|----------|----------|
| 1     | 😊        | 0.023207 |
| 2     | 😄        | 0.015781 |
| 3     | 😂        | 0.013755 |
| ลำดับ | เป็นกลาง | ความถี่  |
| 1     | 😄        | 0.015972 |
| 2     | 😂        | 0.013215 |
| 3     | 😂        | 0.009011 |
| ลำดับ | ด้านลบ   | ความถี่  |
| 1     | 😂        | 0.025229 |
| 2     | 😞        | 0.020167 |
| 3     | 😂        | 0.016575 |

จากตารางที่ 4.11 สัญลักษณ์อันดับแรกที่ปรากฏในบทวิจารณ์ด้านบวกและด้านลบคือ 😊 มีความถี่ร้อยละ 2.32 และ 2.52 ส่วนบทวิจารณ์ที่เป็นกลางมีสัญลักษณ์ 😄 อยู่ในอันดับแรก โดยมีความถี่ร้อยละ 1.60 ทั้งนี้ ยังพบสัญลักษณ์อื่น ๆ ที่ใช้ในการแสดงอารมณ์ เช่น 😂 😄 😞 😂 😂 😂 และ 🙌 ปรากฏในบทวิจารณ์ทั้ง 3 ประเภทอีกด้วย

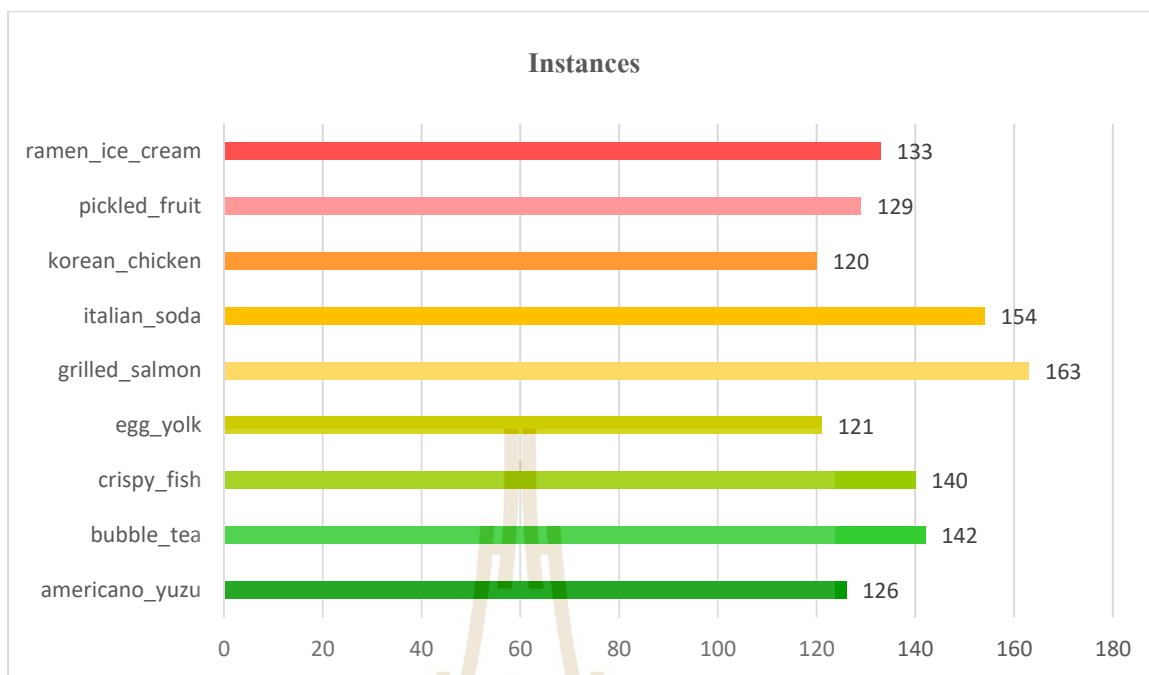
#### 4.4 ผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพ

ในการจำแนกประเภทอาหาร โดยจำแนกวัตถุดิบภาพ งานวิจัยนี้ใช้วิธีการจำแนกภาพด้วยเครื่องมือ YOLO5 เพื่อสร้างแบบจำลอง โดยแบ่งคลาสอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกาน้อย ชูชู ชานมไข่มุก ข้าวปลาแกะ ยำไข่แดงเค็มหมูยอ แซลมอนเสียบไม้ย่าง สปรังคั้งโซดา ไก่เกาหลี ผลไม้ดอง และราเมนไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นประเภทละ 120 ภาพ และทดสอบกับภาพที่เป็นข้อมูลทดสอบเพื่อทดสอบแบบจำลอง จำนวน 270 ภาพ แบ่งเป็นประเภทละ 30 ภาพ ซึ่งกำหนดจำนวนครั้ง (Epochs) ในการเรียนรู้รายละเอียดบนภาพเพื่อสร้างแบบจำลอง (Training) จำนวน 50 ครั้ง โดยผลที่ได้จากการสร้างแบบจำลองแสดงดังตารางที่ 4.8

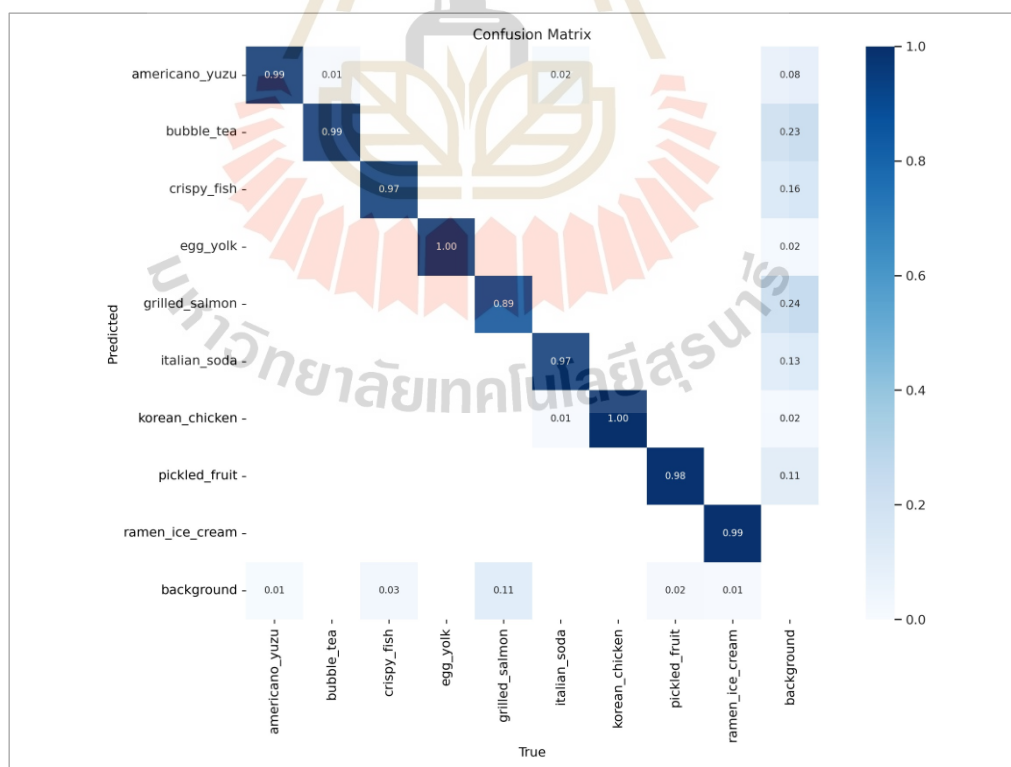
ตารางที่ 4.12 ผลการสร้างแบบจำลองจำแนกประเภทอาหารบนภาพ

| Class           | Images | Instances | Precision | Recall | Accuracy | F-Measure |
|-----------------|--------|-----------|-----------|--------|----------|-----------|
| All             | 1080   | 1228      | 0.94      | 0.96   | 0.98     | 0.95      |
| americano_yuzu  | 120    | 126       | 0.91      | 0.99   | 0.99     | 0.95      |
| bubble_tea      | 120    | 142       | 0.88      | 0.99   | 0.99     | 0.93      |
| crispy_fish     | 120    | 140       | 0.93      | 0.93   | 0.97     | 0.93      |
| egg_yolk        | 120    | 121       | 0.98      | 1.00   | 1.00     | 0.99      |
| grilled_salmon  | 120    | 163       | 0.89      | 0.82   | 0.89     | 0.85      |
| italian_soda    | 120    | 154       | 0.95      | 0.98   | 0.97     | 0.96      |
| korean_chicken  | 120    | 120       | 0.98      | 0.99   | 1.00     | 0.98      |
| pickled_fruit   | 120    | 129       | 0.93      | 0.97   | 0.98     | 0.95      |
| ramen_ice_cream | 120    | 133       | 1.00      | 0.99   | 0.99     | 0.99      |

จากตารางที่ 4.12 ผลการสร้างแบบจำลองประเภทอาหารบนภาพ โดยแบ่งประเภทอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกาโนยuzu (americano\_yuzu) ชานมไข่มุก (bubble\_tea) ข้าวปลาแกะ (crispy\_fish) ยำไข่แดงเค็มหมูยอ (egg\_yolk) แซลมอนเสียบไม้ย่าง (grilled\_salmon) สปร้ากิ้งโซดา (italian\_soda) ไก่เกาหลี (korean\_chicken) ผลไม้ดอง (pickled\_fruit) และราเมนไอศกรีม (ramen\_ice\_cream) ประเภทละ 120 ภาพ รวมทั้งหมด 1,080 ภาพ จำนวน 1,228 อินสแตนซ์ พบว่ามีค่าความแม่นยำ (Precision) โดยรวมเป็นร้อยละ 94 ค่าความระลึก (Recall) ร้อยละ 96 ค่าความถูกต้อง (Accuracy) ร้อยละ 98 และค่าวัดประสิทธิภาพ (F-Measure) ร้อยละ 95 ซึ่งผลการจำแนกประเภทอาหารเพื่อสร้างแบบจำลองของอาหารประเภทไข่แดงเค็มหมูยอ มีค่าความแม่นยำร้อยละ 98 ค่าความระลึกและค่าความถูกต้องร้อยละ 100 ซึ่งมีค่าความระลึกและค่าความถูกต้องในการจำแนกประเภทอาหารสูงที่สุด ส่วนไก่เกาหลี ซึ่งมีค่าความถูกต้องในการทำนายสูงที่สุดเป็นร้อยละ 100 มีค่าความแม่นยำร้อยละ 98 และค่าความระลึกร้อยละ 99 ส่วนราเมนไอศกรีม มีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุด โดยผลจำนวนของอินสแตนซ์ที่ปรากฏบนภาพที่นำไปสร้างแบบจำลองและผลการสร้างแบบจำลองเพื่อจำแนกประเภทอาหารแบบ Confusion matrix แสดงดังรูปที่ 4.1 และ 4.2 ตามลำดับ



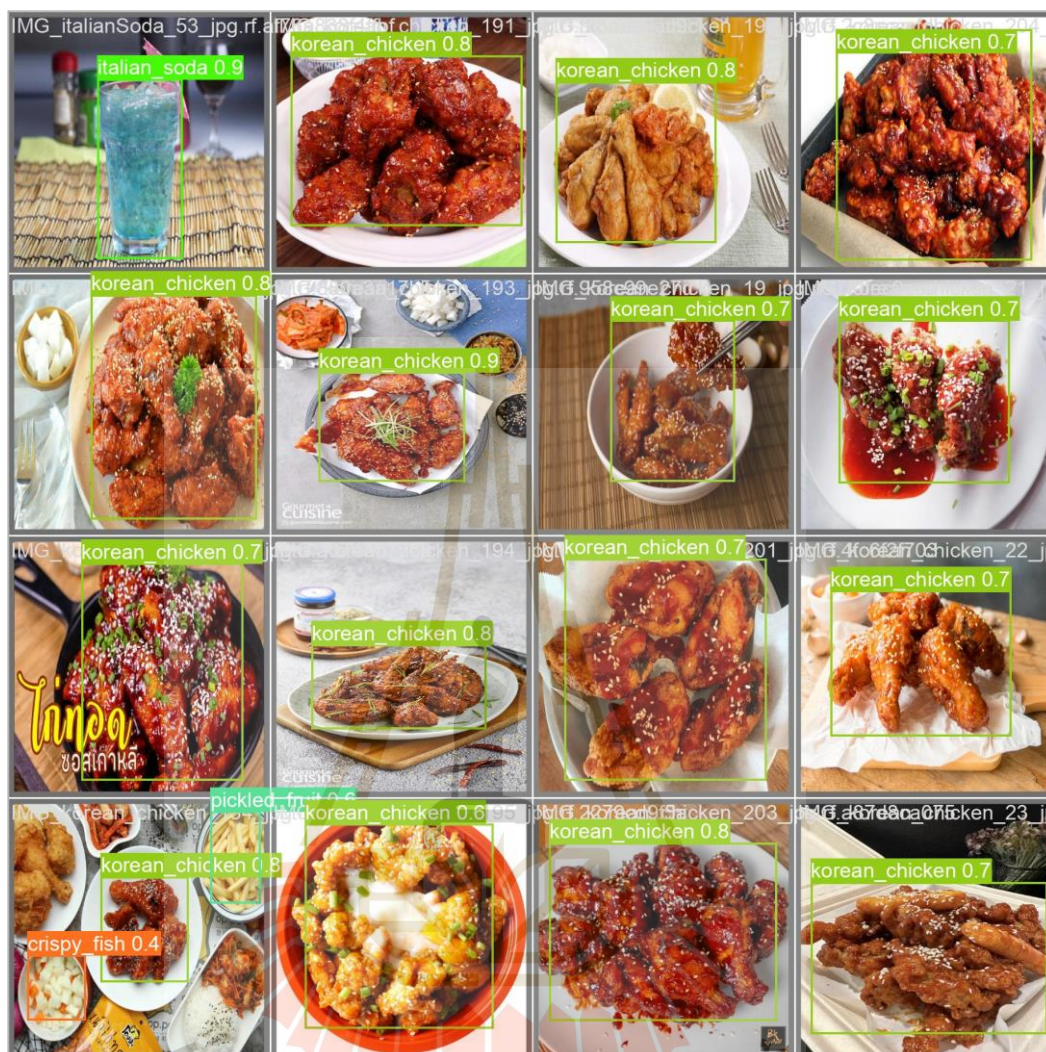
รูปที่ 4.1 จำนวนของอินสแตนซ์ที่ปรากฏบนภาพที่นำไปสร้างแบบจำลอง



รูปที่ 4.2 Confusion matrix แสดงผลการสร้างแบบจำลองจำแนกประเภทอาหาร



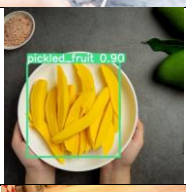
ตัวอย่างภาพจากการสร้างแบบจำลองการจำแนกประเภทอาหาร แสดงดังรูปที่ 4.3



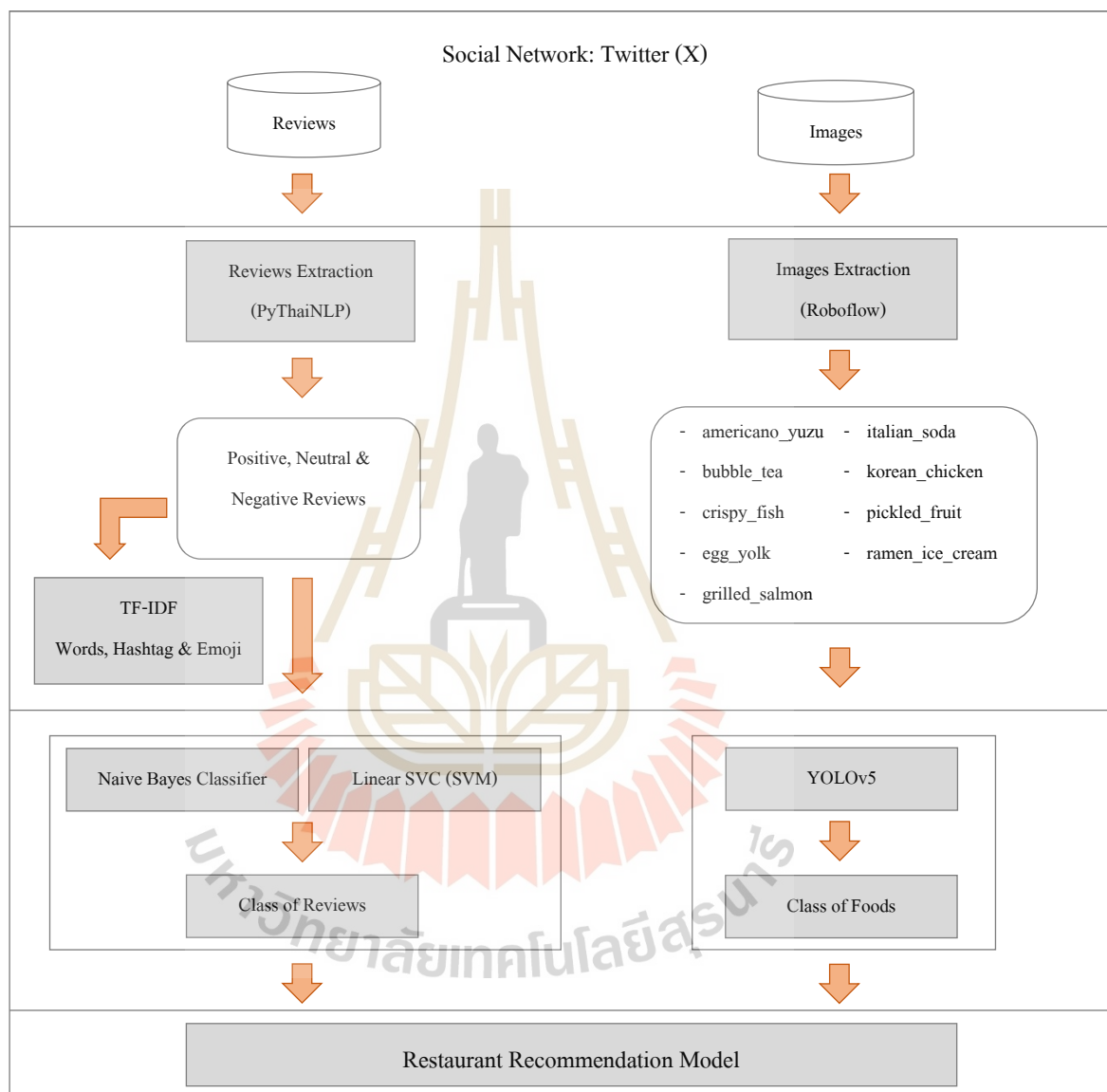
รูปที่ 4.3 ตัวอย่างภาพจากการสร้างแบบจำลองการจำแนกประเภทอาหาร

เมื่อนำภาพจำนวน 270 ภาพ ซึ่งแบ่งเป็นประเภทละ 30 ภาพ มาทดสอบกับแบบจำลองการจำแนกประเภทอาหาร ตัวอย่างผลการจำแนกประเภทอาหารจากภาพที่นำไปทดสอบ แสดงดังตารางที่ 4.13

ตารางที่ 4.13 ตัวอย่างผลการจำแนกประเภทอาหารจากภาพที่นำไปทดสอบ

| Class          | image                                                                              | Accuracy | Class           | image                                                                                 | Accuracy |
|----------------|------------------------------------------------------------------------------------|----------|-----------------|---------------------------------------------------------------------------------------|----------|
| americano_yuzu |   | 0.89     | grilled_salmon  |    | 0.72     |
| bubble_tea     |   | 0.80     | italian_soda    |    | 0.87     |
| crispy_fish    |   | 0.78     | korean_chicken  |    | 0.85     |
| egg_yolk       |  | 0.80     | pickled_fruit   |   | 0.90     |
|                |                                                                                    |          | ramen_ice_cream |  | 0.88     |

จากผลการวิจัยที่ได้จากการสร้างแบบจำลองการจำแนกประเภทบทวิจารณ์และจำแนกประเภทอาหารจากภาพ งานวิจัยนี้จึงได้แบบจำลองสำหรับการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์ ดังรูปที่ 4.4



รูปที่ 4.4 แบบจำลองสำหรับการแนะนำร้านอาหารจากความคิดเห็นบนเครือข่ายสังคมออนไลน์

## 4.5 อภิปรายผลการวิจัย

จากผลการวิจัยที่ได้จากการสร้างแบบจำลองการจำแนกประเภทบทวิจารณ์และจำแนกประเภทอาหารจากภาพ ทำให้สามารถอภิปรายผลการวิจัยโดยมีรายละเอียดดังต่อไปนี้

### 4.5.1 อภิปรายผลการวิจัยการสร้างแบบจำลองการจำแนกประเภทบทวิจารณ์

จากผลการวิเคราะห์โดยการจำแนกบทวิจารณ์ด้วยขั้นตอนวิธี Naive Bayes เปรียบเทียบกับขั้นตอนวิธี SVM ผลที่ได้พบว่า เป็นไปตามสมมติฐานการวิจัยในการจำแนกบทวิจารณ์ซึ่งมีค่าความถูกต้องในการจำแนกบทวิจารณ์มากกว่าร้อยละ 80 โดยขั้นตอนวิธี SVM ให้ค่าความถูกต้องในการทำนายร้อยละ 82.89 มากกว่าขั้นตอนวิธี Naive Bayes ที่ให้ค่าความถูกต้องในการทำนายร้อยละ 80.20 และเมื่อทดสอบกับข้อมูลใหม่ ขั้นตอนวิธี SVM ให้ค่าความถูกต้องในการทำนายร้อยละ 80.87 ส่วน Naive Bayes ให้ค่าความถูกต้องในการทำนายร้อยละ 78.86 ซึ่งสอดคล้องกับงานวิจัยก่อนหน้านี้ที่มีการเปรียบเทียบระหว่าง 2 ขั้นตอนวิธีดังกล่าว โดยการจำแนกบทวิจารณ์ด้วยขั้นตอนวิธี Naive Bayes และ SVM ทั้งสองขั้นตอนวิธีทำนายบทวิจารณ์ที่เป็นลบได้ถูกต้องมากที่สุด โดยระยะเวลาในการประมวลผลขั้นตอนวิธี SVM ใช้เวลาน้อยกว่า Naive Bayes

จากผลการทำนายที่ขั้นตอนวิธี SVM ทำนายถูกต้องมากกว่าขั้นตอนวิธี Naive Bayes อาจสืบเนื่องมาจากคุณลักษณะของแต่ละขั้นตอนวิธี โดย SVM มีการจำแนกข้อมูลออกเป็นมิติ ซึ่งเหมาะสมกับการจำแนกข้อมูลที่มีมิติมากกว่า Naive Bayes ที่เป็นขั้นตอนวิธีซึ่งใช้ความน่าจะเป็นในการทำนาย อีกทั้ง ข้อมูลที่ใช้ในทดสอบมีจำนวนน้อย จึงทำให้การวิเคราะห์ด้วย Naive Bayes ซึ่งเป็นตอนวิธีที่ทำนายสิ่งที่ไม่เคยเกิดขึ้นจากสิ่งที่เคยเกิดขึ้นมาก่อน และไม่ได้มีการทำนายในจำนวนที่เหมาะสม จึงได้ค่าความถูกต้องในการทำนายน้อยกว่า SVM

ทั้งนี้ จากผลการทำนายที่ผิดพลาด อาจเกิดจากจำนวนของบทวิจารณ์ โดยในอนาคตผู้วิจัยอาจเพิ่มจำนวนของข้อมูล และศึกษาขั้นตอนวิธีอื่น ๆ เพิ่มเติมมาเปรียบเทียบกัน เช่น ขั้นตอนวิธี Neural Network หรือการพัฒนาขั้นตอนวิธีในการตัดคำ เพื่อให้ได้ผลลัพธ์ในการทำนายที่ถูกต้องและตรงกับบทวิจารณ์ที่เป็นจริงมากขึ้น อีกทั้ง การนำผลที่ได้จากการวิเคราะห์ไปใช้สำหรับสร้างแบบจำลองในการแนะนำร้านอาหารแก่ผู้ใช้ในงานวิจัยครั้งต่อไป

นอกจากนี้ จากผลที่ได้ พบว่า บทวิจารณ์ที่เป็นกลางถูกทำนายว่าเป็นบทวิจารณ์ด้านบวกเป็นจำนวนมาก อาจเป็นผลมาจากบทวิจารณ์ด้านบวกและบทวิจารณ์ที่เป็นกลาง มีการใช้คำที่บ่งบอกความรู้สึกคล้ายคลึงกัน อีกทั้ง บทวิจารณ์ที่รวบรวมมาส่วนใหญ่เป็นบทวิจารณ์ด้านบวก จึงทำให้ผลการวิเคราะห์บทวิจารณ์ด้านบวกมีความแม่นยำในการทำนายมากกว่าบทวิจารณ์ด้านอื่น ๆ ซึ่งหากมีอัตราส่วนของบทวิจารณ์แต่ละด้านที่ใกล้เคียงกัน อาจทำให้ผลการวิเคราะห์ถูกต้องและแม่นยำมากยิ่งขึ้น



#### 4.5.2 อภิปรายผลการวิจัยการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF)

ในการวิเคราะห์หาความถี่ของคำ แชนแท็ก และสัญลักษณ์ พบว่า คำ แชนแท็ก และสัญลักษณ์อันดับแรก แสดงให้เห็นถึงความรู้สึของผู้วิจารณ์ที่มีต่ออาหาร หรือร้านอาหารได้ และผลที่ได้จากการวิเคราะห์ดังกล่าว สอดคล้องกับประเภทของบทวิจารณ์ โดยข้อค้นพบนี้ทำให้สามารถนำผลที่ได้ไปใช้ในการวิเคราะห์ข้อความในบทวิจารณ์ในงานวิจัยในอนาคต เพื่อจำแนกประเภทของบทวิจารณ์ว่าเป็นบทวิจารณ์ด้านบวก ด้านลบ หรือเป็นกลาง โดยแชนแท็กสามารถปรากฏในบทวิจารณ์ได้ทั้ง 3 ประเภท เนื่องจากเป็นแชนแท็กที่ได้รับความนิยมในขณะนั้น ซึ่งการใส่แชนแท็กลงในบทวิจารณ์บางครั้งอาจไม่เกี่ยวข้องกับประเภทของบทวิจารณ์ ซึ่งนอกจากคำที่ปรากฏในบทวิจารณ์แล้ว การวิเคราะห์จากสัญลักษณ์ในบทวิจารณ์เป็นอีกวิธีหนึ่งในการจำแนกประเภทบทวิจารณ์ได้ ซึ่งจะมีประโยชน์อย่างมากสำหรับงานวิจัยที่มีวัตถุประสงค์เพื่อการวิเคราะห์ความรู้สึกของผู้ใช้ที่มีต่ออาหาร และร้านอาหาร

#### 4.5.3 อภิปรายผลการวิจัยการสร้างแบบจำลองการจำแนกประเภทอาหารจากภาพ

จากผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ด้วยเครื่องมือ YOLO5 โดยแบ่งประเภทอาหารเป็น 9 ประเภท ได้แก่ อเมริกาโนยuzu (americano\_yuzu) ชานมไข่มุก (bubble\_tea) ข้าวปลาแกะ (crispy\_fish) ไข่ไข่แดงเค็มหมูยอ (egg\_yolk) แซลมอนเสียบไม้ย่าง (grilled\_salmon) สปาร์กิ้งโซดา (italian\_soda) ไก่เกาหลี (korean\_chicken) ผลไม้ดอง (pickled\_fruit) และราเมนไอศกรีม (ramen\_ice\_cream) ประเภทละ 120 ภาพ รวมทั้งหมด 1,080 ภาพ จำนวน 1,228 อินสแตนซ์ ผลการวิจัยที่ได้พบว่า มีค่าความแม่นยำ (Precision) โดยรวมร้อยละ 94 ค่าความระลึก (Recall) ร้อยละ 96 ค่าความถูกต้อง (Accuracy) ร้อยละ 98 และค่าวัดประสิทธิภาพ (F-Measure) ร้อยละ 95 ซึ่งผลการจำแนกประเภทอาหารเพื่อสร้างแบบจำลองของอาหารประเภทไข่แดงเค็มหมูยอ มีค่าความแม่นยำร้อยละ 98 ค่าความระลึกและค่าความถูกต้องร้อยละ 100 ซึ่งมีค่าความระลึกและค่าความถูกต้องในการจำแนกประเภทอาหารสูงที่สุด ส่วนไก่เกาหลี ซึ่งมีค่าความถูกต้องในการทำนายสูงที่สุดเป็นร้อยละ 100 มีค่าความแม่นยำ ร้อยละ 98 และค่าความระลึกร้อยละ 99 ส่วนราเมนไอศกรีม มีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุด

จากผลการวิจัยที่ได้ สามารถอภิปรายได้ว่า การสร้างแบบจำลองการจำแนกประเภทอาหารเป็นไปตามสมมติฐานการวิจัยเช่นกัน ซึ่งมีค่าความถูกต้องในการจำแนกประเภทอาหารมากกว่าร้อยละ 80 โดยการใช้ภาพอาหารที่มีจำนวนเท่ากันหรือใกล้เคียง และการแบ่งประเภทอาหารบนภาพที่ชัดเจน เป็นผลให้ผลการวิเคราะห์มีค่าความแม่นยำและค่าความระลึกสูง ซึ่งเมื่อ

วิเคราะห์ผลการจำแนกภาพของอาหารแต่ละประเภทพบว่า ไข่แดงเค็มหมูยอ มีค่าความแม่นยำร้อยละ 98 และค่าความระลึกร้อยละ 100 ซึ่งมีค่าความระลึกในการจำแนกประเภทอาหารสูงที่สุด ส่วนราเมนไอศกรีม มีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุดนั้น อาจเป็นผลมาจากการใช้ภาพที่มีความคล้ายคลึงกันในภาพของอาหารประเภทนั้น ๆ และการกำหนดจำนวนครั้ง (Epochs) ในการสร้างแบบจำลอง (Training) เป็นจำนวนมาก จึงทำให้มีการเรียนรู้รายละเอียดภายในภาพ ซึ่งส่งผลให้มีความแม่นยำและค่าความระลึกสูงเช่นกัน

ทั้งนี้ จำนวนอินสแตนซ์ที่ปรากฏบนภาพสำหรับใช้จำแนกประเภทอาหารนั้น ปรากฏบนภาพที่ต่างประเภทกัน จึงทำให้สามารถจำแนกประเภทของอาหารได้ชัดเจนยิ่งขึ้น แต่อย่างไรก็ตาม จากภาพที่นำไปใช้ในการทดสอบแบบจำลองนั้น ภาพบางภาพที่ทดสอบยังให้ค่าความถูกต้องในการทำนายไม่ถึงร้อยละ 80 อาจเป็นผลมาจากการใช้ภาพที่นำไปสร้างแบบจำลองมีจำนวนน้อย และภาพที่นำไปทดสอบแบบจำลองเป็นภาพที่มีรายละเอียดภายในภาพที่แตกต่างกับภาพที่ใช้ในการสร้างแบบจำลอง จึงทำให้บางภาพที่นำไปทดสอบยังไม่ได้ผลการทดสอบที่มากกว่าร้อยละ 80 อย่างไรก็ตาม การทดสอบภาพกับแบบจำลองเป็นการทดสอบเพื่อทำนายประเภทของอาหารบนภาพเพื่อระบุประเภทอาหารที่ร้านอาหารจำหน่ายเท่านั้น ซึ่งงานวิจัยนี้ได้นำผลการจำแนกประเภทอาหารจากภาพไปใช้เพื่อนำเสนอข้อมูลเกี่ยวกับอาหารที่ร้านอาหารจำหน่าย เพื่อให้สามารถค้นหาร้านอาหารได้ตรงตามความต้องการของผู้ใช้ โดยตัวอย่างการนำเสนอข้อมูลร้านอาหารเพื่อนำร้านอาหารแก่ผู้ใช้แสดงดังรูปที่ 4.5



Show 10 entries Search:

| NO | Text Review                                                                                                                                                                                                                                                                         | Image detected | Review output | Image url                                                                             | Location                 |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|---------------|---------------------------------------------------------------------------------------|--------------------------|
| 1  | บอนซอนแพงมากจริงๆ ไม่ได้เซอร์วิสไรเลยด้วยซ้ำ แต่เท่าที่บอกอาหารมาเสิร์ฟ ออเดิร์ฟบริการเร็วเว่ เทคแคร์ดีสุดๆ หรกว่าอีก ยังคิดแค่10% ไม่เก็บบอนซอนเหมือนกัน บันแต่กินเพราะมันอร่อย แต่คือมันรู้สึกว่าย่างแพงเกินไปนะ                                                                  | โกเกาหลี่      | neg           |    | <a href="#">Location</a> |
| 2  | ร้านมะขิดัง ยูเนียนมอลล์ ชั้น 4 เป็นร้านเกาหลีที่ได้ดีมาก เกาหลีแท้แน่นอนรสชาติอร่อยมาก อร่อยทุกอย่างจริงๆ แนะนำหมีคิมจิคือดี โกเกาหลี่ ร้านนี้ราคาไม่แรงมากนะคะ ฟังตัวนะ @aroi #สองอำมมาป่าของกิน #รีวิวเกาหลี #อร่อยไปแดก #อร่อยจนต้องรีวิว #อร่อยต้องลอง #อร่อย                  | โกเกาหลี่      | pos           |    | <a href="#">Location</a> |
| 3  | ชาไข่มุก 19 บาท ที่ซอยจินดา ม.เทคโนโลยีพระจอมเกล้ากระบี่ อร่อย สิ่งหวานน้อยได้ ชามหอม 19 บ. คือชาเขียว ถ้าเพิ่มไข่มุกก็เป็น 24 บ. แก้วใหญ่ กินหมดก็จุอะ 5555 ก็มีคอกะเคี้ยว กัดกรังเรียก ละจะมีให้รับโทรศัพท์สั่งชาจ่ายดังที่ช่องบน และรับชาที่ช่องล่าง #อร่อยไปแดก #รีวิวลาดกระบัง | ชาไข่มุก       | pos           |    | <a href="#">Location</a> |
| 4  | KAMU fresh milk with brown sugar bubble เมนูอะไรก็ได้ 80บาท อร่อยมากกกกกกกกกก อร่อยกว่าร้านกวางแบบไม่ต้องต่อคิว นมหอมและไข่มุกหวาน อร่อย กาลังดี ผสมกันคือดีมาก ราคาจับได้ไม่แะร่อย ชอบมาก เจอร้านที่เมนูนี้ถูกใจละ สุดท้ายกลับมาดัยังคามู 10/10 จะกินบ่อยๆ #มขพรีวิว               | ชาไข่มุก       | pos           |   | <a href="#">Location</a> |
| 5  | เมื่อวานเห็นคนญี่ปุ่นเอาไอติมใส่ราเม็ง วันนี้เจอเลย์เอาต์ Kfc+อะจิงราเม็ง                                                                                                                                                                                                           | ราเม็ง         | pos           |  | <a href="#">Location</a> |

รูปที่ 4.5 ตัวอย่างการนำเสนอข้อมูลร้านอาหารเพื่อแนะนำร้านอาหารแก่ผู้ใช้



## บทที่ 5

### สรุปผลการวิจัย

จากผลการวิจัยที่ได้จากการวิเคราะห์ข้อมูล ในงานวิจัยเรื่อง การพัฒนาระบบแนะนำร้านอาหารจากความเห็นบนเครือข่ายสังคมออนไลน์ ประกอบด้วยผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes ผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM สำหรับจำแนกประเภทบทวิจารณ์ ผลการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ แสขแท้ก และสัญลักษณ์ที่ปรากฏในบทวิจารณ์ และผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพด้วยโวลี เวอร์ชัน 5 (YOLOv5) สามารถสรุปผลการวิจัยได้ดังหัวข้อดังต่อไปนี้

- 5.1 สรุปผลการวิจัย
  - 5.2 ข้อจำกัดของงานวิจัย
  - 5.3 การประยุกต์ผลการวิจัย
  - 5.4 ข้อเสนอแนะงานวิจัยครั้งต่อไป
- โดยในแต่ละหัวข้อ มีรายละเอียดดังต่อไปนี้

#### 5.1 สรุปผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบการแนะนำร้านอาหารจากความเห็นบนเครือข่ายสังคมออนไลน์ โดยรวบรวมข้อมูลความคิดเห็นประกอบด้วยบทวิจารณ์และภาพเกี่ยวกับอาหารจากเครือข่ายสังคมออนไลน์ เพื่อนำมาการจำแนกประเภทบทวิจารณ์และจำแนกประเภทอาหารบนภาพ ประกอบด้วยผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes เปรียบเทียบกับผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM ซึ่งในการจำแนกประเภทบทวิจารณ์นั้นได้จำแนกประเภทบทวิจารณ์ออกเป็น 3 ประเภท ได้แก่ บทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ จำนวน 1,490 บทวิจารณ์ แบ่งข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองเป็นร้อยละ 80 (1,192 บทวิจารณ์) และ 20 (298 บทวิจารณ์) ข้อมูลที่ใช้ทดสอบแบ่งเป็นด้านบวก เป็นกลาง และด้านลบ จำนวน 212 50 และ 36 บทวิจารณ์ตามลำดับ และมีการเปรียบเทียบเวลาที่ใช้ในการสร้างและทดสอบแบบจำลองในการจำแนกประเภทบทวิจารณ์ของทั้งสองขั้นตอนวิธี นอกจากนี้ได้มีการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับบทวิจารณ์ ประกอบด้วยการวิเคราะห์ความถี่ (TF-IDF) ของคำ แสขแท้ก และสัญลักษณ์ที่ปรากฏในบทวิจารณ์ เพื่อใช้ระบุประเภทของบทวิจารณ์

ส่วนการจำแนกประเภทอาหารบนภาพ งานวิจัยนี้ใช้วิธีการจำแนกประเภทอาหารด้วยวิธีการประมวลผลภาพด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ด้วยเครื่องมือ YOLO5 ได้จำแนกประเภทอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกันโน้ชซู ชานามไข่มุก ข้าวปลาแกะ ข้าไข่แดงเค็มหมุยอ แซลมอนเสียบไม้ย่าง สปราร์กลิ่งโซดา ไข่เกาหลี ผลไม้ดอง และราเมนไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นประเภทละ 120 ภาพ ซึ่งกำหนดจำนวนครั้ง (Epochs) ในการเรียนรู้รายละเอียดบนภาพเพื่อสร้างแบบจำลอง (Training) จำนวน 50 ครั้ง และใช้ภาพสำหรับทดสอบแบบจำลองจำนวน 270 ภาพ แบ่งเป็นประเภทละ 30 ภาพ ซึ่งเมื่อมีการวิเคราะห์จำแนกบทวิจารณ์และประเภทอาหารบนภาพ งานวิจัยนี้สามารถสรุปผลการวิจัยออกเป็น 3 ส่วน ดังนี้

### ส่วนที่ 1 ผลการวิจัยการจำแนกประเภทบทวิจารณ์

ในการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทบทวิจารณ์ งานวิจัยนี้ได้วิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes เปรียบเทียบกับการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี SVM โดยจำแนกประเภทบทวิจารณ์ออกเป็น 3 ประเภท ได้แก่ บทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ จำนวน 1,490 บทวิจารณ์ แบ่งข้อมูลที่ใช้ในการสร้างแบบจำลองร้อยละ 80 และทดสอบแบบจำลองเป็นร้อยละ 20 ซึ่งผลที่ได้จากการวิเคราะห์ดังกล่าวพบว่า ขั้นตอนวิธี SVM ให้ผลลัพธ์ในการจำแนกประเภทบทวิจารณ์ที่มีค่าความถูกต้องมากกว่าขั้นตอนวิธี Naive Bayes โดยผลการวิจัยสามารถสรุปได้ดังตารางที่ 5.1

ตารางที่ 5.1 สรุปผลการวิจัยขั้นตอนวิธี SVM และ Naive Bayes

| ผลการวิเคราะห์/ขั้นตอนวิธี | SVM   | Naive Bayes |
|----------------------------|-------|-------------|
| Accuracy                   | 82.89 | 80.20       |
| Weighted Average Precision | 82.00 | 78.00       |
| Weighted Average Recall    | 83.00 | 80.00       |
| F-Measure                  | 82.00 | 78.00       |
| Time (S)                   | 1.38  | 1.51        |

จากสรุปผลการวิจัยพบว่า ขั้นตอนวิธี SVM มีค่าความถูกต้องในการทำนายร้อยละ 82.89 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 82.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก ร้อยละ 83.00 และค่าวัดประสิทธิภาพ ร้อยละ 82.00 ใช้เวลาในการประมวลผล 1.38 วินาที ส่วนผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes มีค่าความถูกต้องในการทำนายร้อยละ 80.20 ค่าความแม่นยำเฉลี่ยถ่วง

น้ำหนักร้อยละ 78.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนักร้อยละ 80.00 และค่าวัดประสิทธิภาพร้อยละ 78.00 ใช้เวลาในการประมวลผล 1.51 วินาที

เมื่อนำข้อมูลใหม่ (Unseen Data) มาทดสอบแบบจำลอง จำนวน 298 บทความประกอบด้วยบทความด้านบวก เป็นกลาง และด้านลบ จำนวน 212 50 และ 36 บทความตามลำดับ ผลการวิเคราะห์ที่ได้แสดงดังตารางที่ 5.2

ตารางที่ 5.2 สรุปผลการวิจัยขั้นตอนวิธี SVM และ Naive Bayes กับข้อมูลใหม่

| ผลการวิเคราะห์/ขั้นตอนวิธี | SVM   | Naive Bayes |
|----------------------------|-------|-------------|
| Accuracy                   | 80.87 | 78.86       |
| Weighted Average Precision | 80.00 | 77.00       |
| Weighted Average Recall    | 81.00 | 79.00       |
| F-Measure                  | 80.00 | 76.00       |
| Time (S)                   | 1.43  | 1.54        |

จากสรุปผลการวิเคราะห์กับข้อมูลใหม่พบว่า ขั้นตอนวิธี SVM ให้ผลลัพธ์ในการจำแนกประเภทบทความที่มีค่าความถูกต้องมากกว่าขั้นตอนวิธี Naive Bayes เช่นเดียวกัน โดยขั้นตอนวิธี SVM มีค่าความถูกต้องในการทำร้อยละ 80.87 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนัก ร้อยละ 80.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก ร้อยละ 81.00 และค่าวัดประสิทธิภาพ ร้อยละ 80.00 ใช้เวลาในการประมวลผล 1.43 วินาที ส่วนผลการวิเคราะห์ข้อมูลด้วยขั้นตอนวิธี Naive Bayes ค่าความถูกต้องในการทำร้อยละ 78.86 ค่าความแม่นยำเฉลี่ยถ่วงน้ำหนักร้อยละ 77.00 ค่าความระลึกเฉลี่ยถ่วงน้ำหนักร้อยละ 79.00 และค่าวัดประสิทธิภาพร้อยละ 76.00 ใช้ระยะเวลาในการประมวลผล 1.54 วินาที

## ส่วนที่ 2 ผลการวิจัยการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF)

ในการวิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) งานวิจัยนี้ได้วิเคราะห์ข้อมูลโดยการนับความถี่ (TF-IDF) ของคำ แชชแท็ก และสัญลักษณ์ต่าง ๆ โดยจัดอันดับ 3 อันดับแรก เพื่อนำไปใช้ในการระบุประเภทของบทความ โดยผลการวิเคราะห์ความถี่คำพบว่า อันดับแรกของคำในบทความด้านบวก คือ อร่อย มีความถี่ร้อยละ 1.85 ส่วนบทความที่เป็นเป็นกลาง คำที่ปรากฏเป็นอันดับแรกคือ กินได้ มีความถี่ร้อยละ 1.71 และคำในบทความด้านลบเป็นอันดับแรกคือ เสีย ซึ่งเป็นคำสบถ มีความถี่ร้อยละ 3.61 ส่วนผลการวิเคราะห์ความถี่ของแฮชแท็กในบทความด้านบวกและบทความที่เป็นกลางมีความถี่เป็นอันดับแรกคือ #อร่อยไปแดก มีความถี่ร้อยละ 12.51 และ 2.83 ส่วนแฮชแท็กที่ปรากฏในบทความที่ด้านลบที่มีความถี่เป็นอันดับแรกคือ #ไม่อร่อยอย่า

แตก มีความถี่ร้อยละ 2.10 และผลการวิเคราะห์ความถี่สัญลักษณ์ พบว่าสัญลักษณ์อันดับแรกที่เป็น  
ปรากฏในบทวิจารณ์ด้านบวกและด้านลบคือ 😊 มีความถี่ร้อยละ 2.32 และ 2.52 ส่วนบทวิจารณ์ที่เป็น  
กลางมีสัญลักษณ์ 😐 อยู่ในอันดับแรก โดยมีความถี่ร้อยละ 1.60

### ส่วนที่ 3 ผลการวิจัยการวิเคราะห์การจำแนกประเภทอาหารบนภาพ

ในการจำแนกประเภทอาหารบนภาพ งานวิจัยนี้ใช้วิธีการประมวลผลภาพด้วยวิธีโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ด้วยเครื่องมือ YOLO5 เพื่อสร้างแบบจำลองการจำแนกประเภทอาหาร โดยแบ่งคลาสอาหารออกเป็น 9 คลาส ได้แก่ อเมริกาโน ยูซุ สปร้ากลิ่ง โซดา ชานม ไข่มุก ข้าวปลา เกะ ยำไข่แดง เค็ม หมูยอ แซลมอน เสียบ ไม้ย่าง ผลไม้ดอง ไก่เกาหลี และราเมน ไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นคลาสละ 120 ภาพ จำนวน 1,228 อินสแตนซ์ และทดสอบกับภาพที่เป็นข้อมูลทดสอบเพื่อทดสอบแบบจำลอง จำนวน 270 ภาพ แบ่งเป็นคลาสละ 30 ภาพ โดยผลที่ได้จากการสร้างแบบจำลองพบว่า 96 ค่าความถูกต้อง (Accuracy) ร้อยละ 98 และค่าวัดประสิทธิภาพ (F-Measure) ร้อยละ 95 ซึ่งผลการจำแนกประเภทอาหารเพื่อสร้างแบบจำลองของอาหารประเภทไข่แดง เค็ม หมูยอ มีค่าความแม่นยำร้อยละ 98 ค่าความระลึก และค่าความถูกต้องร้อยละ 100 ซึ่งมีค่าความระลึกและค่าความถูกต้องในการจำแนกประเภทอาหารสูงที่สุด ส่วนไก่เกาหลี ซึ่งมีค่าความถูกต้องในการทำนายสูงสุดเป็นร้อยละ 100 มีค่าความแม่นยำร้อยละ 98 และค่าความระลึกร้อยละ 99 ส่วนราเมน ไอศกรีม มีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุด

ทั้งนี้ ผลการสร้างแบบจำลองเพื่อจำแนกประเภทอาหารบนภาพของอาหารประเภทอื่นพบว่า อเมริกาโนยูซุ มีค่าความแม่นยำร้อยละ 91 และค่าความระลึกร้อยละ 99 สปรักลิ่งโซดา มีค่าความแม่นยำร้อยละ 95 และค่าความระลึกร้อยละ 98 ซานมไป่มุก มีค่าความแม่นยำร้อยละ 88 และค่าความระลึกร้อยละ 99 ข้าวปลาแกะ มีค่าความแม่นยำและค่าความระลึกร้อยละ แชลมอนเสียบไม้ย่าง มีค่าความแม่นยำร้อยละ 89 และค่าความระลึกร้อยละ 82 ซึ่งมีค่าความแม่นยำและค่าความระลึกในการทำนายประเภทอาหารน้อยที่สุด ส่วนผลไม้ดอง มีค่าความแม่นยำร้อยละ 93 และค่าความระลึกร้อยละ 97 และไก่เกาหลี มีค่าความแม่นยำร้อยละ 98 และค่าความระลึกร้อยละ 99



## 5.2 ข้อจำกัดของงานวิจัย

ในการนำข้อมูลบทวิจารณ์และภาพจากเครือข่ายสังคมออนไลน์มาเพื่อวิเคราะห์เพื่อจำแนกประเภทบทวิจารณ์และประเภทอาหารนั้น งานวิจัยนี้ได้รวบรวมข้อมูลดังกล่าวจากเครือข่ายสังคมออนไลน์ทวิตเตอร์ ซึ่งเป็นเครือข่ายสังคมออนไลน์ที่ได้รับความนิยมในการใช้งานของผู้ใช้ มักจะมีการทวีตข้อความและอัปโหลดภาพที่เกี่ยวข้องกับอาหารหรือร้านอาหาร ซึ่งได้รวบรวมข้อมูลในเดือนสิงหาคม-ตุลาคม พ.ศ. 2562 โดยรวบรวมข้อมูลแบบออนไลน์ แล้วนำมาวิเคราะห์ข้อมูลแบบออฟไลน์ ซึ่งในขณะนั้นทวิตเตอร์มีการเปิดให้ผู้พัฒนาสามารถใช้งาน Twitter API ได้โดยไม่มีค่าใช้จ่าย แต่ในปัจจุบัน ได้มีการปรับให้มีค่าใช้จ่ายในการใช้งาน Twitter API จึงทำให้ผู้วิจัยไม่สามารถรวบรวมข้อมูลสำหรับการสร้างแบบจำลองการจำแนกประเภทบทวิจารณ์และประเภทอาหารได้แบบเรียลไทม์ ดังนั้นข้อมูลที่นำมาวิเคราะห์จึงเป็นข้อมูลที่มีการรวบรวมเอาไว้ก่อน ซึ่งจะนำไปใช้สร้างแบบจำลองและทดสอบแบบจำลองแบบออฟไลน์เท่านั้น โดยในขั้นตอนของการตัดคำที่นำมาใช้ในงานวิจัยนี้ เป็นการตัดคำแบบที่ตรงตามหลักไวยากรณ์ทางภาษาไทยเท่านั้น ข้อความที่ไม่ถูกต้องตามหลักไวยากรณ์จะไม่ถูกนำมาประมวลผล

ทั้งนี้ ในการรวบรวมข้อมูลบทวิจารณ์และภาพจากทวิตเตอร์นั้น มีการดึงข้อมูลภาพที่เป็นภาพอาหารที่เกี่ยวข้องกับข้อความ แต่ภาพที่นำมาใช้งานวิจัย เป็นภาพที่ถูกอัปโหลดขึ้นบนทวิตเตอร์ ซึ่งภาพดังกล่าวไม่สามารถสกัดข้อมูลที่เป็นตำแหน่งของร้านอาหารได้ จึงทำให้ภาพที่รวบรวมมาใช้ในการวิเคราะห์นั้น ไม่มีพิกัดละติจูดและลองจิจูดเพื่อระบุตำแหน่งของร้านอาหาร ดังนั้น ผู้วิจัยจึงถ่ายรูปอาหารที่เกี่ยวข้องกับบทวิจารณ์ เพื่อนำมาหาตำแหน่งที่ตั้งของร้านอาหารสำหรับแนะนำเส้นทางให้กับผู้ใช้ที่ต้องการค้นหาเส้นทางเพื่อเดินทางไปยังร้านอาหารที่ต้องการ โดยเชื่อมโยงการแนะนำเส้นทางโดยใช้แผนที่จากกูเกิ้ล (Google Map) ซึ่งงานวิจัยนี้เพียงระบุตำแหน่งของร้านอาหารจากการใส่เส้นทางในลิงก์เพื่อให้ผู้ใช้ใช้งานได้เท่านั้น ไม่สามารถนำมาพัฒนาเป็นระบบ เนื่องจากมีข้อจำกัดเรื่องค่าใช้จ่ายในการใช้งาน Google Map API

นอกจากนี้ ในการนำเสนอข้อมูลร้านอาหารแก่ผู้ใช้ งานวิจัยนี้ได้นำเสนอบทวิจารณ์และภาพที่เกี่ยวข้องกับอาหาร โดยผู้ใช้สามารถค้นหาจากบทวิจารณ์ ประเภทของบทวิจารณ์ หรือชื่อต่าง ๆ ที่เป็นข้อความได้เท่านั้น ไม่สามารถค้นหาร้านอาหารที่ต้องการจากภาพได้ ซึ่งเป็นอีกข้อจำกัดหนึ่งในงานวิจัยนี้ ผู้วิจัยได้นำเสนอข้อมูลโดยแสดงภาพอาหารที่เกี่ยวข้องกับร้านอาหารหรือบทวิจารณ์ อีกทั้ง การระบุประเภทอาหารในงานวิจัยนี้ ใช้การระบุประเภทอาหารจากการจำแนกประเภทอาหารบนภาพเท่านั้น ไม่ได้มีการระบุประเภทอาหารจากการวิเคราะห์ข้อความที่อยู่ในบทวิจารณ์เพื่อหาชื่ออาหารหรือร้านอาหาร ซึ่งบทวิจารณ์เกี่ยวกับอาหารนำมาใช้ในการวิเคราะห์หาประเภทของบทวิจารณ์ซึ่งเป็นด้านบวก ด้านเป็นกลาง และด้านลบเท่านั้น



### 5.3 การประยุกต์ผลการวิจัย

งานวิจัยนี้เป็นการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทบทวิจารณ์ประกอบด้วยบทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ ซึ่งจะเป็นประโยชน์ต่องานวิจัยที่มีการวิเคราะห์ความรู้สึก หรือการทำเหมืองความคิดเห็น และสามารถนำความถี่คำ แชนแท็ก และสัญลักษณ์ไปใช้ระบุประเภทบทวิจารณ์ หรือใช้อ้างอิงสำหรับการระบุความรู้สึกด้านต่าง ๆ นอกจากนี้ สามารถนำผลที่ได้จากแบบจำลองไปพัฒนาเป็นระบบสำหรับการวิเคราะห์ความรู้สึก จากบทวิจารณ์บนเครือข่ายสังคมออนไลน์ รวมถึงการนำไปใช้บนแพลตฟอร์มอื่น ๆ เช่น โทรศัพท์มือถือ โดยสามารถดึงข้อมูลที่เป็นข้อความจากเครือข่ายสังคมออนไลน์ แล้วนำมาทดสอบกับแบบจำลองที่ได้จากงานวิจัยนี้ ส่วนของการวิเคราะห์เพื่อจำแนกประเภทอาหารบนภาพ สามารถนำไปต่อยอดการประมวลผลภาพเพื่อดูรายละเอียดที่ปรากฏบนภาพนอกเหนือจากการจำแนกประเภทอาหาร ในงานวิจัยนี้ เช่น การจำแนกอาหารคาว อาหารหวาน เป็นต้น

### 5.4 ข้อเสนอแนะงานวิจัยครั้งต่อไป

ในการวิจัยครั้งนี้เป็นการวิเคราะห์ข้อมูลเพื่อจำแนกประเภทบทวิจารณ์และประเภทอาหารบนภาพ ซึ่งได้แบ่งประเภทบทวิจารณ์ออกเป็น 3 ด้าน ได้แก่ บทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ ซึ่งในการวิจัยครั้งต่อไปอาจเพิ่มประเภทบทวิจารณ์ให้มีความละเอียดมากขึ้น เพื่อให้สามารถวิเคราะห์ความรู้สึกได้ละเอียดมากกว่า 3 ระดับ โดยบทวิจารณ์ที่ใช้ในงานวิจัยนี้มีจำนวน 1,492 บทวิจารณ์ ซึ่งในอนาคตอาจต้องเพิ่มจำนวนบทวิจารณ์ หรือรวบรวมบทวิจารณ์จากเครือข่ายสังคมออนไลน์อื่น ๆ เช่น เฟซบุ๊ก เพื่อให้มีความหลากหลายของบทวิจารณ์ที่เกี่ยวข้องกับอาหาร หรือร้านอาหารมากขึ้น นอกจากนี้ ผลที่ได้จากการวิเคราะห์ความถี่ของคำ แชนแท็ก และสัญลักษณ์ จะถูกนำไปใช้ระบุประเภทบทวิจารณ์ หรือใช้อ้างอิงในงานวิจัยครั้งต่อไปสำหรับการวิเคราะห์ข้อมูลเพิ่มเติมเกี่ยวกับการวิเคราะห์ความรู้สึกจากบทวิจารณ์

ส่วนการวิเคราะห์เพื่อจำแนกประเภทอาหารบนภาพนั้น อาจมีการนำขั้นตอนวิธีอื่น ๆ มาใช้ในการวิเคราะห์ นอกเหนือจากโครงข่ายแบบสังวัตนาการ และเปรียบเทียบความแตกต่างของผลการวิเคราะห์ เพื่อให้ได้ขั้นตอนวิธีที่ให้ค่าความถูกต้องในการทำนายมากที่สุด ทั้งนี้ การจำแนกประเภทอาหารในงานวิจัยครั้งต่อไป อาจจำแนกประเภทอาหารครอบคลุมไปถึงอาหารที่เป็นอาหารท้องถิ่น หรืออาหารที่ได้รับความนิยมในขณะนั้น เพื่อนำเสนอข้อมูลแก่ผู้ใช้ได้อย่างถูกต้องและทันต่อยุคสมัย ซึ่งการพัฒนาการนำเสนอข้อมูล อาจพัฒนาเป็นระบบที่สมบูรณ์ ซึ่งเป็นระบบที่รวบรวมทั้งบทวิจารณ์ที่ผ่านการวิเคราะห์เพื่อจำแนกประเภทบทวิจารณ์ที่มีประเภทบทวิจารณ์มากกว่า 3 ระดับ ร่วมกับประเภทของอาหารที่หลากหลาย อีกทั้ง ยังเป็นระบบที่สามารถค้นหาข้อมูลได้ทั้งจากข้อความและภาพ เพื่อให้ครอบคลุมกับความต้องการของผู้ใช้ ณ ขณะนั้นมากยิ่งขึ้น

## รายการอ้างอิง

- Ch3Thailand. (2561). **เปรี้ยวปาก เช็คอิน ฉลองยิ่งใหญ่ 14 ปี!จัดงาน‘PREAWPAK FESTIVAL 2018’**. [ออนไลน์]. ได้จาก: <http://prd.ch3thailand.com/news/program/11271>.
- CleverKid. (2558). **eat share connect** ไลฟ์สไตล์มาแรงยุค ชิงเกิล แอปฯ. [ออนไลน์]. ได้จาก: <https://ireview.in.th/pr-eat-share-connect-lifestyle-force/>.
- Google Trend. (2017). **Trending How-to**. [ออนไลน์]. ได้จาก: <https://trends.google.com/trends/topcharts#vm=trendingchart&cid=3283692f-1701-47df-9ff0-57664dde073a&geo=TH&date=2017&cat=>.
- Google Trend. (2017). **Trending การซื้อสินค้า**. [ออนไลน์]. ได้จาก: <https://trends.google.com/trends/topcharts#vm=trendingchart&cid=432908e4-2566-4d47-b841-eb1ca2172832&geo=TH&date=2017&cat=>.
- iPick. (2015). **“eat • share • connect”** ไลฟ์สไตล์มาแรงยุค “ชิงเกิล แอปฯ”. [Online]. ได้จาก: <https://www.dailygizmo.tv/2015/12/04/eat-share-connect-lifestyle-ipick-app-stat/>.
- Marketeer. (2560). **5 เทรนด์โลกที่ส่งผลต่อผู้บริโภคไทยในยุค 4.0**. [ออนไลน์]. ได้จาก: <http://marketeer.co.th/archives/122620>.
- Positioningmag. (2560). **หลุดฝั่งอีกรายการ ตลาดสดสนามเป้า ย้ายจากช่อง 5 ไปอยู่ช่อง 3 แล้ว**. [ออนไลน์]. ได้จาก: <https://positioningmag.com/1103942>.
- Thai SME Research. (2558). **Social media** ส่งผลอย่างไรต่อการตัดสินใจซื้อสินค้า. [ออนไลน์]. ได้จาก: <http://www.thaismeresearch.com/social-media-influence-consumer-decision-sme/>.
- Thoth Zocial Obivoc. (2561). **พฤติกรรมการใช้งาน Social Media ของคนไทยปี 2017**. [ออนไลน์]. ได้จาก: <https://www.rainmaker.in.th/interesting-things-in-2017-by-zocial-award-18/>.
- Wongnai. (2565). **9 เมนูฮิตติดกระแส ขายอะไรดี ปี 2022 ทำง่าย ดันเคลิเวอร์พุ่ง**. [ออนไลน์]. ได้จาก: <https://www.wongnai.com/business-owners/social-media-viral-food?>.
- ไทยรัฐ. (2555). **ตำนาน 50 ปี เซลล์ชวนชิม**. [ออนไลน์]. ได้จาก: <https://www.thairath.co.th/content/241098>.
- ไทยรัฐ. (2560). **ไม่กินถือว่าผิด**. [ออนไลน์]. ได้จาก: <https://www.thairath.co.th/content/1154075>.

กฤษฎา โสมนะพันธุ์ และกิตติพันธ์ คงสวัสดิ์เกียรติ. (2556). แรงจูงใจในการรับประทานอาหารไทยนอกบ้านของผู้บริโภคในเขตกรุงเทพมหานคร และปริมณฑล. วารสารการเงิน การลงทุน การตลาด และการบริหารธุรกิจ. 3(4): 695-714.

กานดา รุณนะพงศา และปโยธร อูราธรรมกุล. (2548). การตัดคำภาษาไทยโดยการปรับปรุงปรองกฎและพจนานุกรมแบบใหม่. รายงานวิจัยฉบับสมบูรณ์ คณะวิศวกรรมศาสตร์, มหาวิทยาลัยขอนแก่น.

บุญยาพร วุฒิธรรมคุณ. (2559). การรีวิจร้านอาหารผ่านสื่อสังคมออนไลน์กับการตัดสินใจเลือกร้านอาหารสำหรับการรับประทานนอกบ้านของผู้บริโภคในเขตกรุงเทพมหานคร. การค้นคว้าอิสระ หลักสูตรนิเทศศาสตรมหาบัณฑิต, มหาวิทยาลัยกรุงเทพ.

มติชน สุดสัปดาห์. (2561). ถนัดศรีชวนชิม. [ออนไลน์]. ได้จาก: <https://www.matichonweekly.com/columnist/ถนัดศรีชวนชิม>.

รัชนิ พันธุ์รุ่งจิตติ และคณะ. (2560). ว้าว! Fashion Food แชะ ชิม แซร์. นิตยสาร SME Start Up. 15: 18-19.

วรรณพงษ์ ภัททิย์ไพบูลย์. (2559). คู่มือการใช้งาน PyThaiNLP 1.4. [ออนไลน์]. ได้จาก: <https://pythainlp.readthedocs.io/en/latest/pythainlp-1-4-thai/>.

วรัญญู อินทรกำแหง และพฤกัทร ทรงเที่ยง. (2560). ยังจำได้ไหม จำได้หรือเปล่า...หวนคืนวันวาน สัญลักษณ์ความอร่อยก่อนดาวมิชลิน. [ออนไลน์]. ได้จาก: <https://thestandard.co/delicious-symbol/>.

วิชา ทองลัพท์. (2559). ปัจจัยที่มีผลต่อความพึงพอใจในการใช้บริการร้านอาหารประเภทขาของ ผู้ใช้บริการในเขตกรุงเทพมหานครและปริมณฑล. การค้นคว้าอิสระ หลักสูตรบริหารธุรกิจมหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์.

ศาสตรา วงศ์ชนวสุ. (2560). ปัญญาประดิษฐ์ ทฤษฎี โปรแกรม และการประยุกต์. สำนักพิมพ์มหาวิทยาลัยขอนแก่น.

ศูนย์วิจัยกสิกรไทย. (2560). 5 เทรนด์ธุรกิจ แรงแข่งขันปี 2017. ธนาคารกสิกรไทย.

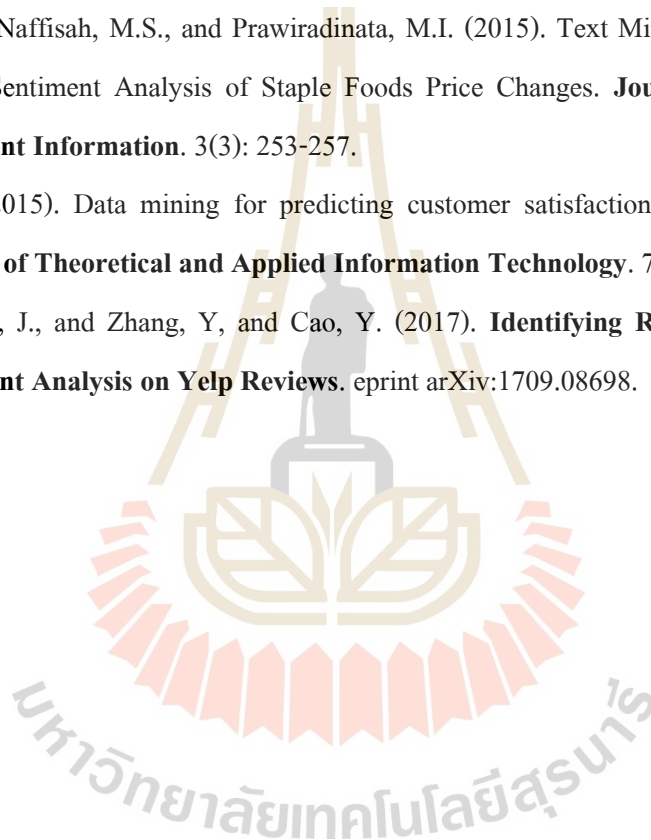
อนุภาค ปัญญาทาสสมบัติ. (2564). ระบบการตรวจสอบองค์ประกอบของฮาร์ดดิสก์ไดรฟ์ในกระบวนการประกอบฮาร์ดดิสก์ไดรฟ์อัตโนมัติด้วยเทคนิคการเรียนรู้เชิงลึก. วิทยานิพนธ์หลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต, มหาวิทยาลัยเทคโนโลยีสุรนารี.

Alana, K. (2012). Instagram & CO.: The Mobile Era of Food Photography. [Online]. Available: <https://www.finedininglovers.com/stories/food-photography-tips-mobile/>.

- Aye, Y.M., and Aung, S.S. (2017). Sentiment analysis for reviews of restaurants in Myanmar text. **In Proceedings** of 18<sup>th</sup> IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD). Kanazawa, Japan. pp. 321-326.
- Bollacker, K. D., Lawrence, S., Giles, C. L. (1998). CiteSeer. **Proceedings of the second international conference on Autonomous agents - AGENTS '98**. pp. 116–123.
- Bright Local. (2017). **Local Consumer Review Survey 2017**. [Online]. Available: <https://www.brightlocal.com/learn/local-consumer-review-survey/>.
- Chu, W.T., and Tsai, Y.L. (2017). A Hybrid Recommendation System Considering Visual Information for Predicting Favorite Restaurants. **World Wide Web Journal: Internet and Web Information Systems manuscript**. 20(6): 1313-1331.
- Cortes, C., Vapnik, V.N. (1995). Support-vector networks. **Machine Learning**. 20 (3): 273–297.
- Dancy, J. (2012). **Dine and Dish: Are Social Media and Food the Perfect Pairing**. [Online]. Available: <https://www.columnfivemedia.com/work-items/infographic-dine-and-dish-are-social-media-and-food-the-perfect-pairing>.
- Davis, J. and Goadrich, M. (2006). **The relationship between Precision-Recall and ROC curves**. Proceedings of the 23rd international conference on Machine learning June 2006; pp. 233–240.
- Demandsage. (2023). **Social Media Users — Global Demographics (2023)**. [Online]. Available: <https://www.demandsage.com/social-media-users/#:~:text=Social%20Media%20User%20Statistics%202023,platform%2C%20with%203.03%20billion%20users>.
- Dietrich, D., Heller, B., & Yang, B. (2015). **Data Science & Big Data Analytics Discovering, Analyzing, Visualizing and Presenting Data** (pp. 420). John Wiley & Sons, Inc.
- Dixon, N., Jaki, B., Lagerweij, R., Mooij, M., and Yudin, E. (2012). FoodMood: Measuring Global Food Sentiment One Tweet at a Time. **In Proceedings** of 6<sup>th</sup> International AAAI Conference on Weblogs and Social Media. Dublin, Ireland. pp. 1-7.
- drench. (2016). **HOTELS.COM Mobile Travel Tracker campaign**. [Online]. Available: <http://www.drench-design.com/project/hotels-com-mobile-travel-tracker/>.

- Global Web Index. (2018). **Social Behaviors**. Global Web Index's flagship report on the latest trends in social media H1 2018.
- Global Web Index. (2018). **Social Research TOP 5 BRAND RESEARCH CHANNELS**. Trends 18 THE TRENDS TO KNOW FOR 2018.
- Govindarajan, M. (2014). Sentiment Analysis of Restaurant Reviews Using Hybrid Classification Method. **International Journal of Soft Computing and Artificial Intelligence**. 2(1): 17-23.
- Higgs, S., and Thomas, J. (2016). Social influences on eating. **Current Opinion in Behavioral Sciences**. 9: 1-6.
- Kalinina, M. and Nikolaev, P. (2020). Research of YOLO Architecture Models in Book Detection. **Advances in Intelligent Systems Research**. 174: 218-221.
- Khoo, K. B, Bun, K. K., Ishizuka, M. (2001). Emerging Topic Tracking System. **Proceedings of Third International Workshop on Advanced Issues of E-Commerce and Web-Based Information Systems**. WECWIS 2001. pp. 2–11.
- Luo, Y. et al, (2023). Detection of Agricultural Pests Based on YOLO. **Journal of Physics: Conference Series**. 2560(1): 1-9.
- MacQueen, J. B. (1967). **Some Methods for classification and Analysis of Multivariate Observations**. Proceedings of 5<sup>th</sup> Berkeley Symposium on Mathematical Statistics and Probability. 1. University of California Press. pp. 281–297.
- Mitchell, T. (1997). **Machine Learning**. New York: McGraw-Hill, p. 52-80.
- Nielsen, Jakob, and Landauer, Thomas K. (1993). A Mathematical Model of The Finding of Usability Problems. **In Proceedings of ACM INTERCHI'93 Conference**. pp. 206 - 213. Netherlands: Amsterdam.
- Perera, I. K. C. U., and Caldera, H.A. (2017). Aspect based Opinion Mining from Restaurant Reviews. **In Proceedings of 2<sup>nd</sup> IEEE International Conference on Computational Intelligence and Applications (ICCIA)**. Beijing, China. pp. 542-546.
- Pugsee, P., and Niyomvanich, M. (2015). Sentiment Analysis of Food Recipe Comments. **ECTI Transactions on Computer and Information Technology**. 9(2): 182-193.
- Raul, N., Shah, Y., and Devyaniya, M. (2016). Restaurant Revenue Prediction using Machine Learning. **International Journal of Engineering and Science**. 6(4): 2319-6483.

- Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. **Journal of Documentation**. 60 (5): 503–520.
- Saha, S. (2018). **A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way**. [Online]. Available: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
- Sparck, J. K. (1972). A Statistical Interpretation of Term Specificity and Its Application in Retrieval. **Journal of Documentation**. 28 (1): 11–21.
- Surjandari, I., Naffisah, M.S., and Prawiradinata, M.I. (2015). Text Mining of Twitter Data for Public Sentiment Analysis of Staple Foods Price Changes. **Journal of Industrial and Intelligent Information**. 3(3): 253-257.
- Tama, B.A. (2015). Data mining for predicting customer satisfaction in fast-food restaurant. **Journal of Theoretical and Applied Information Technology**. 75(1): 18-24.
- Yu, B., Zhou, J., and Zhang, Y, and Cao, Y. (2017). **Identifying Restaurant Features via Sentiment Analysis on Yelp Reviews**. eprint arXiv:1709.08698.





## ประวัติผู้วิจัย

นางสาวลดาวัลย์ เข้มครวญ เกิดเมื่อวันที่ 21 กันยายน 2534 จบการศึกษาระดับปริญญาตรี ในปีการศึกษา 2556 วิทยาการสารสนเทศบัณฑิต (ระบบสารสนเทศเพื่อการจัดการ) สาขาวิชา เทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีสุรนารี

หลังจากสำเร็จการศึกษาระดับปริญญาตรีได้เข้าศึกษาต่อในระดับปริญญาโท ณ มหาวิทยาลัยเทคโนโลยีสุรนารี วิทยาการสารสนเทศมหาบัณฑิต และจบการศึกษาระดับปริญญาโท ในปีการศึกษา 2559 หลังจากนั้นได้ศึกษาต่อในระดับปริญญาเอก ปริญญาปรัชญาดุษฎีบัณฑิต ณ มหาวิทยาลัยเทคโนโลยีสุรนารี และได้ปฏิบัติหน้าที่เป็นผู้ช่วยสอนพิเศษให้กับสำนักวิชาศาสตร์ และศิลปดิจิทัล และสำนักวิชาเทคโนโลยีสังคม มหาวิทยาลัยเทคโนโลยีสุรนารี จนถึงปัจจุบัน

