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 และได้นำแบบจำลองที่ได้ไปทดสอบกับข้อมูลใหม่ (Unseen data) จำนวน 298 บทวิจารณ์ แบ่งเป็นบทวิจารณ์ด้านบวก ด้านเป็นกลาง และด้านลบ จำนวน 209 55 และ 34 บทวิจารณ์ และเปรียบเทียบระยะเวลาที่ใช้ในการทำนายของทั้งสองขั้นตอนวิธี นอกจากนี้ ยังใช้ข้อมูลอื่น ๆ ที่ปรากฏอยู่ในบทวิจารณ์ ได้แก่ คำ แฮชแท็ก และสัญลักษณ์ต่าง ๆ เพื่อใช้ในการวิเคราะห์ความถี่ (TF-IDF) ที่ปรากฏในบทวิจารณ์

ส่วนที่ 2 คือการวิเคราะห์ข้อมูลที่เป็นภาพเพื่อจำแนกประเภทอาหาร โดยแบ่งประเภทอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกาโน่ยูซุ ชานมไข่มุก ข้าวปลาแกะ ยำไข่แดงเค็มหมูยอ แซลมอนเสียบไม้ย่าง สปาร์กลิ้งโซดา 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 เพื่อทดสอบข้อมูลใหม่กับแบจำลองที่มีทำนายประเภทของแบบจำลอง 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 ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก 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 ค่าความระลึกเฉลี่ยถ่วงน้ำหนัก 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 แฮชแท็ก และสัญลักษณ์ต่าง ๆ จากบทวิจารณ์ที่เป็นข้อมูลใหม่ (Unseen Data) โดยจัดอันดับ 3 อันดับแรก เพื่อนำไปใช้ในการระบุประเภทของบทวิจารณ์ โดยผลที่ได้จากการนับความถี่คำ แฮชแท็ก 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	เหี้ย	0.036148
2	ชาติหมา	0.025803
3	อย่าแคก	0.024391

จากผลการวิเคราะห์ความถี่คำพบว่า อันดับแรกของคำในบทวิจารณ์ด้านบวก คือ อร่อย มีความถี่ร้อยละ 1.85 ส่วนบทวิจารณ์ที่เป็นเป็นกลาง คำที่ปรากฏเป็นอันดับแรกคือ กินได้ มีความถี่ร้อยละ 1.71 และคำในบทวิจารณ์ด้านลบเป็นอันดับแรกคือ เหี้ย 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 ส่วนบทวิจาร์ที่เป็นกลางมีสัญลักษณ์ 😋 อยู่ในอันดับแรก โดยมีความถี่ร้อยละ 1.60 ทั้งนี้ ยังพบสัญลักษณ์อื่น ๆ ที่ใช้ในการแสดงอารมณ์ เช่น 😍 🥰 😑 😖 😩 🙄 และ 👍 ปรากฏในบทวิจารณ์ทั้ง 3 ประเภทอีกด้วย

4.4 ผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพ

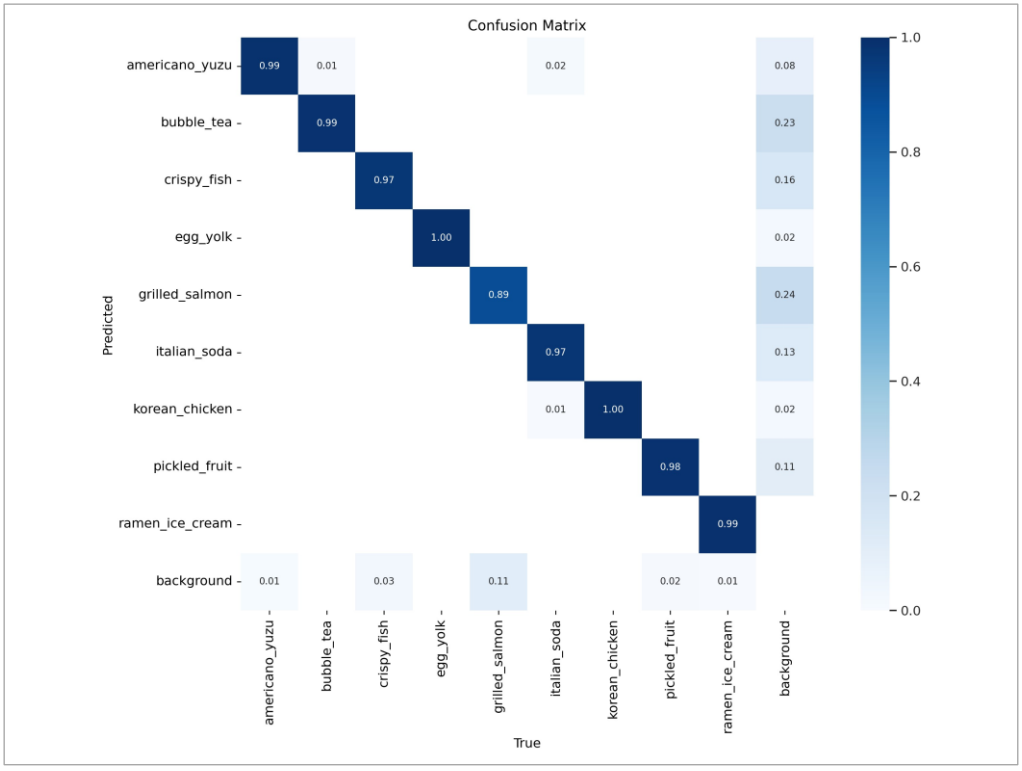
ในการจำแนกประเภทอาหาร โดยจำแนกวัตถุบนภาพ งานวิจัยนี้ใช้วิธีการจำแนกภาพด้วยเครื่องมือ YOLO5 เพื่อสร้างแบบจำลอง โดยแบ่งคลาสอาหารออกเป็น 9 ประเภท ได้แก่ อเมริกาโน่ ยูซุ ชานมไข่มุก ข้าวปลาแกะ ยำไข่แดงเค็มหมูยอ แซลมอนเสียบไม้ย่าง สปาร์กลิ้งโซดา ไก่เกาหลี ผลไม้ดอง และราเมนไอศกรีม จำนวน 1,080 ภาพ แบ่งเป็นประเภทละ 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 อเมริกาโน่ยูซุ (americano_yuzu) ชานมไข่มุก (bubble_tea) ข้าวปลาแกะ (crispy_fish) ยำไข่แดงเค็มหมูยอ (egg_yolk) แซลมอนเสียบไม้ย่าง (grilled_salmon) สปาร์กลิ้งโซดา (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 ซึ่งมีค่าความถูกต้องในการทำนายสูงสุดเป็นร้อยละ 100 มีค่าความแม่นยำ ร้อยละ 98 และค่าความระลึกร้อยละ 99 ส่วนราเมนไอศกรีม มีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุด โดยผลจำนวนของอินสแตนซ์ที่ปรากฏบนภาพที่นำไปสร้างแบบจำลองและผลการสร้างแบบจำลองเพื่อจำแนกประเภทอาหารแบบ Confusion matrix แสดงดังรูปที่ 4.1 และ 4.2 ตามลำดับ

รูปที่ 4.1 จำนวนของอินสแตนซ์ที่ปรากฏบนภาพที่นำไปสร้างแบบจำลอง



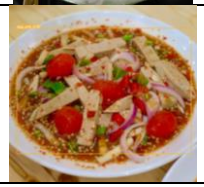
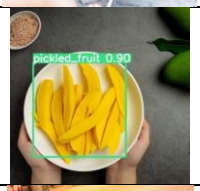
รูปที่ 4.2 Confusion matrix แสดงผลการสร้างแบบจำลองจำแนกประเภทอาหาร

ตัวอย่างภาพจากการสร้างแบบจำลองการจำแนกประเภทอาหาร แสดงดังรูปที่ 4.3

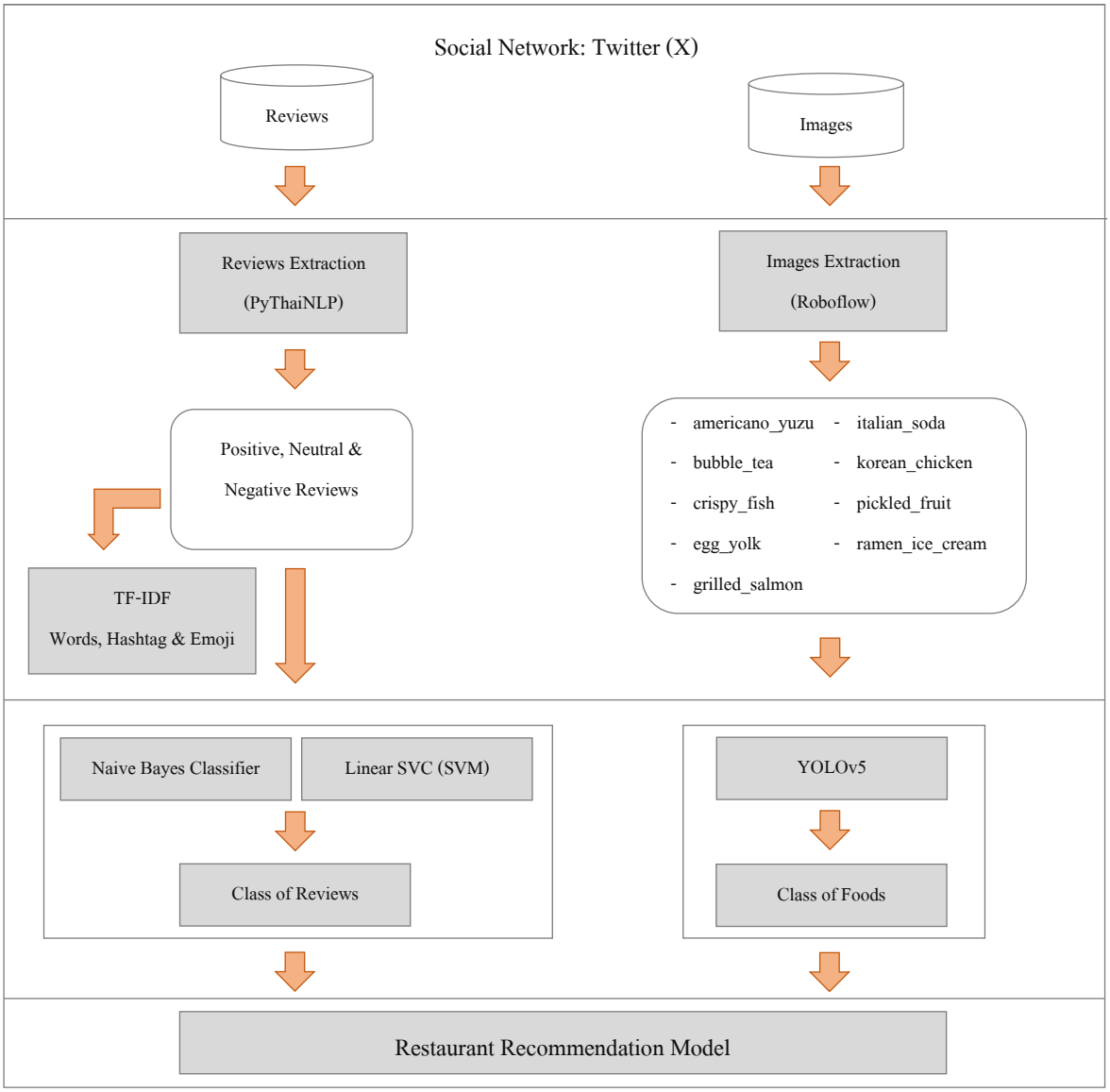
รูปที่ 4.3 ตัวอย่างภาพจากการสร้างแบบจำลองการจำแนกประเภทอาหาร

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 เป็นไปตามสมติฐานการวิจัยในการจำแนกบทวิจารณ์ซึ่งมีค่าความถูกต้องในการจำแนกบทวิจารณ์มากกว่าร้อยละ 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 แฮชแท็ก และสัญลักษณ์ พบว่า คำ แฮชแท็ก และสัญลักษณ์อันดับแรก แสดงให้เห็นถึงความรู้สึกของผู้วิจารณ์ที่มีต่ออาหาร หรือร้านอาหารได้ และผลที่ได้จากการวิเคราะห์ดังกล่าว สอดคล้องกับประเภทของบทวิจารณ์ โดยข้อค้นพบนี้ทำให้สามารถนำผลที่ได้ไปใช้ในการวิเคราะห์ข้อความในบทวิจารณ์ในงานวิจัยในอนาคต เพื่อจำแนกประเภทของบทวิจารณ์ว่าเป็นบทวิจารณ์ด้านบวก ด้านลบ หรือเป็นกลาง โดยแฮชแท็กสามารถปรากฏในบทวิจารณ์ได้ทั้ง 3 ประเภท เนื่องจากเป็นแฮชแท็กที่ได้รับความนิยมในขณะนั้น ซึ่งการใส่แฮชแท็กลงในบทวิจารณ์บางครั้งอาจไม่เกี่ยวข้องกับประเภทของบทวิจารณ์ ซึ่งนอกจากคำที่ปรากฏในบทวิจารณ์แล้ว การวิเคราะห์จากสัญลักษณ์ในบทวิจารณ์เป็นอีกวิธีหนึ่งในการจำแนกประเภทบทวิจารณ์ได้ ซึ่งจะมีประโยชน์อย่างมากสำหรับงานวิจัยที่มีวัตถุประสงค์เพื่อการวิเคราะห์ความรู้สึกของผู้ใช้ที่มีต่ออาหาร และร้านอาหาร

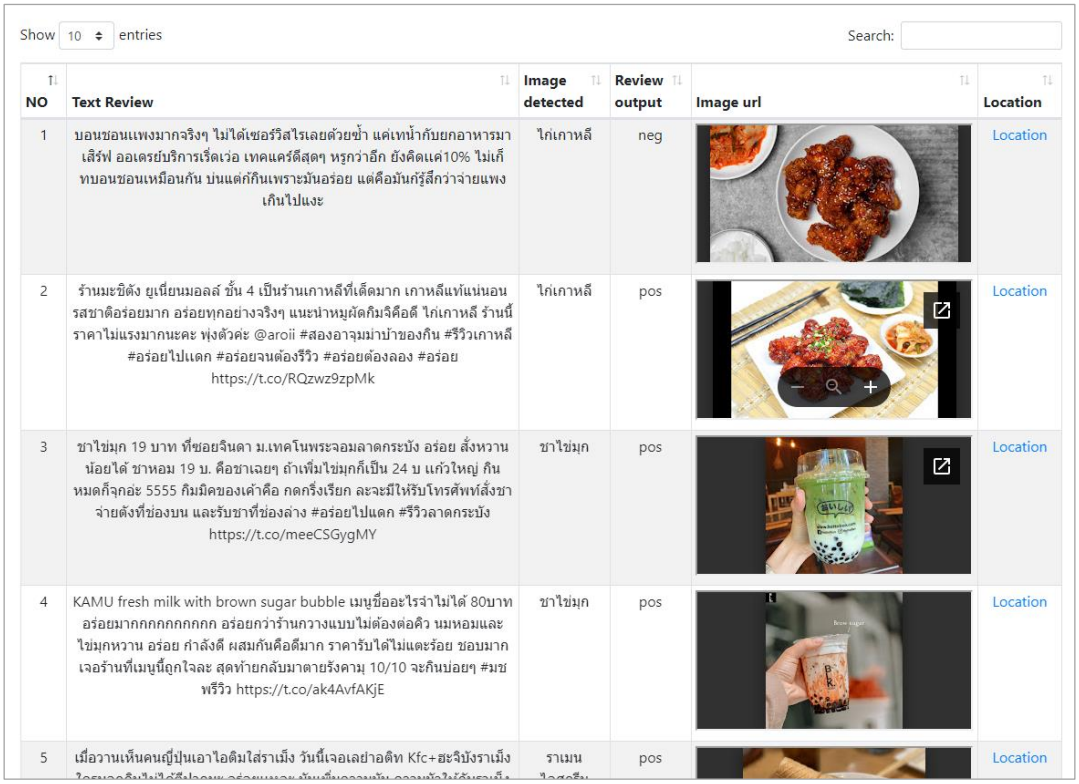
4.5.3 อภิปรายผลการวิจัยการสร้างแบบจำลองการจำแนกประเภทอาหารจากภาพ

จากผลการวิเคราะห์การจำแนกประเภทอาหารบนภาพด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Network: CNN) ด้วยเครื่องมือ YOLO5 โดยแบ่งประเภทอาหารเป็น 9 ประเภท ได้แก่ อเมริกาโน่ยูซุ (americano_yuzu) ชานมไข่มุก (bubble_tea) ข้าวปลาแกะ (crispy_fish) ยำไข่แดงเค็มหมูยอ (egg_yolk) แซลมอนเสียบไม้ย่าง (grilled_salmon) สปาร์กลิ้งโซดา (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 ซึ่งมีค่าความถูกต้องในการทำนายสูงสุดเป็นร้อยละ 100 มีค่าความแม่นยำ ร้อยละ 98 และค่าความระลึกร้อยละ 99 ส่วนราเมนไอศกรีมมีค่าความแม่นยำร้อยละ 100 และค่าความระลึกร้อยละ 99 ซึ่งมีค่าความแม่นยำในการจำแนกประเภทอาหารสูงที่สุด

จากผลการวิจัยที่ได้ สามารถอภิปรายได้ว่า การสร้างแบบจำลองการจำแนกประภทอาหารเป็นไปตามสมติฐานการวิจัยเช่นกัน 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 ซึ่งส่งให้มีค่าความแม่นยำและค่าความระลึกสูงเช่นกัน

ทั้งนี้ จำนวนอินสแตนซ์ที่ปรากฏบนภาพสำหรับใช้จำแนกประเภทอาหารนั้นปรากฏบนภาพที่ต่างประเภทกัน จึงทำให้สามารถจำแนกประเภทของหารได้ชัดเจนยิ่งขึ้น แต่อย่างไรก็ตาม จากภาพที่นำไปใช้ในการทดสอบแบบจำลองนั้น ภาพบางภาพที่ทดสอบยังให้ค่าความถูกต้องในการทำนาย ไม่ถึงร้อยละ 80 อาจเป็นผลมาจากการใช้ภาพที่นำไปสร้างแบบจำลองมีจำนวนน้อย และภาพที่นำไปทดสอบแบบจำลองเป็นภาพที่มีรายละเอียดภายในภาพที่แตกต่างกับภาพที่ใช้ในการสร้างแบบจำลอง จึงทำให้บางภาพที่นำไปทดสอบยังไม่ได้ผลการทดสอบที่มากกว่าร้อยละ 80 อย่างไรก็ตาม การทดสอบภาพกับแบบจำลองเป็นการทดสอบเพื่อทำนายประเภทของอาหารบนภาพเพื่อระบุประเภทอาหารที่ร้านอาหารจำหน่ายเท่านั้น ซึ่งงานวิจัยนี้ได้นำผลการจำแนกประเภทอาหารจากภาพไปใช้เพื่อนำเสนอข้อมูลเกี่ยวกับอาหารที่ร้านอาหารจำหน่าย เพื่อให้สามารถค้นหาร้านอาหารได้ตรงตามความต้องการของผู้ใช้ โดยตัวอย่างการนำเสนอข้อมูลร้านอาหารเพื่อแนะนำร้านอาหารแก่ผู้ใช้แสดงดังรูปที่ 4.5



Show 10 entries Search:

NO	Text Review	Image detected	Review output	Image url	Location
1	บอนชอนแพงมากจริงๆ ไม่ได้เซอร์วิสไรเลยด้วยซ้ำ แค่เทน้ำกับยกอาหารมาเสิร์ฟ ออเดรย์บริการเร็ดเว่อ เทคแคร์ดีสุดๆ หรูกว่าอีก ยังคิดแค่10% ไม่เก็ทบอนชอนเหมือนกัน บ่นแต่กินเพราะมันอร่อย แต่คือมันก็รู้สึกว่าจ่ายแพงเกินไปแงะ	ไก่เกาหลี	neg		Location
2	ร้านมะชิดัง ยูเนี่ยนมอลล์ ชั้น 4 เป็นร้านเกาหลีที่เด็ดมาก เกาหลีแท้แน่นอน รสชาติอร่อยมาก อร่อยทุกอย่างจริงๆ แนะนำหมูผัดกิมจิคือดี ไก่เกาหลี ร้านนี้ราคาไม่แรงมากนะคะ ฟุ่งตัวค่ะ @aroii #สองอาจุมม่าบ้าของกิน #รีวิวเกาหลี #อร่อยไปแดก #อร่อยจนต้องรีวิว #อร่อยต้องลอง #อร่อย https://t.co/RQzwz9zpMk	ไก่เกาหลี	pos		Location
3	ชาไข่มุก 19 บาท ที่ซอยจินดา ม.เทคโนพระจอมลาดกระบัง อร่อย สั่งหวานน้อยได้ ชาหอม 19 บ. คือชาเฉยๆ ถ้าเพิ่มไข่มุกก็เป็น 24 บ แก้วใหญ่ กินหมดก็จุกอะ 5555 กิมมิคของเค้าคือ กดกริ่งเรียก ละจะมีให้รับโทรศัพท์สั่งชาจ่ายตังที่ช่องบน และรับชาที่ช่องล่าง #อร่อยไปแดก #รีวิวลาดกระบัง https://t.co/meeCSGygMY	ชาไข่มุก	pos		Location
4	KAMU fresh milk with brown sugar bubble เมนูชื่ออะไรจำไม่ได้ 80บาท อร่อยมากกกกกกกกก อร่อยกว่าร้านกวางแบบไม่ต้องต่อคิว นมหอมและไข่มุกหวาน อร่อย กำลังดี ผสมกันคือดีมาก ราคารับได้ไม่แดะร้อย ชอบมาก เจอร้านที่เมนูนี้ถูกใจละ สุดท้ายกลับมาตายรังคามุ 10/10 จะกินบ่อยๆ #มชพรีวิว https://t.co/ak4AvfAKjE	ชาไข่มุก	pos		Location
5	เมื่อวานเห็นคนญี่ปุ่นเอาไอดิมใส่ราเม็ง วันนี้เจอเลยอ่าอดิท Kfc+ฮะจิบังราเม็ง	ราเมน	pos		Location

รูปที่ 4.5 ตัวอย่างการนำเสนอข้อมูลร้านอาหารเพื่อแนะนำร้านอาหารแก่ผู้ใช้