



การระบุนิ่วที่นำสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์
ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง



เพ็ญนภา จุฬพลพงษ์

วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรบริหารธุรกิจมหาบัณฑิต
สาขาวิชาการบริหารเทคโนโลยีสารสนเทศเชิงกลยุทธ์

ปีการศึกษา 2564

ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

การระบุวิธีที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์
ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง



วิทยานิพนธ์เสนอบัณฑิตวิทยาลัย มหาวิทยาลัยนเรศวร
เพื่อเป็นส่วนหนึ่งของการศึกษา หลักสูตรบริหารธุรกิจมหาบัณฑิต
สาขาวิชาการบริหารเทคโนโลยีสารสนเทศเชิงกลยุทธ์
ปีการศึกษา 2564
ลิขสิทธิ์เป็นของมหาวิทยาลัยนเรศวร

วิทยานิพนธ์ เรื่อง "การระบุวิธีที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้า
ออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง"

ของ เพ็ญภา จุมพลพงษ์

ได้รับการพิจารณาให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาบริหารธุรกิจมหาบัณฑิต สาขาวิชาการบริหารเทคโนโลยีสารสนเทศเชิงกลยุทธ์

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการสอบวิทยานิพนธ์
(ผู้ช่วยศาสตราจารย์ ดร.อนุพงศ์ อวิรุทธา)

..... ประธานที่ปรึกษาวิทยานิพนธ์
(ดร.อนิรุทธิ์ อัสวสกุลสร)

..... กรรมการที่ปรึกษาวิทยานิพนธ์
(ผู้ช่วยศาสตราจารย์ ดร.วศิน เหลี่ยมปรีชา)

..... กรรมการผู้ทรงคุณวุฒิภายใน
(ผู้ช่วยศาสตราจารย์ ดร.วรเดช ณ กรม)

..... กรรมการผู้ทรงคุณวุฒิภายใน
(ดร.มณีนรัตน์ ขาดีรังสรรค์)

อนุมัติ

.....
(รองศาสตราจารย์ ดร.กรองกาญจน์ ชูทิพย์)

คณบดีบัณฑิตวิทยาลัย

ชื่อเรื่อง	การระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
ผู้วิจัย	เพ็ญนภา จุ่มพลพงษ์
ประธานที่ปรึกษา	ดร.อนิรุทธิ์ อัสวสกุลศร
กรรมการที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.วศิน เหลี่ยมปรีชา
ประเภทสารนิพนธ์	วิทยานิพนธ์ บธ.ม. การบริหารเทคโนโลยีสารสนเทศเชิงกลยุทธ์, มหาวิทยาลัยนเรศวร, 2564
คำสำคัญ	รีวิวที่น่าสงสัย, ผลิตภัณฑ์ดูแลผิวหน้า, ตลาดซื้อขายสินค้าออนไลน์, การเรียนรู้ของเครื่อง

บทคัดย่อ

งานวิจัยฉบับนี้นำเสนอวิธีการระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง สามารถช่วยให้ลูกค้าที่ต้องการอ่านรีวิวก่อนตัดสินใจในการซื้อผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ สามารถรู้ได้ว่ารวิวนั้นเป็นรีวิวจริงหรือเป็นรีวิวที่น่าสงสัย โดยเก็บรวบรวมข้อมูลจากรีวิวผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์แห่งหนึ่งที่เป็นภาษาไทย ระหว่างเดือนมกราคมถึงเดือนพฤศจิกายน 2564 จากผลิตภัณฑ์ดูแลผิวหน้า 10 ยี่ห้อที่มีจำนวนขายสูงสุดตามลำดับ จำนวน 60,081 รีวิว และนำมาพัฒนาแบบจำลองโดยใช้เทคนิควิธีการเรียนรู้ของเครื่อง (Machine Learning) ประกอบด้วย 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เทคนิคต้นไม้ตัดสินใจ (Decision Tree) จากนั้นทำการทดสอบและเปรียบเทียบประสิทธิภาพของแบบจำลอง ผลการวิจัยพบว่าเทคนิคที่เหมาะสมสำหรับการสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ คือ เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ให้ค่าความถูกต้อง (Accuracy) มากที่สุดเท่ากับ 85%

Title	IDENTIFYING SUSPICIOUS CUSTOMER REVIEWS OF FACIAL SKINCARE PRODUCTS ON AN ONLINE MARKETPLACE WITH MACHINE LEARNING TECHNIQUES
Author	PHENNAPA JOOMPONPONG
Advisor	DR. Anirut Asawasakulsorn
Co-Advisor	Assistant Professor DR. Wasin Liampreecha
Academic Paper	M.B.A. Thesis in Strategic Information Technology Management, Naresuan University, 2021
Keywords	IDENTIFYING SUSPICIOUS REVIEWS, FACIAL SKINCARE PRODUCTS, ONLINE MARKETPLACE, MACHINE LEARNING

ABSTRACT

This research presents a method for Identifying Suspicious Customer Reviews of Facial Skincare Products on an Online Marketplace with Machine Learning Techniques. It can help customers to read reviews before deciding whether to buy face care products on the online marketplace to make sure the reviews are real or a questionable review. by collecting data from reviews of facial care products on an online marketplace that is in Thai Between January to November 2021. Out of the 10 best-selling facial skincare brands, a total of 60,081 reviews and used to development a using machine learning techniques. There were 3 techniques used Naives Bayes Technique, Support Vector Machine Technique, and Decision Tree Technique. Then test and compare the performance of the model. The results showed that the ideal technique for modeling identifying suspicious customer reviews of facial skincarre products on an online marketplace with machine learning techniques. The technique supports vector machine, which gives the accuracy as much as 85%.

ประกาศคุณูปการ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดี โดยการสนับสนุนจากอาจารย์ ดร.อนิรุทธิ์ อัครสกุลศร อาจารย์ที่ปรึกษาวิทยานิพนธ์ และ ผู้ช่วยศาสตราจารย์ ดร.วศิน เหลี่ยมปรีชา อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม ซึ่งท่านได้ให้คำแนะนำ ช่วยเหลือ แก้ไขปัญหา เสนอแนะแนวทางและข้อคิดเห็นอันเป็นประโยชน์อย่างยิ่งในการทำวิทยานิพนธ์

ขอขอบคุณ ผู้ช่วยศาสตราจารย์ ดร.วรเดช ณ กรม และ ดร.มณีนรัตน์ชาติรังสรรค์ สำหรับการดูแลและให้คำแนะนำในการทำวิทยานิพนธ์ ขอขอบคุณ ดร.อนุพงศ์ สุขประเสริฐ สำหรับแนวคิดและเทคนิคในการวิเคราะห์ข้อมูล ขอขอบคุณเจ้าหน้าที่ประจำหลักสูตรที่อำนวยความสะดวกในด้านการประสานงาน ดำเนินการเอกสารต่าง ๆ ที่เกี่ยวข้อง และขอขอบคุณกัลยาณมิตรทุกคนที่ส่งกำลังใจมาให้

สุดท้ายนี้ผู้วิจัยขอขอบพระคุณบิดา มารดาและสมาชิกในครอบครัวทุกคนที่คอยช่วยเหลือ ส่งเสริม สนับสนุนและให้กำลังใจมาโดยตลอด

เพ็ญนภา จุมพลพงษ์



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
ประกาศคุุณูปการ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฉ
สารบัญภาพ.....	ฉ
บทที่ 1 บทนำ.....	15
1.1 ความเป็นมาและความสำคัญของปัญหา.....	15
1.2 คำถามของการวิจัย.....	19
1.3 วัตถุประสงค์ของการศึกษา.....	20
1.4 ประโยชน์ที่จะได้รับ.....	20
1.5 ขอบเขตของงานวิจัย.....	20
ขอบเขตด้านประชากร.....	20
ขอบเขตด้านเนื้อหา.....	20
ขอบเขตด้านพื้นที่.....	20
ขอบเขตด้านเวลา.....	20
1.6 นิยามศัพท์เฉพาะ.....	20
1.7 วิธีดำเนินการวิจัย.....	22
ประชากรและกลุ่มตัวอย่าง.....	22

ตัวแปรที่ใช้ในการวิจัย	22
เครื่องมือและการพัฒนาเครื่องมือ	22
การเก็บรวบรวมข้อมูล.....	22
วิธีวิเคราะห์ข้อมูล.....	22
1.8 แผนการดำเนินงาน.....	22
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	23
2.1 ผลิตภัณฑ์ดูแลผิวหน้า (Skincare).....	23
คุณลักษณะผลิตภัณฑ์ดูแลผิวหน้า.....	23
ประโยชน์ของผลิตภัณฑ์ดูแลผิวหน้า.....	23
ประเภทของผลิตภัณฑ์ดูแลผิวหน้า	23
ประเภทของผลิตภัณฑ์ดูแลผิวหน้า	26
งานวิจัยเกี่ยวกับการผลิตภัณฑ์ดูแลผิวหน้า.....	28
2.2 ตลาดซื้อขายสินค้าออนไลน์ (Online Marketplace)	29
ความหมายของ Online Marketplace	40
บทบาทของ Online Marketplace	44
ธุรกิจ e-Commerce ไทยอยู่ในมือของ e-Marketplace ต่างชาติ	45
จุดเด่น ข้อดี ที่ช่วยตอบโจทย์ให้ผู้ซื้อ-ผู้ขาย.....	46
งานวิจัยเกี่ยวกับตลาดซื้อขายสินค้าออนไลน์ (Online Marketplace)	47
2.3 ช่องทางการให้บริการแบบดิจิทัล (Digital Touchpoint).....	49
การเชื่อมโยงแต่ละจุดสัมผัส	51
Touch Point ที่ตอบโจทย์	52
บทบาทของรีวิวสินค้ากับการทำการตลาดที่เปลี่ยนไป	53

งานวิจัยการวิเคราะห์ช่องทางการให้บริการแบบดิจิทัล (Digital Touchpoint).....	54
2.4 รีวิวออนไลน์.....	54
การป้องกันการเปิดเผยข้อมูล.....	57
รีวิวปลอม	58
งานวิจัยรีวิวออนไลน์ (Online Reviews)	61
2.5 การเรียนรู้ของเครื่อง (Machine Learning).....	63
วิวัฒนาการของ Machine learning.....	63
ความเชื่อมโยงระหว่าง Machine Learning และ Artificial Intelligence.....	64
ความสำคัญของระบบการเรียนรู้ของเครื่อง.....	64
สิ่งจำเป็นในการสร้างระบบเพื่อการเรียนรู้ของเครื่องที่มีประสิทธิภาพ.....	65
Machine Learning ในโลกปัจจุบัน	65
วิธีการ Machine Learning.....	66
กระบวนการของ Machine Learning.....	71
งานวิจัยการเรียนรู้ของเครื่อง (Machine Learning).....	72
2.6 การวิเคราะห์เหมืองข้อมูล (Data Mining).....	75
ความสำคัญของการทำเหมืองข้อมูล	76
เทคนิคเหมืองข้อมูล (Data mining).....	77
กระบวนการในการทำเหมืองข้อมูล (Data Mining).....	77
ส่วนประกอบที่สำคัญของระบบการทำเหมืองข้อมูล	79
กลยุทธ์ในการทำเหมืองข้อมูล.....	79
กระบวนการขุดข้อมูลทำงานอย่างไร	79
องค์ประกอบการทำเหมืองข้อมูล (Data Mining).....	80

ลักษณะเฉพาะของข้อมูลที่มีการใช้ทำ Data Mining	81
คุณลักษณะของ Data Mining	81
งานวิจัยการวิเคราะห์เหมืองข้อมูล (Data Mining)	82
2.7 การประมวลผลภาษาธรรมชาติ (Natural Language Processing)	83
ประโยชน์ของการใช้ NLP ที่นำไปใช้ในชีวิตประจำวัน	84
การทำงานของ NLP ที่มีผลต่อ Digital Marketing	84
งานวิจัยการประมวลผลภาษาธรรมชาติ (Natural Language Processing)	87
2.8 การวิเคราะห์ความรู้สึก (Sentiment Analysis)	88
การวิเคราะห์ความรู้สึกด้วย Sentiment Analysis	89
งานวิจัยการวิเคราะห์ความรู้สึก (Sentiment Analysis)	90
บทที่ 3 วิธีดำเนินงานวิจัย	92
3.1 ข้อมูลเบื้องต้น	92
ประชากรและกลุ่มตัวอย่าง	92
ตัวแปรที่ใช้ในการวิจัย	92
เครื่องมือและการพัฒนาเครื่องมือ	92
3.2 วิธีดำเนินการ (Methods)	93
การได้มาซึ่งข้อมูล (Data Acquisition)	93
การเตรียมข้อมูล (Data Preparation)	95
การสร้างแบบจำลอง (Modelling)	102
การประเมินผล (Evaluation)	104
การปรับใช้ (Deployment)	106
ปฏิบัติการ (Operations)	107

การเพิ่มประสิทธิภาพ (Optimization)	107
บทที่ 4 ผลการวิจัย	108
4.1 ผลการวิเคราะห์รีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้า ออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง.....	108
4.2 ผลการสร้างแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขาย สินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง	112
4.3 ผลการทดสอบประสิทธิภาพแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบน ตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง.....	115
บทที่ 5 บทสรุป.....	116
5.1 สรุปผลการวิจัย.....	116
5.2 อภิปรายผล	117
5.3 ข้อเสนอแนะ	118
บรรณานุกรม	118
ประวัติผู้วิจัย	124

สารบัญตาราง

	หน้า
ตาราง 1 ตารางแสดงการวิเคราะห์แบบพยากรณ์.....	108
ตาราง 2 ตารางแสดงการวิเคราะห์ความรู้สึก.....	109
ตาราง 3 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิคนาอิวเบย์ (Naives Bayes).....	112
ตาราง 4 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine).....	113
ตาราง 5 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree).....	113
ตาราง 6 ตารางแสดงผลลัพธ์การเปรียบเทียบการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่อง ทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน.....	114
ตาราง 7 ตารางแสดงผลลัพธ์การทดสอบประสิทธิภาพด้วยวิธีการเรียนรู้ของเครื่องทั้ง 3 เทคนิค ด้วยวิธี K-fold cross validation.....	115

สารบัญภาพ

	หน้า
ภาพ 1 ภาพรวมการใช้ดิจิทัลและอินเทอร์เน็ตของไทย.....	29
ภาพ 2 อิทธิพลของอีมาร์เก็ตเพลส.....	30
ภาพ 3 ช่องทางอีมาร์เก็ตเพลสที่นิยมซื้อขายมากที่สุด.....	32
ภาพ 4 E-commerce Channels (Marketplace).....	34
ภาพ 5 E-commerce Channels (Social Commerce).....	36
ภาพ 6 E-commerce Channels (E-tailer).....	36
ภาพ 7 E-commerce Channels (Payment).....	37
ภาพ 8 มูลค่า E-Commerce ปี 2560-2564.....	39
ภาพ 9 มูลค่า E-Commerce แบบ B2C.....	39
ภาพ 10 มูลค่า E-Commerce จากผู้ประกอบการ SMEs.....	39
ภาพ 11 มูลค่าตลาด E-Commerce.....	40
ภาพ 12 มูลค่าตลาด E-Commerce ในสถานการณ์โควิด-19.....	41
ภาพ 13 จำนวนร้านค้า Brand Official Shop.....	42
ภาพ 14 การตลาดกับ Influencer.....	43
ภาพ 15 Convergence of Platform in Thailand Digital Landscape.....	43
ภาพ 16 วงจรชีวิตการเรียนรู้ของเครื่อง.....	67
ภาพ 17 วิธีการเรียนรู้ของเครื่อง.....	68
ภาพ 18 การเรียนรู้แบบจำลอง.....	68
ภาพ 19 การอนุมาน.....	69

ภาพ 20 Training Data	71
ภาพ 21 กระบวนการวิธีในการทำเหมืองข้อมูล	77
ภาพ 22 องค์ประกอบการทำเหมืองข้อมูล	80
ภาพ 23 การเติบโตของตลาดที่ใช้งานเทคโนโลยี NLP ทั่วโลก.....	83
ภาพ 24 การวิเคราะห์ความรู้สึกเชิงบวกและเชิงลบ	85
ภาพ 25 ขั้นตอนการเตรียมข้อมูลประเภท text.....	89
ภาพ 26 ภาพรวมของวิธีดำเนินการ.....	93
ภาพ 27 การดึงข้อมูลรีวิวผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ ..	94
ภาพ 28 ความไม่สมบูรณ์ของข้อมูลรีวิวในผลิตภัณฑ์ดูแลผิวหน้า	94
ภาพ 29 Download Library.....	95
ภาพ 30 ตัวอย่างการจำแนกข้อมูลรีวิว (Classification) ทำนายหมวดหมู่ของรีวิว.....	97
ภาพ 31 การทำ Text Cleaning	97
ภาพ 32 การลบข้อมูลที่มีสิ่งรบกวน (noise).....	98
ภาพ 33 ตัวอย่างรีวิวที่ถูกลบหรือลดข้อมูลที่มีสิ่งรบกวน (noise).....	98
ภาพ 34 การทำ Tokenization.....	100
ภาพ 35 ตัวอย่างการทำ Tokenization.....	100
ภาพ 36 การนำ Scikit-learn มาใช้งาน	101
ภาพ 37 การจับกลุ่มรีวิวที่แบ่งแล้วโดยใช้ spacebar คั่น.....	102
ภาพ 38 การกำหนด TfidfVectorizer จาก sklearn library	103
ภาพ 39 การทำ Vectorizer.....	103
ภาพ 40 การทำ Machine Learning ด้วยเทคนิคทั้ง 3 วิธี	104
ภาพ 41 การทำ Pipeline.....	106

ภาพ 42 การค้นหาแบบกริด GridSearch 106



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

จากภาวะเศรษฐกิจโลกที่ชะลอตัวลงเนื่องจากผลกระทบการระบาดของโรคโควิด-19 รวมถึงราคาน้ำมันที่ปรับลดลงจนกระทบต่อประเทศผู้ผลิตน้ำมันถือเป็นความท้าทายต่อไทยในฐานะประเทศผู้ส่งออกเครื่องสำอางที่สำคัญของโลก ซึ่งนอกจากลูกค้าจะมีความระมัดระวังการใช้จ่ายแล้วในช่วงที่ประชาชนมีการชะลอในการทำกิจกรรมนอกบ้าน ยังมีอุปสรรคจากการจัดกิจกรรมส่งเสริมการขายในแต่ละประเทศที่ทำได้ยากขึ้น ซึ่งมูลค่าเครื่องสำอางไทยได้รับความนิยมในตลาดอาเซียน ญี่ปุ่น ออสเตรเลีย ฮองกง และจีน ส่งผลให้ไทยเป็นประเทศส่งออกเครื่องสำอางอันดับ 2 อาเซียนและอันดับ 10 ของโลก ใช้ข้อได้เปรียบด้านภาษีภายใต้เอฟทีเอ เพิ่มโอกาสการส่งออกและขยายตลาดต่างประเทศ พบว่า สินค้าเครื่องสำอาง อาทิ ผลิตภัณฑ์บำรุงผิวและตกแต่งหน้า สบู่ แชมพู ผลิตภัณฑ์เพื่ออนามัยในช่องปากและฟัน และวัตถุดิบที่ใช้ทำเครื่องสำอาง เป็นหนึ่งในสินค้าที่ได้รับการยอมรับจากลูกค้าในต่างประเทศเนื่องจากในช่วงหลายปีที่ผ่านมามูลค่าการส่งออกขยายตัวอย่างต่อเนื่อง และในปัจจุบันไทยเป็นประเทศผู้ส่งออกเครื่องสำอางอันดับที่ 2 ของอาเซียน รองจากสิงคโปร์ และเป็นอันดับที่ 10 ของโลก (กรมเจรจาการค้าระหว่างประเทศ, 2564)

ทั้งนี้ปี 2564 โรคระบาดยังอยู่ แต่ผู้ประกอบการ ผู้บริโภคปรับตัวให้เข้ากับสถานการณ์สมาคมผู้ผลิตเครื่องสำอางไทย ประเมินแนวโน้มตลาดเครื่องสำอางเติบโตอย่างต่ำ 3% นับเป็นการฟื้นตัวกลับมาอีกครั้ง ประกอบกับสถานการณ์ตลาดก่อนสงกรานต์ ยอดออเดอร์สินค้าจากต่างประเทศมีการเติบโตอย่างมาก สร้างมูลค่าถึง 6,000-7,000 ล้านบาท ส่วนตลาดในประเทศ ปฏิเสธไม่ได้ว่าพฤติกรรมของผู้บริโภคมีความรักสวยรักงาม ยิ่งกว่านั้นมาพร้อมกับการดูแลสุขอนามัยเข้มข้นขึ้นส่งผลให้สินค้าหมวดดูแลผิวพรรณหรือสกินแคร์ ยังขยายตัว รวมถึงหมวดผลิตภัณฑ์ดูแลเส้นผมมีการเติบโตได้ โดยทั้ง 2 หมวดยังเป็นตลาดใหญ่มูลค่า 80,000 ล้านบาท และ 30,000 ล้านบาทตามลำดับ รวมถึงผลิตภัณฑ์กลุ่มทำความสะอาด เช่น ครีมอาบน้ำด้วย ขณะที่เครื่องสำอางที่หัตถ์ยังเกิดขึ้นต่อเนื่อง ได้แก่ กลุ่มผลิตภัณฑ์สีส้นหรือเมกอัพ มีมูลค่าราว 20,000 ล้านบาท หดตัว 10% น้ำหอม ผลิตภัณฑ์กันแดด และเครื่องสำอางจับกลุ่มผู้ชาย มูลค่าหลักพันล้านบาท เมื่อผู้บริโภคต้องทำงานที่บ้าน อยู่บ้านมากขึ้น การใช้ผลิตภัณฑ์ดังกล่าวจึงลดน้อยลง แต่ด้วยขนาดตลาดไม่ใหญ่ การหดตัวที่เกิดขึ้นจึงไม่กระทบภาพรวมมากนัก ตลาดเครื่องสำอางมีการเติบโตมาโดยตลอด แต่ปี 2564 เป็นครั้งแรกที่เจอการหดตัว

ในปัจจุบันลูกค้าได้เปลี่ยนพฤติกรรมการซื้อสินค้าโดยผ่านช่องทางออนไลน์ที่ได้รับความนิยมสูงสุดคือ eMarketplace มีส่วนแบ่งมากถึง 47% ของมูลค่าตลาดทั้งหมด รองลงมา ได้แก่ โซเชียลมีเดีย และแบรนด์ดอทคอม (GroupM Focal, 2020) แม้ว่าผลการสำรวจพฤติกรรมการเลือกใช้แพลตฟอร์ม E-commerce ของผู้บริโภค พบว่า สัดส่วนการซื้อสินค้าแต่ละกลุ่มผ่านแพลตฟอร์มออนไลน์ของผู้บริโภคในปี 2565 มีแนวโน้มเพิ่มขึ้น เมื่อเทียบกับการใช้จ่ายในภาพรวมของสินค้าแต่ละประเภท รวมถึงความไม่แน่นอนของโควิด-19 ส่งผลให้คาดว่าผู้บริโภคยังคงมีการใช้จ่ายอย่างระมัดระวัง ผู้บริโภคได้มีการปรับเปลี่ยนพฤติกรรมโดยการวางแผนปรับลดการใช้จ่ายผ่านหน้าร้าน (Physical stores) มาเป็นการใช้จ่ายผ่านแพลตฟอร์มออนไลน์ (E-commerce) มากขึ้น ภายใต้งบประมาณการใช้จ่ายโดยรวมที่ยังคงจำกัด โดยเฉพาะสินค้าในกลุ่มอาหารและของใช้ส่วนตัวที่เดิมทีผู้บริโภคซื้อผ่านช่องทางหน้าร้าน ไม่ว่าจะเป็นไฮเปอร์มาร์เก็ต ซูเปอร์มาร์เก็ตและคอนวีเนียนสโตร์ จากพฤติกรรมผู้บริโภคที่เปลี่ยนไปและการปรับตัวของผู้ประกอบการค้าปลีกผ่านการทำกลยุทธ์ Omni-channel (Offline to Online) ส่งผลให้ผู้บริโภคหันมาซื้อสินค้าผ่านแพลตฟอร์มออนไลน์มากขึ้น สะท้อนได้จากยอดขายของผู้ประกอบการค้าปลีกบางรายที่ยอดขายผ่านแพลตฟอร์มออนไลน์ในปีที่ผ่านมาโต 2-3 เท่าตัว แต่ยอดขายในภาพรวมกลับไม่ได้เพิ่มขึ้นมากนักหรือโตในกรอบจำกัด (และเป็นผลของราคาเป็นหลัก) สอดคล้องไปกับผลสำรวจของผู้บริโภคที่คาดว่าปีนี้ หันมาใช้จ่ายซื้อสินค้าผ่านแพลตฟอร์มออนไลน์มากขึ้นกว่า 35% ขณะที่งบประมาณยังคงเท่าเดิมหรือใช้จ่ายอย่างจำกัด ส่งผลให้ในเบื้องต้นธุรกิจ B2C E-commerce กลุ่มสินค้าปี 2565 มีมูลค่าตลาดราว 5.65 แสนล้านบาท ขยายตัว 13.5% (YoY) ซึ่งถือเป็นอัตราการขยายตัวที่ต่ำสุดเมื่อเทียบกับในช่วง 3 ปีก่อนหน้า (ปี 2562-2564) ที่ขยายตัวเฉลี่ยปีละ 40% โดยการขยายตัวของ B2C E-commerce ส่วนหนึ่งเป็นเพราะราคาสินค้าที่ไม่สูงและทำให้รู้สึกถึงความคุ้มค่าโดยเฉพาะจากการจัดทำโปรโมชั่นของผู้ประกอบการที่อยู่บนแพลตฟอร์มออนไลน์ซึ่งมีความถี่ขึ้น อีกทั้งผู้บริโภคก็มีความสะดวกและคุ้นเคยกับการซื้อสินค้าออนไลน์ในช่วง 1-2 ปีที่ผ่านมาจากการระบาดของโควิด-19 นอกจากนี้ส่วนแบ่งตลาดของธุรกิจ B2C E-commerce กลุ่มสินค้าขยับเพิ่มขึ้นจาก 14% ในปี 2564 เป็น 16% ในปี 2565 เมื่อเทียบกับมูลค่าตลาดค้าปลีกสินค้ารวมทั้งหมด โดยเฉพาะกลุ่มอาหารเครื่องดื่ม และของใช้ส่วนตัว ซึ่งแม้ว่ายังมีสัดส่วนออนไลน์ที่น้อยมาก แต่เนื่องจากเป็นกลุ่มสินค้าจำเป็นจึงคาดว่าผู้บริโภคยังคงมีการใช้จ่ายผ่านช่องทางดังกล่าวเพิ่มขึ้น (ศูนย์วิจัยกสิกรไทย, 2565)

พฤติกรรมลูกค้าชาวไทยในปัจจุบัน เมื่อต้องการซื้อสินค้าทั้งออนไลน์และออฟไลน์ มากถึง 90% จะหาข้อมูลประกอบการตัดสินใจก่อนซื้อและส่วนใหญ่หาข้อมูลผ่านช่องทางออนไลน์สิ่งเหล่านี้ทำให้เห็นว่าแพลตฟอร์มมีความสำคัญอย่างมากในการช่วยให้ลูกค้าตัดสินใจซื้อสินค้า (Marketeer, 2020) ปัจจุบันการรีวิวสินค้ากลายเป็นอีกหนึ่งในกลยุทธ์ยอดนิยมในการทำตลาดออนไลน์ของนักการตลาดทั่วโลก เพราะสามารถช่วยเพิ่มประสิทธิภาพทางการตลาดได้มหาศาลกระแส Social Network ที่

เข้ามามีบทบาทอย่างมากในชีวิตประจำวันของผู้คนยุคนี้ ซึ่งปัจจุบันสื่อสังคมออนไลน์เกือบกลายเป็นสื่อหลักของทุกประเทศไปแล้ว ทั้งนี้การตลาดออนไลน์เชิงอิทธิพล (Influencer Marketing) แบบรีวิวลินค้าโดย บล็อกเกอร์ เน็ตไอดอล หรือเซเลบ เริ่มเป็นที่ยอมรับมากขึ้นในหมู่แบรนด์สินค้าต่างๆ เนื่องจากภาพคือตัวแทนของลูกค้า จึงทำให้รู้สึกว่ามีการใช้งานจริง จากผู้มีประสบการณ์จริง การเขียนรีวิวลินค้า ผลิตภัณฑ์ต่างๆของเหล่า Influencer บนโลกออนไลน์เลยดูเป็นมิตร มีความเป็นธรรมชาติมากกว่า การรีวิวลินค้า โดย ดารา นักร้อง นักแสดง สอดคล้องกับผลวิจัยด้านการตลาดที่ยืนยันว่าการใช้วิธี รีวิวลินค้า มีส่วนสำคัญต่อการตัดสินใจของลูกค้าสูง เพราะพฤติกรรมของผู้คนยุคนี้ส่วนใหญ่เชื่อคำแนะนำแบบปากต่อปากจากเพื่อนฝูงและการบอกต่อกันในโลกออนไลน์มากกว่าโฆษณา ขายเป็นออนไลน์จำเป็นต้องมีรีวิว ซึ่งจากข้อมูลกรุงศรี กูรู ได้กล่าวถึงข้อมูลการสำรวจของ Bright Local ที่เกี่ยวข้องกับผู้บริโภคยุคนี้ตัดสินใจเลือกซื้อสินค้าจากอะไร พบว่า ผู้บริโภคจำนวนมากกำลังอ่านรีวิวมากกว่าที่เคยเป็นมา ซึ่งในปี 2020 ผู้บริโภคอ่านรีวิว 60% และในปี 2021 ผู้บริโภคอ่านเป็นประจำ 77% สำหรับเปอร์เซ็นต์ของผู้คนที่ไม่เคยอ่านรีวิว ในปี 2020 ไม่เคยอ่าน 13% ในปี 2021 ไม่เคยอ่านรีวิวเหลือเพียง 2% และ 88% ของผู้บริโภคนั้นเชื่อการรีวิวลินค้าบนโลกออนไลน์มากพอ ๆ กับการแนะนำจากครอบครัว ช่างแบบออฟไลน์ (รวีศ หาญอุตสาหกรรม, 2563)

การมี “รีวิว” (Review) แสดงประกอบรายละเอียดสินค้าในเว็บไซต์ หรือร้านค้าบนแพลตฟอร์มออนไลน์ต่าง ๆ นั้นจะช่วยเสริมการพิจารณาของลูกค้าที่กำลังดูร้านค้าอยู่ สิ่งลูกค้าในยุคนี้มองหาเมื่อสนใจสินค้าหรือบริการก่อนตัดสินใจซื้อนั่นคือรีวิวจากลูกค้าก่อนหน้านั้นนั่นเอง นอกจากนี้ความสำคัญอื่นๆ ที่ได้รับจากการมีรีวิวลินค้าหรือรีวิวบริการ ได้แก่ 1) ความน่าเชื่อถือ เพราะก่อนเสียเงินไปซื้อของอะไรก็ตาม หากดูแล้วร้านค้าที่กำลังสนใจมีรายละเอียดสินค้าที่ไม่น่าเชื่อถือ ไม่มีรีวิวจากลูกค้าว่าสั่งซื้อแล้วได้รับของมาสภาพไหน , การจัดส่งไวหรือไม่ , ตอบคำถามลูกค้าดีไม่ได้อย่างไร หากเป็นอย่างนี้ความเชื่อมั่นต่อร้านค้าก็ไม่เกิดขึ้น 2) ได้รับความคิดเห็นโดยตรงจากลูกค้า บ่อยครั้งมีความคิดเห็นดี ๆ เพื่อนำไปพัฒนาปรับปรุงสินค้าและบริการมาจากลูกค้าโดยตรง หลังจากที่ได้ทดลองใช้สินค้าแล้ว ซึ่งเป็นประสบการณ์จากผู้ใช้งานจริง ๆ ที่มีคุณค่าอาจช่วยให้พัฒนาสินค้าที่ตอบสนองความต้องการของลูกค้าได้มากที่สุด 3) ช่วยค้นหาสินค้าเจอได้ง่าย เครื่องมือค้นหาเสิร์ชเอนจินที่คุ้นเคยกันดีอย่างกูเกิ้ล (Google) ดึงเอาสินค้าที่มีคนรีวิวมาแสดงให้นักท่องเที่ยวเห็นด้วย ยิ่งสินค้าได้รับการรีวิวหรือพูดถึงจากลูกค้ามีจำนวนมาก ยิ่งเป็นตัวชี้วัดถึงความน่าเชื่อถือและความน่าสนใจของสินค้ามากตามไปด้วย จนทำอยากคลิกเข้าไปดูเว็บไซต์เพิ่มขึ้น 4) รีวิวด้านลบบางครั้งส่งผลดี ความผิดพลาดเป็นสิ่งที่มีโอกาสเกิดขึ้นได้เสมอ หากเจ้าของร้านรับมือกับบางรีวิวที่ตำหนิการให้บริการได้ดีและแสดงออกถึงความจริงใจในการขอภัยลูกค้า รับผิดชอบในความผิดพลาดอย่างเหมาะสม ลูกค้าที่เข้ามาอ่านรีวิวรู้สึกถึงความน่าเชื่อถือ ช่วยเพิ่มการตัดสินใจในการใช้บริการมากขึ้น เพราะได้เห็นถึงการบริการหลังการขายอีกมุมหนึ่ง หากร้านค้ามีแต่คำชมอย่างเดียว ทำให้ความรู้สึกน่าสงสัยมากกว่าเกิดความ

ไว้วางใจ 5) สินค้ามีรีวิว คนยิ่งสนใจ เมื่อลูกค้ารายใหม่ ๆ เข้ามาเห็นรีวิวก่อนหน้านี้ก็เกิดแรงจูงใจและความกล้าที่แชร์ความเห็นเพิ่มเติมทิ้งไว้ให้คนอื่น ๆ ได้รับรู้ต่อไป ยิ่งสินค้าขึ้นใหม่มีรีวิวใหม่ ๆ เพิ่มเติมสม่ำเสมอ คนให้ความสนใจเพิ่มตามไปด้วย (SMEONE, 2020) ช่องทางการซื้อของออนไลน์ที่ค่อนข้างปลอดภัยคงหนีไม่พ้นการซื้อผ่านแอปพลิเคชันชอปปิงออนไลน์ เนื่องจากระบบมีการคัดกรองร้านค้าให้เบื้องต้น และยังสามารถอ่านรีวิวของลูกค้าที่ซื้อของได้อีกด้วย อย่างไรก็ตาม ลูกค้าต้องสังเกตด้วยว่ารีวิวที่เห็นจำนวนมากนั้น มีความจริงปะปนอยู่มากน้อยเพียงใด เพราะบางทีการให้ดาวเยอะ ๆ หรือคอมเมนต์รีวิวแสบกวัก ซึ่งอาจไม่ได้สะท้อนถึงคุณภาพที่แท้จริงของสินค้าก็เป็นได้ จากกระแสในการเลือกซื้อผลิตภัณฑ์ดูแลผิวหน้ามีความต้องการเป็นอย่างสูง และทำให้ธุรกิจจำเป็นต้องใช้ช่องทางออนไลน์เป็นหน้าร้านในการซื้อขายสินค้า ปัจจัยที่สำคัญอย่างหนึ่งในการตัดสินใจเลือกซื้อผลิตภัณฑ์ดูแลผิวหน้าของลูกค้า คือการอ่านรีวิวเพื่อใช้เป็นข้อมูลสนับสนุนในการตัดสินใจโดยตรงตามรีวิวหรือความคิดเห็นที่เขียนโดยผู้อื่นตามประสบการณ์ ในการแข่งขันทุกคนสามารถเขียนอะไรก็ได้ และลูกค้าจะทราบได้อย่างไรว่ารีวิวจากสินค้าดังกล่าว เป็นรีวิวจริง หรือเป็นรีวิวที่มีความน่าสงสัย เมื่อในแต่ละวันก่อนที่ลูกค้าจะเลือกซื้อผลิตภัณฑ์ดูแลผิวหน้าผ่านตลาดซื้อขายสินค้าออนไลน์ ต้องมีการอ่านรีวิวเป็นจำนวนมาก การพิจารณาตัดสินใจเชื่อรีวิวเป็นเรื่องยากที่ลูกค้าอาจไม่ทราบว่ารีวิวนั้นเป็นรีวิวจริงหรือรีวิวที่น่าสงสัย เนื่องจากธุรกิจหลายแห่งจ้างผู้คนให้เขียนรีวิวเชิงบวกปลอมเกี่ยวกับบริการหรือผลิตภัณฑ์ของตน หรือรีวิวเชิงลบที่ไม่เป็นธรรมเกี่ยวกับบริการหรือผลิตภัณฑ์ของคุณแข่ง กระบวนการนี้ให้ข้อมูลที่ผิดกับลูกค้าที่ต้องการซื้อสินค้า

สำหรับรีวิวของลูกค้าไม่ได้มีบทบาทเหมือนกันในทุกธุรกิจ โดยภาคอุตสาหกรรมร้านอาหาร โรงแรม การแพทย์ ยานยนต์ และร้านเสื้อผ้า เป็นอุตสาหกรรม 5 อันดับแรกที่ได้รับผลกระทบมากที่สุด จากความคิดเห็นของลูกค้าออนไลน์ แต่ไม่ว่าจะอยู่ในอุตสาหกรรมใด มีลูกค้าเพียง 48% เท่านั้นที่พิจารณาใช้บริการจากธุรกิจที่ได้รับการจัดอันดับเป็นดาวน้อยกว่า 4 ดวง ในทางกลับกันรีวิวของลูกค้าออนไลน์ที่ไม่ดีก็สามารถส่งผลดีต่อชื่อเสียงทางธุรกิจได้ โดยการที่ลูกค้าในอนาคตเห็น รีวิวเชิงลบ อาจจะช่วยเพิ่มความน่าเชื่อถือให้กับธุรกิจได้ และมีแนวโน้มที่จะเชื่อถือธุรกิจมากขึ้น มีลูกค้ามากถึง 82% ค้นหารีวิวเชิงลบประกอบการตัดสินใจ ส่วนสาเหตุว่าทำไมถึงทำเช่นนั้น นั่นก็เพราะไม่มีใครเชื่อถือว่ามีผลิตภัณฑ์ทุกชิ้นเป็นผลิตภัณฑ์ระดับห้าดาวที่สมบูรณ์แบบ รีวิวเชิงลบจึงเป็นพื้นฐานสำหรับสถานการณ์ที่ผู้บริโภคอาจมีต่อผลิตภัณฑ์ ซึ่งการฟังเสียงของลูกค้าแล้วนำมาใช้ให้เป็นประโยชน์ จะช่วยให้ธุรกิจพัฒนาสินค้าและบริการให้เป็นที่พึงพอใจของลูกค้า ซึ่งเป็นปัจจัยสำคัญในการดำเนินธุรกิจให้ติดตลาดปูทางสู่ความสำเร็จ แต่การเขียนรีวิวปลอมเป็นสิ่งที่ไม่ควรทำที่สุด จะส่งผลกระทบร้ายแรงกับแบรนด์ อีกทั้งยังทำให้ความน่าเชื่อถือของแบรนด์ลดลง หรืออาจจะหายไปพร้อมกับยอดขายซ้ำร้ายคือ ธุรกิจอาจไปไม่รอดกันเลยทีเดียว เชื่อว่าลูกค้าก็สังเกตเห็นอยู่ว่า อันไหนเป็นรีวิวปลอม แต่บ่อยครั้งที่ลูกค้าไม่ทันสังเกตเห็น (SME Update, 2021)

ซึ่งเจ้าของแบรนด์ต้องมีวิธีการกับรีวิวในเชิงที่ไม่เหมาะสม หรือเป็นเชิงลบ คือความสามารถในการแก้ไขสถานการณ์ที่รวดเร็ว มีการตอบสนองอย่างทันทีหากเป็นการบกพร่อง หรือความเสียหายที่เกิดจากฝั่งของผลิตภัณฑ์ต้องยอมรับข้อผิดพลาด รับฟังความคิดเห็นของลูกค้า มีการพูดคุยกับผู้เกี่ยวข้องดำเนินการแก้ไขให้ถูกต้องในระยะเวลาที่เหมาะสม แม้ว่าการรีวิวผลิตภัณฑ์นั้น จะไม่ใช่ปัจจัยเดียวที่ทำให้แบรนด์ประสบความสำเร็จเพราะการตลาดของผลิตภัณฑ์จะได้ดีนั้น ขึ้นกับองค์ประกอบหลายประการ แต่แนวโน้มของการตลาดได้ปรับเปลี่ยนกลยุทธ์โดยการใช้ Influencer หรือว่าผู้ที่มีความน่าสนใจ ทั้งที่ผ่านสื่อต่างๆ และ Social Media เพื่อรีวิวผลิตภัณฑ์เพื่อจูงใจให้กลุ่มเป้าหมายหันมาใช้ผลิตภัณฑ์แบรนด์ของตนเองให้มากยิ่งขึ้น การขยายผลแบบปากต่อปาก หรือประสบการณ์การใช้งานเป็นสิ่งที่สำคัญมาก การรีวิวและความคิดเห็น จึงเป็นการสะท้อนบทบาทของการสื่อสารแบรนด์ได้เป็นอย่างดี ด้วยเหตุผลดังกล่าวทำให้ผู้วิจัยได้ศึกษาและทดลองการระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง ที่อาจบ่งชี้ถึงประสิทธิภาพว่าจะจะเป็นรีวิวจริง หรือรีวิวที่มีความน่าสงสัย เป็นสินค้าที่ต้องระมัดระวังเป็นพิเศษ และไม่ให้เกิดความเข้าใจผิดหรือหลงเชื่อไป แต่พอได้รับสินค้ามาแล้วกลับไม่เป็นไปอย่างที่คาดหวังไว้ โดยใช้ทฤษฎีและหลักการกระบวนการเรียนรู้ของเครื่อง (Machine learning) โดยการดึงรีวิว (Scrape reviews) ในส่วนของผลิตภัณฑ์ดูแลผิวหน้าที่ต้องการจากตลาดซื้อขายสินค้าออนไลน์ เป็นการได้มาซึ่งข้อมูลรีวิว การทำความสะอาดข้อมูล (Data Cleaning) ซึ่งเป็นการทำงานที่เกี่ยวข้องกับการตรวจสอบและแก้ไขข้อมูล เพื่อให้ข้อมูลอยู่ในรูปแบบที่ถูกต้องและสมบูรณ์ และนำมาสร้างแบบจำลองด้วยเทคนิควิธีการเรียนรู้ของเครื่อง เลือกอัลกอริทึมที่ให้ผลลัพธ์ที่ดีที่สุด และการประเมินผลหลังจากฝึกแบบจำลองเกี่ยวกับข้อมูลรีวิว โดยประเมินผลลัพธ์สุดท้ายเพื่อดูว่าแบบจำลองทำงานได้ดีหรือเชื่อถือได้เพียงใด เพราะการรีวิวเป็นอีกคำตอบสำคัญที่ลูกค้าสามารถค้นหาได้ และกลายมาเครื่องมือช่องทางทางการตลาดให้เลือกใช้ในรูปแบบที่เหมาะสมทั้งกับผลิตภัณฑ์ ช่วงระยะเวลา โอกาส และกลยุทธ์ทางการตลาดที่กำหนดไว้

1.2 คำถามของการวิจัย

1. การวิเคราะห์รีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์มีลักษณะอย่างไร
2. การค้นหาแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่องมีวิธีการใดบ้าง
3. การทดสอบประสิทธิภาพแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่องมีรูปแบบอย่างไร

1.3 วัตถุประสงค์ของการศึกษา

1. เพื่อวิเคราะห์รีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
2. เพื่อสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
3. เพื่อทดสอบประสิทธิภาพแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง

1.4 ประโยชน์ที่จะได้รับ

1. ได้แบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
2. สามารถทดสอบประสิทธิภาพแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
3. สามารถนำแบบจำลองไปใช้ในการพัฒนาแพลตฟอร์มเพื่อระบุรีวิวที่น่าสงสัยได้

1.5 ขอบเขตของงานวิจัย

ขอบเขตด้านประชากร

ประชากรที่ใช้ในการทำวิจัย ได้แก่ รีวิวผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์
ขอบเขตด้านเนื้อหา

งานวิจัยนี้เลือกเก็บข้อมูล 8 คุณลักษณะ ได้แก่ ยี่ห้อสินค้า ราคา ชื่อร้านค้า ความเร็วในการตอบกลับ ชื่อลูกค้า ข้อมูลรีวิว วันเวลาที่รีวิว จำนวนดาว

ขอบเขตด้านพื้นที่

ตลาดซื้อขายสินค้าออนไลน์

ขอบเขตด้านเวลา

ระหว่างเดือนพฤศจิกายน - ธันวาคม พ.ศ.2564

1.6 นิยามศัพท์เฉพาะ

ผลิตภัณฑ์ดูแลผิวหน้า ได้แก่ ผลิตภัณฑ์สกินแคร์ (Skincare) ผลิตภัณฑ์ครีมบำรุงผิวและฟื้นฟูผิวหน้า เช่น ลดเลือนริ้วรอย จุดต่างดำ เพิ่มความชุ่มชื้น ลดการเกิดสิว เป็นต้น มีผลิตภัณฑ์สกินแคร์หลายประเภทซึ่งแต่ละประเภทยังให้เลือกใช้ตามสภาพและปัญหาผิวที่แตกต่างกัน ซึ่งครีม (Cream) เป็นสกินแคร์ที่ได้รับความนิยมที่สุดและมีส่วนผสมของน้ำมันเยอะที่สุดที่ช่วยคงความชุ่มชื้นเหมาะสมสำหรับผู้ที่มีผิวแห้ง และต้องการสร้างความแข็งแรงให้กับผิวหน้า (ไลฟ์ สไตลิสซู่, 2019)

ตลาดซื้อขายสินค้าออนไลน์ (Online Marketplace) เป็นแพลตฟอร์มสำหรับการซื้อขายผลิตภัณฑ์ดูแลผิวหน้าบนโลกออนไลน์ ที่รวบรวมสินค้าและร้านค้าจำนวนมากไว้ด้วยกันผ่านเครือข่ายอินเทอร์เน็ต เปรียบเสมือนห้างสรรพสินค้าหรือร้านค้าที่เปิดตลอด 24 ชั่วโมง ไม่ว่าผู้ซื้อต้องการซื้ออะไร อยากได้สินค้าแบบไหน หรือผู้ขายอยากขายสินค้าอะไรสามารถทำได้ง่าย ๆ Online Marketplace ที่มีชื่อเสียง (สินีนาฏ เพิ่มสวัสดิ์, 2564)

การรีวิว (Review) เป็นหัวใจสำคัญของการขายของ การรีวิวจะช่วยให้ร้านค้ามีโอกาสขายของได้มากขึ้นเพราะเป็นสิ่งที่แสดงให้เห็นลูกค้าเข้าใจสินค้ามากขึ้นการรีวิวที่ดีทำให้ลูกค้ารู้สึกอยากได้สินค้ามากขึ้น และมีความเชื่อถือในสินค้าของร้านได้มากขึ้น ทั้งนี้ การรีวิวสามารถทำได้ทั้งในแบบที่ “เจ้าของร้านรีวิวเอง” และ “ให้บุคคลที่ 3 หรือลูกค้ารีวิวให้” ซึ่งผลลัพธ์ที่ได้แตกต่างกัน ซึ่งรีวิวด้วยการเขียนบรรยายถือเป็นการรีวิวที่ใครก็ทำได้ โดยไม่จำเป็นต้องกังวลว่าเขียนดีไม่ เขียนแล้วมีคนอ่านหรือไม่ หลักของการเขียนรีวิวนั้นอาศัยความจริงใจเป็นสำคัญ คือ รู้สึกต่อสินค้าอย่างไรให้เขียนบรรยายลงไปแบบนั้น อะไรดีกว่าดี อะไรไม่ดีกว่าไป อาจเลี่ยงไปก็ได้ แต่การเขียนแต่สิ่งดีๆ อย่างเดียว อาจทำให้คนอ่านรู้สึกไม่สมจริงและทำให้ความน่าเชื่อถือน้อยลง (quickservice, 2019)

การจำแนกข้อมูล (Classification) หมายถึง การจำแนกข้อมูลประกอบไปด้วยสองกระบวนการหลัก เป็นกระบวนการสร้างตัวจำแนกข้อมูลจากชุดของข้อมูลที่เป็นอินพุตที่ซึ่งแต่ละเรคคอร์ดของข้อมูลที่ทำการศึกษาประกอบไปด้วยเซตของแอทริบิวต์ที่บอกถึงคุณลักษณะของบุคคล และหมวดหมู่ของบุคคลนั้นๆ ว่ามีความปลอดภัยหรือมีประสิทธิภาพหรือไม่ โดยกระบวนการสร้างตัวจำแนกข้อมูลถูกเรียกว่า ‘learning’ หรือ ‘training’ ที่เกิดจากการนำเอาขั้นตอนวิธีสำหรับการจำแนกข้อมูลมาดำเนินการกับข้อมูล ยังมีข้อมูลอีกหนึ่งแอทริบิวต์ที่บ่งบอกถึงหมวดหมู่ของข้อมูล (class label attribute) โดยแอทริบิวต์หมวดหมู่ข้อมูลเป็นข้อมูลแบบไม่ต่อเนื่อง (discrete-valued) โดยชุดข้อมูลที่เป็นอินพุตสำหรับการสร้างตัวจำแนกข้อมูลถูกเรียกว่า “ชุดข้อมูลสำหรับสอน (training data) ตัวอย่าง (samples/instances) ชุดข้อมูล (data points) หรือ สิ่งของ (objects)” เป็นต้น (โกเมศ อัมพวัน, 2561)

การเรียนรู้ของเครื่อง เป็นที่ยอมรับในด้านการประยุกต์ใช้งานกับแก้ไขปัญหาหลายด้าน โดยแต่ละกระบวนการเรียนรู้มีความเหมาะสมกับข้อมูลในรูปแบบเฉพาะ กระบวนการเรียนรู้ของเครื่องที่แก้ปัญหาด้วยการสอนโดยใช้ข้อมูลที่มีคำตอบที่แท้จริงของปัญหา โดย Supervised Learning เป็นโมเดลที่เรียนรู้จากข้อมูลตัวอย่างที่ป้อนเข้าไป โดยข้อมูลตัวอย่างมีการระบุคำตอบอย่างชัดเจน (Labeled, Structured Data) โมเดลในกลุ่มนี้ที่พบบ่อย ได้แก่ การทำนายผล, การระบุประเภท (Classification) และ Regression โดยเทคนิคที่ใช้ในการสร้างแบบจำลองการจำแนกข้อมูล ที่ได้รับความนิยมใช้กัน ได้แก่ Naïve Bayes, Decision Tree และ Support Vector Machine (SVM) (Nessessence, 2018)

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ ผู้วิจัยได้ศึกษาค้นคว้าข้อมูล บทความทางวิชาการ เอกสารจากสื่อ และงานวิจัยที่เกี่ยวข้อง ดังรายละเอียดต่อไปนี้

2.1 ผลิตภัณฑ์ดูแลผิวหน้า (Skincare)

ผลิตภัณฑ์ที่เป็นสิ่งปรุงแต่งเพื่อใช้บนผิวหนังส่วนใดส่วนหนึ่งของร่างกาย โดยใช้ทา ถู นวด ฟน หรือโรย มีจุดประสงค์เพื่อทำความสะอาด หรือส่งเสริมให้เกิดความสวยงาม หรือเพื่อเปลี่ยนแปลงรูปลักษณ์(Lifestyleissue, 2021) ซึ่งผลิตภัณฑ์ดูแลผิวหน้าจัดอยู่ในกลุ่มเครื่องสำอาง โดยมีลักษณะต่าง ๆ ดังนี้

คุณลักษณะผลิตภัณฑ์ดูแลผิวหน้า

ในการผลิตเครื่องสำอาง มีลักษณะการเตรียมหรือการผลิตมีลักษณะที่เฉพาะเด่นชัดที่แตกต่างจากการผลิตยาอยู่ 3 ส่วน ได้แก่ 1. ผลิตภัณฑ์ที่มีกลิ่นหอมชวนดม 2. มีลักษณะสวยงาม ทั้งลักษณะของผลิตภัณฑ์ รวมถึงการบรรจุหีบห่อ 3. ใช้งานได้ง่าย สะดวกต่อการพกพาเครื่องสำอางโดยทั่วไป ต้องบอกคุณลักษณะของเครื่องสำอางนั้นๆ ไว้ ตามมาตรฐานผลิตภัณฑ์อุตสาหกรรม (มอก.) เช่น ส่วนประกอบของผลิตภัณฑ์ วิธีใช้ ข้อควรระวัง ภาชนะและการบรรจุ รวมถึงการทดสอบ การตรวจหาปริมาณและการวิเคราะห์ต่างๆ

ประโยชน์ของผลิตภัณฑ์ดูแลผิวหน้า

1. ตกแต่งให้ผิวดูเนียนใส ผุดผ่อง เช่น แป้งแต่งหน้า ดินสอเขียนคิ้ว ครีมต่างๆ
2. ทำความสะอาดรักษาอนามัย สุขภาพผิวของปากและฟัน เช่น สบู่และยาสีฟัน
3. กลบเกลื่อนให้แลดูเป็นธรรมชาติ เช่น กลบฝ้าและไฝต่างๆ
4. ตกแต่งทรงผมให้อยู่ทรง และสวยงามตามที่ต้องการ
5. ทำให้สบายผิว แก้ความอับชื้น เช่น แป้งฝุ่นโรยตัว
6. ทำให้จิตใจสดชื่น รู้สึกผ่อนคลาย เนื่องจากกลิ่นหอมของเครื่องสำอาง

ประเภทของผลิตภัณฑ์ดูแลผิวหน้า

สามารถแบ่งเป็นประเภทใหญ่ๆ ได้ 2 ประเภท คือ

1. เครื่องสำอางที่ไม่ได้ใช้แต่งสีของผิว เครื่องสำอางประเภทนี้ ใช้สำหรับการทำความสะอาดผิวหน้า หรือใช้เพื่อป้องกันผิวหน้าไม่ให้เกิดอันตรายจากสิ่งแวดล้อม เครื่องสำอางประเภทนี้ได้แก่ สบู่ แชมพู ครีมล้างหน้า ครีมกันแดด น้ำยาช่วยกระชับผิวให้ตึง เป็นต้น

2. เครื่องสำอางที่ใช้แต่งสีผิวเครื่องสำอางประเภทนี้ ใช้สำหรับการแต่งสีของผิวให้มีสีสดใสยิ่งขึ้น จากผิวธรรมชาติที่เป็นอยู่ เช่น แป้งแต่งผิวหน้า ลิปสติก เป็นต้น

ผลิตภัณฑ์ที่พบในท้องตลาดแบ่งออกเป็น 10 ประเภท ดังนี้

1. เครื่องสำอางผิวหน้า ได้แก่ ครีมทาผิว ผลิตภัณฑ์ขจัดสิว ผลิตภัณฑ์ขจัดสีผิวและขจัดฝ้า ผลิตภัณฑ์ระงับเหงื่อและขจัดกลิ่นตัว ผลิตภัณฑ์ป้องกันแสงแดด ผลิตภัณฑ์ป้องกันแมลงกัดต่อย
2. เครื่องสำอางผมและขน ได้แก่ แชมพูและครีมนวดผม ผลิตภัณฑ์ตกแต่งผม ผลิตภัณฑ์สำหรับโกนหนวดและกำจัดขน
3. เครื่องสำอางแต่งตาและคิ้ว
4. เครื่องสำอางแต่งใบหน้า ได้แก่ ผลิตภัณฑ์พอกและลอกหน้า ผลิตภัณฑ์ทกลบเกลื้อน ผลิตภัณฑ์รองพื้นแต่งหน้า แป้งผัดหน้าและแป้งโรยตัว
5. เครื่องสำอางแต่งแก้ม
6. เครื่องสำอางแต่งปาก
7. เครื่องสำอางทำความสะอาดผิว ปากและฟัน ครีมล้างหน้าและครีมล้างมือ ยาสีฟันและน้ำยาบ้วนปาก
8. เครื่องสำอางเล็บ
9. เครื่องสำอางเด็ก
10. ผลิตภัณฑ์น้ำหอม

ปี 2563 เป็นปีแห่งความท้าทายสำหรับตลาดเครื่องสำอาง ซึ่งต้องเผชิญกับภาวะ เศรษฐกิจไม่เอื้อ อย่างไม่รู้ทิศทาง จากการที่รัฐบาลของประเทศต่างๆ ได้ออกมาตรการทั้งการเงินและการคลัง เพื่อพยุงเศรษฐกิจและกระตุ้นให้เกิดการบริโภคในประเทศมากขึ้น คาดว่าทำให้เศรษฐกิจของบางประเทศที่สามารถฟื้นตัวจากโรคระบาดได้เร็ว อาทิ จีน ญี่ปุ่น เกาหลีใต้ รวมถึงประเทศเพื่อนบ้านในกลุ่ม CLMV (หากไม่เกิดการระบาดของโรคที่รุนแรงในภายหลัง) เป็นตลาดลำดับต้นๆ ที่ผู้ประกอบการเครื่องสำอางต้องเร่งเข้าไปทำตลาด ขณะเดียวกันการระบาดของโรค COVID-19 อาจส่งผลให้ผู้ประกอบการเครื่องสำอาง ใช้โอกาสนี้ปรับประสิทธิภาพการดำเนินการด้วยการบริหารจัดการต้นทุนควบคู่ไปกับการเร่งพัฒนาช่องทาง การขายสินค้าจากช่องทางหน้าร้าน (Offline) ที่เคยเป็นตลาดหลัก ขยายไปสู่ช่องทางออนไลน์ (Online) ให้มากขึ้น ซึ่งช่วยให้สินค้าเครื่องสำอางสามารถขยายโอกาสการส่งออก ภายหลังการระบาดของโรคคลี่คลายเป็นปกติได้กว้างขวางมากขึ้น ผลจากตลาดต่างประเทศมีความเชื่อมั่นและยอมรับในคุณภาพเครื่องสำอางของไทย ทั้งในส่วนที่ผลิตโดยผู้ประกอบการไทยเอง รวมถึงผู้ประกอบการจากต่างประเทศที่เข้ามาใช้ไทยเป็นฐานผลิตเพื่อการส่งออก ทำให้ไทยมีการส่งออกเครื่องสำอางไปตลาดต่างประเทศเป็นมูลค่าปีละกว่า 1 แสนล้านบาท เติบโตเฉลี่ยร้อยละ 5.0

ต่อปี ในช่วง 5 ปีที่ผ่านมา (ปี 2558-2562) แม้ในช่วงที่เงินบาทแข็งค่าในปี 2562 ที่ผ่านมา ผลกระทบต่อตลาดส่งออกยังมีน้อย การส่งออกเครื่องสำอางยังเติบโตสูงถึงร้อยละ 7.2 โดยมีตลาดส่งออกที่สำคัญคือตลาดในเอเชียที่มีสัดส่วนสูงถึงร้อยละ 82.8 (ตลาดหลัก ได้แก่ อาเซียนร้อยละ 38.9 ญี่ปุ่นร้อยละ 12.7 จีนร้อยละ 10.7 เกาหลีใต้ ร้อยละ 3.7) และตลาดที่มีศักยภาพ ได้แก่ ตะวันออกกลาง ออสเตรเลีย และตลาดสหภาพยุโรป ขณะที่ประเภท เครื่องสำอางที่ส่งออกสูง ได้แก่ สิ่งปรุงแต่งที่ใช้แต่งหน้า หรือบำรุงผิวและสิ่งปรุงแต่งที่ใช้กับผม ซึ่งมีสัดส่วนรวมกันสูงถึงร้อยละ 52.9 รองลงมาคือ วัตถุดิบที่ใช้ทำเครื่องสำอางสัดส่วนร้อยละ 18.7 สินค้าที่ใช้ชำระล้างในห้องน้ำ อาทิ สบู่ แชมพู รวมถึง วัตถุดิบที่ใช้ทำเครื่องสำอาง อาจได้รับผลกระทบน้อยกว่า เครื่องสำอางกลุ่มอื่น ที่ส่วนใหญ่เกี่ยวข้องกับ การเข้าถึงคน ได้แก่ ตลาดส่งออกหลักอย่าง จีน ญี่ปุ่น เกาหลีใต้ ซึ่งสามารถควบคุมการแพร่ระบาดของโรคได้เร็วรวมถึงประเทศเพื่อนบ้านใน CLMV ซึ่งการแพร่ระบาดยังไม่รุนแรง รวมถึงเศรษฐกิจมีการพึ่งพา ภาคราชการท่องเที่ยวไม่สูงมาก แต่คงต้องติดตามสถานการณ์ ขณะที่ตลาดยุโรปและสหรัฐฯ รวมถึงตลาด ตะวันออกกลาง อาจต้องใช้ระยะเวลาในการฟื้นตัว เนื่องจากการระบาดที่ยืดเยื้อรวมถึงมีปัจจัยที่กระทบ ต่อเศรษฐกิจซับซ้อนมากกว่า (ศูนย์วิจัยกสิกรไทย, 2563)

การระบาดของเชื้อไวรัส COVID-19 จากปีที่แล้วได้เกิดการเปลี่ยนแปลงครั้งใหญ่ใน อุตสาหกรรมความงามทั่วโลก ลูกค้านำมาให้ความสำคัญกับสุขภาพและชีวิตความเป็นอยู่ที่ดี เห็นได้ว่า ลูกค้านำมาใส่ใจดูแลผิวพรรณมากขึ้น นำไปสู่ยอดขายผลิตภัณฑ์สกินแคร์ที่เติบโต ซึ่ง คาดการณ์ว่าอุตสาหกรรมผลิตภัณฑ์บำรุงผิวจะขับเคลื่อนไปในทิศทาง ดังต่อไปนี้ (nsgcosmetics, 2021)

1. Growing Demand for Personalized Skincare ด้วยลักษณะสภาพผิวที่แตกต่างกัน ของแต่ละบุคคล ดังนั้นลูกค้าจึงต้องการการดูแลที่เหมาะสมและมีความเฉพาะเจาะจงมากขึ้นกว่าเดิม โดยผู้เชี่ยวชาญด้านการดูแลผิวออกมาให้คำแนะนำในเรื่องนี้และในหลายบริษัทเองกำลังตั้งเป้าหมาย ไปที่กลุ่มลูกค้าที่ต้องการดูแลผิวแบบส่วนตัว เช่น แอปพลิเคชันสแกนใบหน้าเพื่อแนะนำผลิตภัณฑ์ที่ เหมาะกับสภาพผิวของแต่ละบุคคล เป็นการแข่งขันหาข้อได้เปรียบให้กับผลิตภัณฑ์เพื่อตอบสนองต่อ เทรนด์นี้ ตลาดสำหรับ Personalized Beauty คาดว่าเพิ่มขึ้นจาก 38 พันล้านดอลลาร์ เป็น 72 พันล้านดอลลาร์ จากปี 2019 ถึง 2028

2. Skinfluencers Take Center Stage การโฆษณาผ่านช่องทาง Social Media อย่าง Facebook, Instagram, TikTok และ YouTube เพื่อดึงดูดใจและให้เป็นไปแบบธรรมชาติมากขึ้น มี Influencer มาช่วยในการทำการตลาด ด้วยความคิดที่ว่า Influencer เหล่านี้ก็เป็นลูกค้าของแบรนด์เอง เช่นกัน ช่วยเพิ่มความไว้วางใจและได้รับความสนใจอย่างรวดเร็ว แต่การใช้ “Skinfluencers” ให้เข้า กับแนวทางการตลาดของแบรนด์ สิ่งสำคัญคือ ต้องหา Influencer ที่เหมาะสมตรงตามค่านิยม หรืออัตลักษณ์ของแบรนด์ มีกลุ่มเป้าหมายเดียวกัน ทำหน้าที่ร่วมกัน และให้ความทุ่มเทอย่างเต็มที่ สร้างสรรค์เนื้อหาที่ครอบคลุมแบบองค์รวม

3. Creating Products that are Both Skin and Planet-Healthy ลูกค้ายุคใหม่มีความใส่ใจต่อสิ่งแวดล้อมเพิ่มขึ้น ให้ความสนใจข้อมูลบนฉลากผลิตภัณฑ์ โดยมองหาข้อมูล “Cruelty-free” และ “Natural Ingredients” ทำให้เป็นคำติดอันดับสูงสุดในการค้นหาผลิตภัณฑ์ดูแลผิว จึงได้เห็นหลายบริษัทปรับวิธีการใหม่เพื่อสื่อสารถึงการเห็นความสำคัญต่อสิ่งแวดล้อม เช่น การปรับเปลี่ยนส่วนประกอบเป็น Vegan Product การใช้บรรจุภัณฑ์ที่รีไซเคิล การเปลี่ยนพลังงานการผลิตไปใช้พลังงานลม ลดขนาดบรรจุภัณฑ์ให้เล็กหรือเปลี่ยนเป็นแบบรีฟิล และเพื่อตอบสนองความต้องการด้านสิ่งแวดล้อมของลูกค้าได้อย่างมีประสิทธิภาพผลิตภัณฑ์ควรมีความโปร่งใส มีข้อมูลฉลากที่เข้าใจได้ง่าย ตรวจสอบแหล่งที่มาของส่วนผสมและบรรจุภัณฑ์ กระบวนการจัดส่ง กระบวนการจัดการของเสียในกระบวนการผลิตได้

4. Skincare Below the Surface ลูกค้ายุคใหม่มีความเข้าใจความเชื่อมโยงระหว่างสิ่งที่รับประทานเข้าไปภายในร่างกายส่งผลต่อสุขภาพผิว ทำให้แนวโน้มของผลิตภัณฑ์ดูแลผิวจากภายในพร้อมเติบโตอย่างต่อเนื่อง ที่เห็นได้ชัดเจนคือ Collagen Peptide ในอนาคตตลาดสกินแคร์อาจต้องมองหาผลิตภัณฑ์ที่สามารถให้คุณค่าได้เพิ่มเติมนอกเหนือจากคอลลาเจน มุ่งเน้นไปที่การผสมผสานสารสำคัญ และให้คุณประโยชน์แบบองค์รวมมากขึ้น ซึ่งรวมถึงสุขภาพผิวด้วย

5. DIY Skincare จากการกักตัวอยู่บ้าน ผู้คนเริ่มหาซื้อหรือดัดแปลงอุปกรณ์ต่างๆจากภายในบ้าน ซึ่งเกิดกับการดูแลผิวด้วยเช่นเดียวกัน แต่ก่อนนั้นเครื่องมือการดูแลผิวพบได้ตามคลินิกหรือสปา แต่ด้วยเทคโนโลยีที่ทันสมัยในปัจจุบันสามารถใช้เครื่องมือเหล่านี้ได้เองที่บ้าน เช่น เครื่องนวดหน้า อุปกรณ์การบำบัดด้วยแสง และเมื่อสถานการณ์โรคระบาดรุนแรง นวัตกรรมในการดูแลผิวใหม่ๆต่างเป็นที่ต้องการมากขึ้น การเข้าใจลูกค้ามุ่งมั่นตอบสนองและแก้ปัญหาให้ลูกค้าเพื่อตอบสนองต่อเทรนด์นี้ได้เป็นสิ่งสำคัญ

ประเภทของผลิตภัณฑ์ดูแลผิวหน้า

ผลิตภัณฑ์ดูแลผิวหน้ามีความจำเป็นอย่างยิ่งในยุค COVID-19 เพราะทุกคนต่างให้ สกินแคร์ (Skincare) หรือผลิตภัณฑ์บำรุงผิวมีมากมายหลายประเภท และแต่ละประเภทมีแยกไปอีกหลายรูปแบบให้เลือกใช้ตามสภาพผิวและปัญหาผิวที่แตกต่างกันไป โดยมีผลิตภัณฑ์ในรูปแบบต่าง ๆ ดังนี้ (Kaewsiri Srisam-ang, 2020)

1. Cleanser กฎพื้นฐานอย่างหนึ่งที่ชาวเกาหลีปฏิบัติตามอย่างขยันขันแข็งคือกิจวัตรการทำความสะอาด และไม่ใช่แค่การ “ทำความสะอาด” ธรรมดา แต่คือ “การทำความสะอาดสองครั้ง” ปัญหาผิวส่วนใหญ่เกิดจากผิวสกปรก ซึ่งต้องล้างหน้าทุกวันวันละ 2 ครั้งดูสะอาดหมดจด แต่สิ่งที่เห็นมีเพียงเครื่องสำอางที่หายไป ไม่สามารถมองเห็นสิ่งที่อยู่ลึกลงไปใต้รูขุมขนเหล่านั้น เห็นได้ชัดว่าการทำความสะอาดสองครั้งหมายถึงการล้างหน้าสองครั้ง แต่ในครั้งแรกคือให้ล้างเครื่องสำอางชั้นพื้นฐานออกไป และครั้งที่สองต้องล้างทำความสะอาดสิ่งสกปรกและน้ำมันอุดตันรูขุมขนบนใบหน้า

2. Toner เมื่อทำความสะอาดใบหน้าเสร็จแล้วต้องใช้โทนเนอร์ เพราะรูขุมขนที่เปิดอยู่มีแนวโน้มที่ดูดซับสิ่งสกปรกและฝุ่นละอองและนี่คือเหตุผลที่โทนเนอร์ต้องเป็นส่วนสำคัญในขั้นตอนการดูแลผิว โทเนอร์ยังสามารถมีประโยชน์อื่น ๆ เช่น การผลิตเซลล์ผิวและการล้างเครื่องสำอางที่ยังติดอยู่
3. Serum / Essence ตอนนี้ผิวได้รับการทำความสะอาดและเตรียมการแล้วซึ่งได้เวลาให้สารอาหารที่จำเป็นมากในรูปแบบของเซรัมที่อุดมสมบูรณ์ เซรัมเปรียบเสมือนยาอายุวัฒนะสำหรับผิวแห้งและผิวที่แก่ก่อนวัยเซรัมสามารถซึมลึกเข้าสู่ภายในได้ การใช้เซรัมเป็นประจำช่วยให้ผิวมีความชุ่มชื้นนุ่มนวลเนียนนุ่ม
4. Cream หลังจากเซรัมต้องมีครีมบำรุงผิวหน้าที่ดีที่ช่วยให้ผิวชุ่มชื้นและคงไว้ได้อย่างนั้นตลอดทั้งวันแน่นอนว่าไม่ต้องการครีมที่ทำให้ผิวแห้งภายใน 2-3 ชั่วโมงและต้องทาซ้ำ ครีมที่ดีควรมีการเพิ่มส่วนผสมที่ช่วยบำรุงผิวและปลอดภัยด้วย
5. Mask มาส์กเป็นส่วนสำคัญในการดูแลผิวเป็นเหมือนการบำรุงเพิ่มเติมที่สามารถทำให้ผิวหมองคล้ำ มีชีวิตชีวาดูแลปัญหาบางอย่างและทำให้ผิวนุ่มและเปล่งประกาย ใช้มาส์กหน้าอย่างน้อยสัปดาห์ละครั้งและถ้าเป็นไปได้ให้บ่อยขึ้น ทาบาง ๆ บนผิวที่สะอาดทิ้งไว้ประมาณ 10-15 นาที
6. Exfoliating คือการขัดผิวขจัดเซลล์ผิวที่ตายแล้วออกจากชั้นบนสุดของผิวทำให้ผิวเรียบเนียนและยังช่วยป้องกันไม่ให้รูขุมขนอุดตัน ให้ขัดผิว 2-3 ครั้งต่อสัปดาห์ เพื่อผลลัพธ์ที่ดีที่สุดหลังการทำความสะอาด

Forecast แนวโน้มผลิตภัณฑ์ดูแลผิวหน้าปี 2564 (comets, 2020)

1. Minimalism สกินแคร์เน้นความเรียบง่าย ส่วนผสมจากธรรมชาติ เป็นผลิตภัณฑ์โฮมเมดหรือสินค้าออร์แกนิก ไม่มีสารเคมี เนื่องจากมีการวิจัยของสารเคมีที่มีต่อร่างกายอยู่เสมอ ดังนั้นลูกค้ายุคใหม่จะระมัดระวังเกี่ยวกับส่วนผสมในผลิตภัณฑ์ที่มีผลกระทบต่อสุขภาพ
2. Trendy Skin Care Ingredients ส่วนผสมที่เป็นมิตรกับสิ่งแวดล้อม เช่น น้ำมัน CBD ได้รับความนิยมมากขึ้นในช่วงหลายปีที่ผ่านมา นอกจากสรรพคุณทางการแพทย์ CBD ในผลิตภัณฑ์บำรุงยังมีสรรพคุณช่วยลดการระคายเคืองผิวหนัง รักษาสิวและยังให้ความชุ่มชื้น ต่อต้านริ้วรอย ถูกแทนที่ด้วย bakuchiol สารสกัดจากพืชที่ไม่ทำให้ผิวระคายเคือง มีความปลอดภัยสูงกว่า ซึ่งกรด glycolic (AHA) และกรดซาลิไซลิก (PHA) ที่อ่อนโยนกว่า
3. Personalized Solution เทรนด์ดูแลผิวที่ตอบโจทย์ได้ตรงจุดมากยิ่งขึ้นกำลังมาแรง ลูกค้าหันมาโฟกัสปัญหาผิวตามลักษณะของแต่ละบุคคล ไม่ได้เจาะจงที่ความขาวหรือผลิตภัณฑ์ Whitening แล้ว ผลิตภัณฑ์บำรุงผิวที่เจาะจงวัตถุประสงค์ เช่น สกินแคร์เพื่อการผลัดเซลล์ผิว (Exfoliation) เติมความชุ่มชื้น (hydrating) รักษาสิว (Anti-Acne) ต่อต้านริ้วรอย (Anti-Aging) มาส์กหน้า น้ำยาล้างเครื่องสำอาง ครีมบำรุงรอบดวงตา เป็นต้น

งานวิจัยเกี่ยวกับการผลิตภัณฑ์ดูแลผิวหน้า

อธิปดี คำเจริญ และสุมาลย์ ปานคำ (2563) รูปแบบความสัมพันธ์เชิงสาเหตุ ความตั้งใจซื้อครีมยูมิคซ์ สกินแคร์ ผ่านเพจเฟซบุ๊ก ของลูกค้าในเขตกรุงเทพมหานครและปริมณฑล งานวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนารูปแบบความสัมพันธ์เชิงสาเหตุความตั้งใจซื้อ ครีมยูมิคซ์ สกินแคร์ ผ่านเพจเฟซบุ๊กของลูกค้าในเขตกรุงเทพมหานครและปริมณฑล และเพื่อ ตรวจสอบความสอดคล้อง ของรูปแบบความสัมพันธ์เชิงสาเหตุที่พัฒนาขึ้นกับข้อมูลเชิงประจักษ์ เครื่องมือที่ใช้ในการวิจัย ได้แก่ แบบสอบถามออนไลน์ กลุ่มตัวอย่าง ได้แก่ ผู้ที่เคยซื้อครีมยูมิคซ์ สกินแคร์ ผ่านเพจเฟซบุ๊ก umixskincare ซึ่งได้มาจากการสุ่มตัวอย่างแบบสะดวก จำนวน 400 คน สถิติ ที่ใช้ในการวิเคราะห์ ข้อมูล ได้แก่ การวิเคราะห์โมเดลสมการโครงสร้าง พบว่า รูปแบบความสัมพันธ์เชิงสาเหตุความตั้งใจซื้อครีมยูมิคซ์ สกินแคร์ผ่านเพจเฟซบุ๊กของลูกค้าในเขตกรุงเทพมหานครและปริมณฑล ประกอบด้วย 5 ปัจจัย ได้แก่ 1) การตลาดผ่านสื่อสังคมออนไลน์ 2) คุณค่าตราสินค้า 3) มูลค่าสินค้า 4) คุณค่าความสัมพันธ์ และ 5) ความตั้งใจซื้อ ซึ่งปัจจัยแต่ละปัจจัยส่งผลต่อความตั้งใจซื้อครีมยูมิคซ์ สกินแคร์ ผ่านเพจเฟซบุ๊ก สามารถอธิบายความแปรปรวนของรูปแบบ ได้ร้อยละ 93 พบว่า ด้านคุณค่าความสัมพันธ์ส่งผลต่อความตั้งใจซื้อครีมยูมิคซ์ สกินแคร์ผ่านเพจเฟซบุ๊กมากที่สุด

ณัฐพร เหล่าสีพงษ์ (2562) กลยุทธ์การจัดการทางธุรกิจของผลิตภัณฑ์อินทรีย์ (organic) ประเภทเครื่องสำอางและสกินแคร์ในเขตกรุงเทพมหานคร มุมมองจากลูกค้า งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาความแตกต่างของความคิดเห็นใน กระบวนการตัดสินใจ พฤติกรรม และความคิดเห็นด้านปัจจัยส่วนประสมทางการตลาดที่มีผลต่อพฤติกรรมการซื้อของลูกค้ากลุ่มตัวอย่างที่เคยซื้อและไม่เคยซื้อผลิตภัณฑ์อินทรีย์ประเภทเครื่องสำอางและสกินแคร์ ในเขต กรุงเทพมหานคร ทำการวิจัยในเชิงปริมาณเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 250 คน โดยการสุ่มแบบบังเอิญและนำมาวิเคราะห์ข้อมูลทางสถิติ จากการศึกษาวิจัย พบว่า ปัจจัยที่มีผลต่อพฤติกรรมการซื้อของกลุ่มตัวอย่างที่เคยซื้อผลิตภัณฑ์อินทรีย์ (Organic) ประเภทเครื่องสำอางและสกินแคร์ คือ ปัจจัยด้านผลิตภัณฑ์ โดยเฉพาะผลิตภัณฑ์ที่มีคุณสมบัติที่ดีและมีความแตกต่างไม่เหมือนใคร ทำให้ความถี่ในการซื้อของลูกค้าเพิ่มมากขึ้น ดังนั้นต้องผลิตผลิตภัณฑ์ที่ดีและมีความแตกต่าง ส่วนกลุ่มตัวอย่างที่ไม่เคยซื้อผลิตภัณฑ์ค่านึงถึงปัจจัยด้านช่องทางการจัดจำหน่าย โดยสามารถหาซื้อผลิตภัณฑ์ได้ง่าย รวมถึงด้านผลิตภัณฑ์ ต้องมีการชี้แจงส่วนประกอบชัดเจน และมีการส่งเสริมการตลาดโดยมีส่วนร่วมในการซื้อผลิตภัณฑ์

จากการพิจารณาบทความและงานวิจัยนั้น นักวิจัยได้ใช้ข้อมูลจากแบบสอบถามและมุ่งเน้นไปที่หาความสัมพันธ์เชิงสาเหตุ ว่ามีปัจจัยใดบ้างที่ส่งผลต่อความตั้งใจซื้อครีมยูมิคซ์ สกินแคร์ และปัจจัยส่วนประสมทางการตลาดที่มีผลต่อพฤติกรรมการซื้อของลูกค้ากลุ่มตัวอย่างที่เคยซื้อและไม่เคยซื้อผลิตภัณฑ์อินทรีย์ (Organic) ประเภทเครื่องสำอางและสกินแคร์ เป็นต้น แต่ยังไม่ได้มุ่งเน้นไปที่การวิจัยผลิตภัณฑ์

ดูแลผิวหน้าที่มีผลต่อการตัดสินใจซื้อหรือไม่ ทำให้ผู้วิจัยมีความสนใจในส่วนของ การรีวิวลิตภัณฑ์ดูแลผิวหน้าว่ามีผลต่อการตัดสินใจซื้อของลูกค้าหรือไม่ และมีรีวิวลิตที่มีความน่าสงสัยหรือแนวโน้มความเสี่ยงเพียงใดจากการสั่งซื้อสินค้าผ่านทางช่องทาง Social Media หรือ e-Marketing

2.2 ตลาดซื้อขายสินค้าออนไลน์ (Online Marketplace)

การซื้อขายสินค้าแบบอีคอมเมิร์ซและโซเซียลคอมเมิร์ซ ต่างกลายมาเป็นช่องทางการขายที่คนไทยหันมาใส่ใจและเปิดใจใช้งานกันมากขึ้น ส่วนหนึ่งเพราะต้องการลดการเดินทางจากการแพร่ระบาดของโควิด-19 อีกส่วนหนึ่งมาจากพฤติกรรมที่ปรับเปลี่ยนไป คือ อยากรลดเวลาในการออกไปเดินหาสินค้า มาเป็นค้นหาผ่านปลายนิ้วมากขึ้น (ธนวัฒน์ มาลาบุปผา, 2564)

ภาพรวมวงการ E-commerce ของปี 2021

ภาพรวมของ Digital ประเทศไทยปี 2021 นี้ เห็นว่าตัวเลขไปในทิศทางเดียวกันว่าคนไทยยังใช้เวลากับอินเทอร์เน็ตมากยิ่งขึ้น จากข้อมูลพบว่า 74 เปอร์เซ็นต์ของผู้ใช้อินเทอร์เน็ตชาวไทยยอมรับว่าใช้เวลากับอินเทอร์เน็ตเฉลี่ย 8 ชั่วโมงต่อวันแล้ว แปลว่ายังมีคนเพิ่มเวลาใช้อินเทอร์เน็ต ซึ่งกิจกรรมหลักที่คนไทยใช้กับอินเทอร์เน็ต โดยใช้เกี่ยวกับการเช็คข้อมูลออนไลน์ ส่วน“ชอปปิง” ซึ่งสัดส่วนของช่องทางการขายสินค้าออนไลน์ พบว่า มาร์เก็ตเพลส มีจำนวน 32% ซึ่งเป็นช่องทางที่มีความนิยมมากที่สุด และเหตุผลที่ผู้บริโภคมาจับจ่ายสินค้าออนไลน์มากขึ้น ดังภาพที่ 1



ภาพ 1 ภาพรวมการใช้ดิจิทัลและอินเทอร์เน็ตของไทย
ที่มา : <https://www.thumbsup.in.th/ecommerce-2021>

การปรับตัวของธุรกิจค้าปลีก และการแข่งขันของแพลตฟอร์ม

ภาพรวมของตลาดอีคอมเมิร์ซประเทศไทยในปี 2021 อ้างอิงจาก Google และ Temasek ที่ทำข้อมูลตลาดดิจิทัล พบว่า มูลค่าตลาดอีคอมเมิร์ซอยู่ที่ประมาณ 4 แสนล้านบาทปรับตัวขึ้นจากปีที่แล้วและปี 2021 นี้ เติบโตขึ้นประมาณ 75% ไปแตะที่ประมาณ 7 แสนล้านบาท เทียบกับก่อนหน้าเมื่อปี 2019-2020 ที่โควิดเข้ามามีผลกระทบอย่างจังกับธุรกิจ

ปี 2021 อีคอมเมิร์ซไทยโต 140% ทำให้คาดการณ์ว่าในมุมของมาร์เก็ตเพลสยังเป็นแพลตฟอร์มที่มีอิทธิพลเป็นหลัก มีสัดส่วนถึงเกือบครึ่งเลยทีเดียว ผู้เล่น 3 รายหลักอย่าง ช้อปปี้ เป็นปีที่เติบโตเร็วมาก พบว่า ปีที่แล้วเป็นปีจุดตัดที่ยอดขายแข่งลาซาด้าและปีนี้ช่องว่างนั้นขยายห่างขึ้น แต่ขณะเดียวกันลาซาด้าพยายามรุกหนักมากในขยายกิจการต่างๆ ขณะที่ช้อปปี้มีการขึ้นค่าบริการทางฝั่งผู้ขาย ผู้ขายจำนวนมากทุ่มน้ำหนักไปทางฝั่งลาซาด้า ในส่วนของทางฝั่งของ JD Central เห็นว่ามีความเคลื่อนไหวมากขึ้น แต่ยังห่างจากผู้เล่นทางฝั่งของลาซาด้ากับช้อปปี้แต่ยังมีอีกจุดหนึ่งที่น่าจับตามอง คือ อีมาร์เก็ตเพลส ซึ่งมีอิทธิพลกับนักช้อปมากยิ่งขึ้น ดังภาพที่ 2



ภาพ 2 อิทธิพลของอีมาร์เก็ตเพลส

ที่มา : <https://www.thumbsup.in.th/ecommerce-2021>

Social Commerce เทรนด์มาแรง คู่แข่งอีคอมเมิร์ซ

คาดการณ์ว่าโซเชียลคอมเมิร์ซกินสัดส่วนประมาณ 40 เปอร์เซ็นต์ของมูลค่าอีคอมเมิร์ซ 7 แสนล้านบาทในปีนี้ พฤติกรรมการซื้อสินค้าแบบ Social commerce เห็นได้จาก Digital User ว่าคนไทยใช้เวลาอยู่กับโซเชียลมีเดียแพลตฟอร์มเยอะมากๆ เห็นว่าโซเชียลมีเดียแพลตฟอร์มต่างๆ มีแนวโน้มที่

ขยับขยายตัวไปทำตัวเองเป็นคอมเมิร์ซแพลตฟอร์มมากขึ้นเรื่อยๆ ไม่ว่าจะ เป็น TikTok ที่ปีนี้มีการประกาศฟีเจอร์ที่เกี่ยวกับคอมเมิร์ซมากขึ้น อ้างอิงที่อินโดนีเซีย Tiktok มีการเปิดทำให้แอปเป็นคอมเมิร์ซแพลตฟอร์มเป็นที่เรียบบ่อยแล้วแปลว่าเวลาคนอยากซื้อสินค้าในเวลาดูเสร็จสามารถกดซื้อจบบนในฟอร์มของตัวเองได้เลย คาดการณ์ว่าคอมเมิร์ซแพลตฟอร์มกลายเป็นเทรนด์ของปีหน้าเลยละกันว่าแนวโน้มของโซเชียลคอมเมิร์ซเป็นเทรนด์ในการที่ซื้อของต่างๆ ครอบจบใน Social Media Platform ไปเลย

Facebook กับการรุกตลาดคอมเมิร์ซ

ทิศทางของคอมเมิร์ซเป็นทิศทางหลักที่เฟซบุ๊กผลักดันมาตลอด แต่ค่อนข้างช้าเหมือนกันนะ เพราะผลักดันไปในภาพรวมหลายแพลตฟอร์ม ความเคลื่อนไหวล่าสุดเห็นว่า Facebook Group เริ่มมีฟีเจอร์ในเรื่องของ commerce ผนวกเข้าไปด้วย พยายามปรับให้ตัว Facebook Group กลายเป็นคอมมูนิตี้ของคนที่มาทำการซื้อขายกันได้ครบจบอยู่ในคอมมูนิตี้ แทบไม่ใช้โซเชียลมีเดีย แต่เป็นโคเรกซ์ชันใหม่ในการหารายได้ เป็นการเสริมฟีเจอร์ทำให้ประสบการณ์ในการซื้อของ seamless หรือว่าไร้รอยต่อมากยิ่งขึ้น ณ ปัจจุบันโซเชียลมีเดียแพลตฟอร์ม คือเรื่องของโซเชียลคอนเนกชัน เรื่องของการเชื่อมโยงกัน ระหว่าง คอนเทนต์ คือเอา Content เป็นตัวนำ เป็น Social Content based ในฝั่งของ commerce ยังไม่ได้มีฟีเจอร์ที่ทำให้การซื้อขายไร้รอยต่อสักทีเดียว ปกติเวลาอยากซื้อสินค้าบางทีในหลายๆ กรณี ต้องมีการทักแชทไปหรือบางทีต้อง Add LINE คุยกัน หรือต้องกดออกไปซื้อที่เว็บ กดออกไปซื้อที่มาร์เก็ตเพลสซึ่งมีแรงเสียดทานอยู่ ทำให้มุมของนักช้อปหลุดโอกาสทางการซื้อขายออกไปได้ สิ่งที่โซเชียลมีเดียแพลตฟอร์มพยายามทำคือ ลดแรงเสียดทานลง ยิ่งลดลงได้เท่าไรจะเป็นผลประโยชน์ต่อนักการตลาดหรือมุมของผู้ขายที่ลงโฆษณาให้กับโซเชียลมีเดียแพลตฟอร์มนั่นเองยิ่งทำให้ประสิทธิภาพของโฆษณาดีขึ้น ขายของดีขึ้น ทำให้มีแนวโน้มที่นักการตลาดโยกงบโฆษณามาลงทุนกับโซเชียลมีเดียแพลตฟอร์มมากขึ้น มูลค่าตลาดอีคอมเมิร์ซปีนี้อยู่ที่ 7 แสนล้าน เกินสัดส่วนอยู่ประมาณสัก 5 – 10 เปอร์เซ็นต์ของมูลค่าปลีกทั้งประเทศ เห็นได้ว่าสัดส่วนตรงนี้น้อยน้อยเมื่อเทียบกับมูลค่าค้าปลีกมวลรวมทั้งประเทศ

พฤติกรรมกรรมการซื้อของผู้บริโภคมีตั้งแต่ขั้นตอนการสรรหาสินค้าที่อยากซื้อ การตัดสินใจซื้อเข้ามาเมื่อมีอิทธิพลตั้งแต่ต้นกระบวนการ การหาแรงบันดาลใจ ตั้งแต่การหาข้อมูลประกอบการตัดสินใจซื้อ การเปรียบเทียบแล้วตัดสินใจซื้อ ตอนนี้แนวโน้มที่มีอิทธิพลที่สุดคือขั้นการหาแรงบันดาลใจ แล้วการตัดสินใจ ในการซื้อคนส่วนใหญ่ยังคงอยู่กับช่องทางออฟไลน์ แต่แนวโน้มเติบโตสูงขึ้น เพราะฉะนั้นการขายสินค้าผ่านหน้าร้านไม่ได้หายไป แต่กลับกลายเป็นว่าจะส่งเสริมซึ่งกันและกันด้วยซ้ำ โดยธุรกิจที่สามารถออกแบบประสบการณ์การซื้อให้ส่งเสริมการตัดสินใจซื้อได้ทั้งผ่านช่องทางของออนไลน์ ดิจิทัลแพลตฟอร์มต่างๆ รวมทั้งหน้าร้าน มีแนวโน้มที่ได้ใจลูกค้าไปมากกว่าที่ไปก็สไปแค่ช่องทางใดช่องทางเดียว เป็นเรื่องของออมนิแชนแนล

วิเคราะห์ 3 แพลตฟอร์ม

ในมุมมองของผู้เล่น marketplace 3 รายหลักๆ เจ้าที่แข่งขันกันดุเดือด คือ ลาซาด้ากับช้อปปี้ เชื่อว่าปี 2565 ทั้ง 2 รายนี้ยังแข่งขันกันดุเดือดอยู่ แต่ยังไม่มีการที่แบบชิงตลาดไปแล้วแข่งโดดไปเลย เหตุผลเพราะว่า Lazada มองว่าตลาดประเทศไทย เป็นตลาดที่สำคัญตลาดหนึ่งเลย Lazada เสียโอกาสในตลาดใหญ่ๆ ตลาดหนึ่งไปในภูมิภาคนี้คือ อินโดนีเซีย ที่อินโดนีเซีย Lazada ถูก Shopee ทิ้งห่างไป ต้องยอมรับว่า Shopee เป็นที่นิยมของทั้งผู้ซื้อผู้ขายในตลาดประเทศไทย ดังภาพที่ 3



ภาพ 3 ช่องทางอีมาร์เก็ตเพลสที่นิยมซื้อขายมากที่สุด

ที่มา : <https://www.thumbsup.in.th/ecommerce-2021>

Quick Commerce คือการซื้อสินค้าผ่าน On demand platform ต่างๆ ไม่ว่าจะเป็น Grab LINEMAN FoodPanda ที่ทุกวันนี้เปิด marketplace ของตัวเองแล้ว เชิญชวนให้ร้านค้าต่างๆที่เป็นร้านค้าแบบหน้าร้านให้เอาสินค้าเข้ามาลง สามารถซื้อขายสินค้าผ่านทางแพลตฟอร์มของตัวเองได้เลย จุดเด่นคือได้สินค้าเร็ว เชื่อว่าแนวโน้ม 3 ช่องทางนี้จากยังเป็นช่องทางที่น่าจับตามองแล้วเป็นช่องทางหลักๆ ของ e-commerce ประเทศไทยในปี 2565 กลุ่มสินค้ามาแรงที่ยังครองใจนักช้อปภาพรวมกลุ่มสินค้าที่เป็นที่นิยมในช่วงโควิดนี้ คือสินค้าโฮมแอนด์ลิฟวิ่งเลย ของใช้ในบ้านเฟอร์นิเจอร์ของตกแต่งบ้าน เกี่ยวกับใช้ชีวิตการใช้ชีวิตที่บ้านที่ทำงานการทำงานที่บ้านให้เป็นที่ทำงานได้ ส่วนสินค้าเกี่ยวกับสุขภาพยังคงมีการเติบโตสูงเหมือนกัน

สิ่งที่เกิดขึ้นในปี 2565

ตลาดอีคอมเมิร์ซพิสูจน์ตัวเองแล้วว่า ไม่ได้เป็นช่องทางเลือกอีกต่อไป จะเป็นช่องทางที่สำคัญที่ไม่ว่าจะเป็นค้าปลีก แทบทุยกู้หื้อเปิดช่องทางอีคอมเมิร์ซเป็นของตัวเอง ในปี 2565 จะลงทุนเพิ่มไปมากกว่าปัจจุบันอีกสองคือแบรนด์ ซึ่งการสร้างช่องทางขายของตัวเองทำให้เกิดความซับซ้อนของช่องทางในการขายสินค้า เริ่มซับซ้อนมากยิ่งขึ้น และยังเชื่อว่าตลาดดิจิทัลเม็ดเงินโฆษณาดิจิทัลเติบโตขึ้นอีกแน่นอน ซึ่งในปี 2565 นี้เป็น 10-11 ชั่วโมงและวัตถุประสงค์ของเม็ดเงินโฆษณาดิจิทัลที่ต้องจัดเน้นไปที่ conversion หรือว่ายอดขายมากขึ้น ส่งผลไปถึงตลาดของ influencer Marketing ที่เน้นในการวัดผลไปถึงเรื่องของการยอดขายมากขึ้น ทำให้แบรนด์เวลาตัดสินใจจ้างอยากวัดผลได้ง่ายเงินตามผลงานคิดว่าปีหน้าจะเป็นปีที่สำคัญของ affiliate Marketing ประเทศไทยที่มีอัตราการเติบโตที่สูงขึ้น ตลาดอีคอมเมิร์ซจะเป็นตลาดที่เติบโตไปแน่นอน เพราะมีการขับเคลื่อนกันระหว่างแพลตฟอร์มที่ต่อเนื่อง ไม่ใช่เฉพาะมาร์เก็ตเพลสที่นำโดยประเทศจีนแต่จะเป็นแพลตฟอร์มของฝั่งอเมริกาด้วย ไม่ว่าจะเป็น Google Search, Social Media ต่างๆ จะมีการเพิ่มพีเจอร์เกี่ยวกับ commerce เข้ามามากยิ่งขึ้นแน่นอน และช่องทางการขายยิ่งแตกออกไปเยอะมาก ในมุมมองของผู้ขายเป็นทั้งโอกาสและความท้าทายอยู่ที่การปรับตัวและเรียนรู้ในการเอาช่องทางต่างๆ มาใช้

มูลค่าของตลาดขายสินค้าออนไลน์ ปี 2020

การช้อปปิ้งออนไลน์ และสถานการณ์การแพร่ระบาดโควิด 19 (COVID-19) ในประเทศไทยเป็นแรงผลักดันให้ลูกค้าหันมาช้อปปิ้งสินค้าผ่านช่องทางออนไลน์มากขึ้น ส่งผลให้ธุรกิจกลุ่ม E-commerce มีแนวโน้มเติบโตขึ้นทุกปี คาดการณ์มูลค่าตลาด E-commerce ปี 2020 อาจมีมูลค่าสูงถึง 220,000 ล้านบาท หรือเติบโตขึ้น 35% จากปีก่อนหน้าเลยทีเดียว จากการแข่งขันทางธุรกิจกลุ่ม E-commerce ที่มากขึ้นทุกปี จาก Thailand E-commerce Landscape ปี 2020 สามารถแบ่งผู้เล่นทั้งหมดออกเป็น 4 ส่วนใหญ่ๆ ได้แก่ (Techsauce, 2020)

1. เครื่องมือการทำงานการตลาด (Marketing Tools)

กว่าที่แบรนด์จะเป็นที่รู้จักและประสบความสำเร็จนั้น ไม่ได้ใช้เวลาเพียงแค่วันสองวันจะสามารถโด่งดังเป็นพลุแตกได้ แต่แบรนด์ต้องมีการวางแผนการตลาดต่างๆ ไม่ว่าจะเป็นการลงโฆษณา การคิดรายการส่งเสริมการขาย การบริหารจัดการดูแลฐานลูกค้าที่ดี วางแผนช่องทางในการโปรโมท หรือจะเป็นประชาสัมพันธ์ให้แบรนด์เป็นที่รู้จักต่อลูกค้ามากขึ้น เพราะสิ่งเหล่านี้ล้วนแล้วแต่เป็นองค์ประกอบสำคัญที่ทำให้แบรนด์ของคุณประสบความสำเร็จ ดังนั้นการเลือกใช้เครื่องมือการทำงานการตลาดต่างๆ จึงสำคัญกับแบรนด์ผู้ผลิตเป็นอย่างมาก

- **อีคอมเมิร์ซ (E-commerce Channel)**

เมื่อแบรนด์ผู้ผลิตได้มีการเลือกใช้เครื่องมือในการทำ Marketing เรียบร้อยแล้ว มาในส่วนของช่องทางการขายผ่านแพลตฟอร์ม E-commerce ได้เลย เรียกได้ว่า ช่องทาง E-commerce เป็นหนึ่งในช่องทางที่ผู้ผลิตต่างๆ ควรเข้ามาแข่งขันในช่องทางนี้ได้แล้ว เนื่องจากว่าปัจจุบันเทคโนโลยีและความก้าวหน้าของอินเทอร์เน็ตทำให้ไม่มีช่องว่างระหว่างผู้ค้าและผู้ซื้ออีกต่อไป ดังภาพที่ 4



ภาพ 4 E-commerce Channels (Marketplace)

ที่มา : <https://contentshifu.com/news/thailand-ecommerce-landscape-2020>

เมื่อแบรนด์เลือกใช้เครื่องมือในการทำการตลาด อีกหนึ่งช่องทางขายออนไลน์ที่เป็นที่นิยมในประเทศไทยเป็นอย่างมาก คือ E-commerce ในที่นี้ โพรซ์ชา แบ่งออกเป็น 3 ช่องทาง ได้แก่

1. Marketplace

C2C ให้บริการติดต่อซื้อขายระหว่างพ่อค้าแม่ค้ารายย่อยกับลูกค้า

เจ้าใหญ่ผู้นำตลาด Marketplace อย่าง Lazada และ Shopee นอกจากนี้ยังมีแพลตฟอร์ม Marketplace C2C ที่รวบรวมสินค้าเกษตรกรไทยอย่าง Thailand postmart ที่เป็นช่องทางการกระจายสินค้าเกษตร สินค้าชุมชน และสินค้าโอท็อปที่ใหญ่ที่สุดในประเทศไทย พร้อมให้บริการจัดส่งถึงมือลูกค้าในระยะเวลา 1 – 2 วันทำการ โดยลูกค้าสามารถตรวจสอบรายละเอียดของสินค้า และสั่งซื้อ พร้อมชำระเงินค่าสินค้าออนไลน์ได้ทันที

B2C ให้บริการติดต่อซื้อขายระหว่างธุรกิจผู้จัดจำหน่ายสินค้าโดยตรงถึงลูกค้า

ตัวอย่างเช่น LazMall / Shopee Mall / JD Central / WeMall / NocNoc

B2B ให้บริการติดต่อซื้อขายระหว่างหน่วยงานธุรกิจถึงหน่วยงานธุรกิจ

ตัวอย่างเช่น OfficeMate / pantavanij / thaitrade

Cross-boarder

ซื้อขายสินค้าข้ามพรมแดนยักษ์ใหญ่ของจีนอย่าง Alibaba และแพลตฟอร์มอื่น เช่น ebay / jd.com / amazon / Taobao

2. Social Commerce

ธุรกิจออนไลน์ไทย 95% ขายสินค้าผ่าน Social Commerce ซึ่งเป็นค่าเฉลี่ยมากที่สุดในโลก แพลตฟอร์มที่นิยมเป็นอย่างมากคงหนีไม่พ้น Facebook / Twitter / IG ทั้งยังมีบริการใหม่อย่าง Facebook Marketplace บริการที่ให้สมาชิกสามารถค้นหาและโพสต์สินค้าและบริการที่ต้องการซื้อหรือขายระหว่างผู้คนที่อยู่ในละแวกเดียวกัน โดยใช้การส่งข้อความผ่านบริการ Messenger เพื่อติดต่อและนัดพบกันตามสถานที่ต่างๆ ที่ใช้ในการแลกเปลี่ยนซื้อขายสินค้าหรือบริการนั้น นอกจากนี้แพลตฟอร์มที่ให้บริการด้าน Chatbot ดูแลเรื่องการตอบข้อความ อำนวยความสะดวกให้ผู้ค้ามีให้เลือกใช้อีกมากมาย เช่น Chatpify/ Chatfuel และบริษัทสัญชาติไทยอย่าง botio และยังมีบริการด้าน Order Management ดูแลจัดการเรื่องสต็อกสินค้าให้อย่าง xCommerce / Zort / Page365 เป็นต้น

3. e-Tailer

การขายสินค้าและบริการให้กับลูกค้าโดยตรง ซึ่งนั่นคือจะไม่ผ่านแพลตฟอร์มกลาง โดยแบ่งเป็น 3 ประเภท ดังภาพที่ 5 และภาพที่ 6 ได้แก่

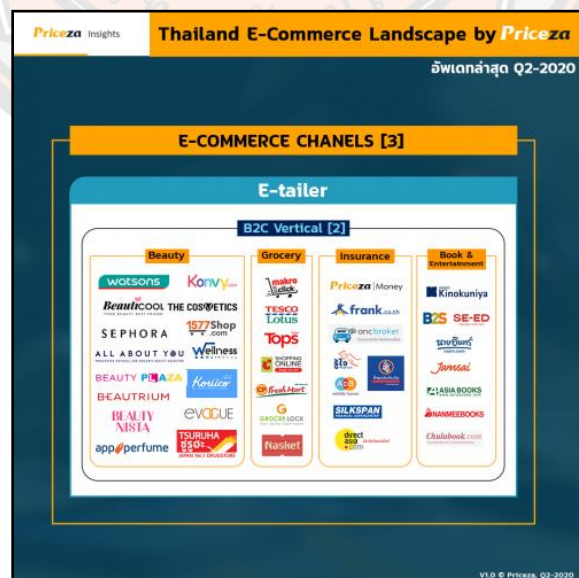
1. Webstore Platform/Software แพลตฟอร์มบริการเปิดร้านค้าออนไลน์ให้แก่พ่อค้าแม่ค้าอาทิ LnwShop / BentoWeb / Tarad
2. B2C Multi-Category เว็บไซต์รวบรวมหลากหลายหมวดให้เลือก เช่น Central / Shopat24 / Robinson / TV direct
3. B2C Vertical เว็บไซต์ที่โดดเด่นเฉพาะตามหมวดต่างๆ ไม่ว่าจะเป็น
 - กลุ่ม Electronics เครื่องใช้ไฟฟ้า อย่าง Power buy / Jib ที่มีบริการส่งฟรีภายใน 3 ชั่วโมง / Advice
 - กลุ่ม Home & Living รวมเฟอร์นิเจอร์และของตกแต่งบ้าน Homepro / Concept furniture SB Design Square
 - กลุ่ม Fashion & Apparel เสื้อผ้าแฟชั่น Pomelo/Zara/Uniqlo/ H&M

- กลุ่ม Beauty ความสวยความงาม Watson / Konvy / Beauticool
- กลุ่ม Book & Entertainment หนังสือ Kinokuniya / B2S / Se-ed
- กลุ่ม Grocery สินค้าอุปโภคบริโภค Makroclick / Tops / Tesco lotus / Big C / CP Fresh Mart และ Grocer lock
- กลุ่ม Insurance ประกันออนไลน์ Priceza Money / Tip insure / Frank



ภาพ 5 E-commerce Channels (Social Commerce)

ที่มา : <https://contentshifu.com/news/thailand-ecommerce-landscape-2020>



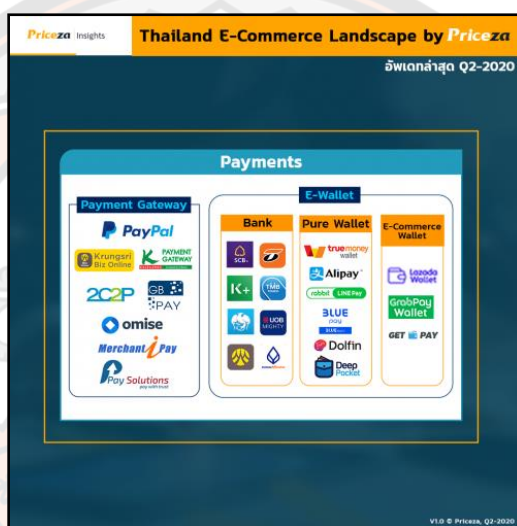
ภาพ 6 E-commerce Channels (E-tailer)

ที่มา : <https://contentshifu.com/news/thailand-ecommerce-landscape-2020>

- ระบบการชำระเงิน (Payment)

ต่อมาเข้าสู่ขั้นตอนระบบการชำระเงิน ปัจจุบันเนื่องจากธุรกิจ E-commerce ได้รับความนิยม จึงทำให้ระบบ Payment มีการแข่งขันกันมากขึ้น ทั้งจากผู้เล่นธนาคารและกลุ่มบริษัทเอกชน ในที่นี้ขอแบ่งออกเป็น 2 ประเภทใหญ่ๆ คือ Payment Gateway และ E-Wallet

กลุ่ม Payment Gateway เป็นระบบรับชำระเงินและจัดการเงิน โดยสามารถควบคุมได้เองทุกอย่าง ดูแลแบบครบวงจรตั้งแต่จบการขายที่หน้าเว็บไซต์ของร้านค้า ไม่มี Popup กวนใจ หรือการส่งผู้ซื้อไปยังหน้าต่างใหม่ / 2C2P บริการชำระเงินออนไลน์เช่นเดียวกัน เป็นที่นิยมมากเห็นได้ว่าบริษัทใหญ่ๆ ที่ต้องมีระบบชำระเงินออนไลน์ส่วนใหญ่ใช้บริการจากเจ้าหนี้ ดังภาพที่ 7



ภาพ 7 E-commerce Channels (Payment)

ที่มา : <https://contentshifu.com/news/thailand-ecommerce-landscape-2020>

กลุ่ม E-Wallet ได้แก่

Bank ระบบการชำระเงินจากแอปพลิเคชัน (Mobile Banking) ตรงจากธนาคาร เริ่มด้วยแม่แม่ณีจาก SCBeasy จากการสร้างความไว้วางใจให้แก่พ่อค้าแม่ค้า คู่ของมงคล ที่ต้องมีแม่แม่ณีตั้งไว้หน้าร้านแทบทุกร้านจนกลายเป็นคาแรกเตอร์ของ SCBeasy และแอปพลิเคชันธนาคารอื่นๆ เช่น KPlus / Krungthai NEXT / KMA (กรุงศรี) / Thanachart Connect / TMB Touch / UOB Mighty Thailand / Bualuang mBanking

Pure Wallet กระเป๋าเงินออนไลน์สะดวกสบาย ไม่ต้องพกเงินสด จ่ายบิลค่าน้ำค่าไฟฟ้าบัตรเครดิตตัวอย่าง true money wallet / rabbit line pays ขึ้นสต็อกเกอร์ชิมไลน์ต่างๆและไว้เติมเงินขึ้นขนส่งสาธารณะอย่าง BTS เป็นต้น

E-commerce wallet กระเป๋าเงินสำหรับแอปพลิเคชันของ E-commerce อย่าง LAZADA ที่มี LAZADA Wallet สามารถเติมเงินและชำระเงินสำหรับไว้ใช้ซื้อของในแอปได้

- **Logistics & Fulfilment**

การแข่งขันในธุรกิจระบบขนส่ง โพรซซาได้แบ่งกลุ่มไว้ 6 กลุ่มดังนี้

3PL หรือ Third Party Logistics การให้บริษัทที่เชี่ยวชาญด้านการขนส่งมาทำหน้าที่ บางอย่างหรือเกือบทั้งหมด เช่น ระบบขนส่ง เพื่อเพิ่มประสิทธิภาพให้ดียิ่งขึ้น อาทิ ไปรษณีย์ไทย / Kerry / Flash express เป็นต้น

3PL Export ที่เชี่ยวชาญด้านการขนส่งต่างประเทศอย่าง DHL / smeshipping

Fulfillment บริการช่วยอำนวยความสะดวกสำหรับพ่อค้าแม่ค้าออนไลน์ มีทั้งระบบจัดการสต็อกสินค้า การแพ็คสินค้า และจัดส่งให้ถึงมือลูกค้า เช่น siam outlet / meowlogis / shipyours เป็นต้น

Shipping Aggregator ตัวกลางบริษัทขนส่งสินค้า ที่ได้รวบรวมบริการขนส่งพัสดุเจ้าต่างๆ เข้ามาอยู่ในระบบ เช่น shippop / easyparcels / Smartship

On Demand แพลตฟอร์มบริการรับส่งสินค้าตามความต้องการ อาทิ Grab / Lineman / Lalamove / เป็นต้น

Parcel Locker ระบบตู้ล็อคเกอร์ทางเลือกใหม่ในการจัดส่งสินค้า เหมือนตู้ไปรษณีย์ที่ผสมผสานเทคโนโลยี สร้างความสะดวกสบายให้ผู้ใช้งานมากขึ้น ผู้ให้บริการได้แก่ Kerry / Lockbox / box24corp

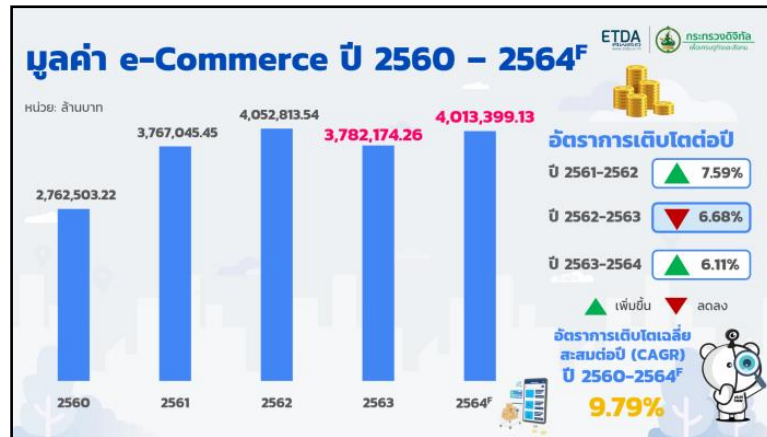
E-commerce Trends 2021 โอกาสใหม่ของธุรกิจไทยที่รออยู่ข้างหน้า นักการตลาดคาดการณ์เห็นความเปลี่ยนแปลงเหล่านี้เกิดขึ้นกับธุรกิจไทยในปี 2021 (Chulamane Wiphawee, 2020)

- ธุรกิจส่วนใหญ่หันมาโฟกัสที่การขายสินค้าออนไลน์ผ่านช่องทาง E-commerce มากขึ้น
- เว็บไซต์, E-marketplace (เช่น Shopee, lazada) และ โซเชียลมีเดีย (เช่น Facebook, Instagram) จะกลายเป็น shopping malls ของโลกยุคใหม่

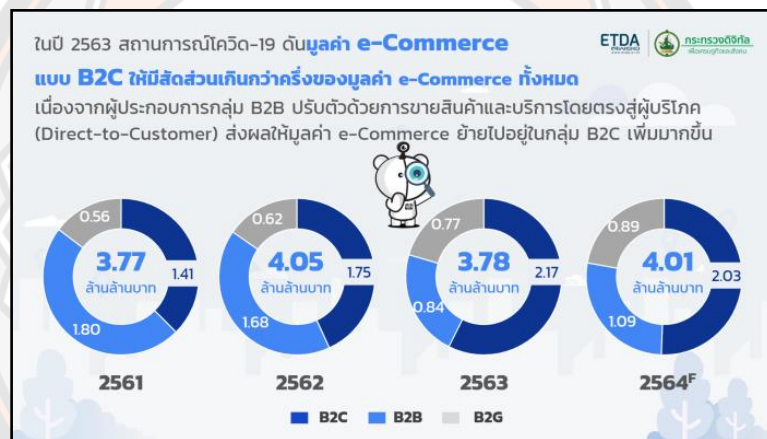
- การซื้อ-ขายสินค้าออนไลน์ง่ายและสะดวกรวดเร็วขึ้นจากการเติบโตของ digital payments

- การผสมผสานแพลตฟอร์มต่าง ๆ เข้าด้วยกัน สร้างประสบการณ์ช้อปปิ้งที่ยอดเยี่ยมให้กับลูกค้า

หากแบ่งสัดส่วน e-Commerce ในประเทศไทย ในปี 2564 พบว่ามีอัตราเติบโตเพิ่มขึ้น 6.11% ดังภาพที่ 8 และส่งผลให้มูลค่า e-Commerce ย้ายไปอยู่กลุ่ม B2C เพิ่มมากขึ้น 2.03 ดังภาพที่ 9 โดยผู้ประกอบการ SMEs เติบโตอย่างต่อเนื่อง 13.94% (นกยง, 2564) ดังภาพที่ 10



ภาพ 8 มูลค่า E-Commerce ปี 2560-2564
ที่มา : <https://www.marketingoops.com/>



ภาพ 9 มูลค่า E-Commerce แบบ B2C
ที่มา : <https://www.marketingoops.com/>



ภาพ 10 มูลค่า E-Commerce จากผู้ประกอบการ SMEs
ที่มา : <https://www.marketingoops.com/>

ความหมายของ Online Marketplace

Online Marketplace คือ เว็บไซต์สื่อกลางการติดต่อซื้อ – ขายที่รวบรวมสินค้าและร้านค้าต่างๆ จำนวนมาก ลักษณะคล้ายกับตลาดนัดหรือห้างสรรพสินค้า แต่จะเป็นในรูปแบบตลาดนัดออนไลน์ ซึ่งรูปแบบบริการของ Online Marketplace เป็นการบริการในรูปแบบของเว็บไซต์ที่เปิดให้ผู้ที่สนใจเปิดร้านค้าของตัวเอง แล้วสามารถนำสินค้าไปวางขายได้ โดยทาง Online Marketplace มีบริการรวมถึงการจ่ายเงินซื้อขายสินค้า,การจัดส่งสินค้าจากร้านค้าไปยังลูกค้าถึงบ้าน นอกจากนี้ยังมีการทำการตลาด การโปรโมท การวางแผนธุรกิจ และศูนย์ส่งเสริมธุรกิจอีกด้วย ดังภาพที่ 11

ตลาดซื้อออนไลน์ในปีที่ผ่านมาไม่รวมบริการจองที่พักโรงแรม ตัวเครื่องบิน ดิจิทัลคอนเทนต์และบริการอื่น ๆ ที่จับต้องไม่ได้ มูลค่าทั้งสิ้น 163,300 ล้านบาท สัดส่วน 3% ของค่าปลีกทั่วประเทศ และคาดการณ์ปีนี้เติบโต 35% ด้วยมูลค่า 220,000 ล้านบาท สัดส่วน 4-5% ของค่าปลีกทั่วประเทศ มูลค่านี้เป็นการอ้างอิงจากโพรซซา โดยในปีที่ผ่านมามูลค่าตลาดออนไลน์ที่เกิดจาก Online Marketplace มีมูลค่าทั้งสิ้น 76,600 ล้านบาท และเชื่อว่าจากการรุกตลาดของ Online Marketplace ที่มาพร้อมกับ New Normal ใหม่ ๆ ที่ดึงให้คนซื้อสินค้าบนโลกออนไลน์มากขึ้น



ภาพ 11 มูลค่าตลาด E-Commerce

ที่มา : <https://www.thumbsup.in.th/priceza-ecommerce-2019>

ธุรกิจ E-Commerce มีมานานแล้วในประเทศไทย ต้องยอมรับว่าในปี 2020 ที่ผ่านมาจากปัญหาโรคระบาดโควิด19 ทำให้เกิดการเติบโตในธุรกิจประเภทนี้มากขึ้น ซึ่งเรียกได้ว่าเป็นปีทองของ E-Commerce ซึ่งจากความสำคัญของธุรกิจนี้เอง “E-Commerce Trends 2021” เป็นกระแสที่น่าสนใจของธุรกิจ E-Commerce ออกเป็น 5 เทรนด์ ดังนี้ (pigabyte, 2021)

Trend 1 E-Commerce Boom สร้างการแข่งขันที่สูงขึ้น

จากสถานการณ์ COVID-19 ทำให้มูลค่าตลาด E-Commerce ในปี 2020 มีมูลค่า 294,000 ล้านบาท โดยในปี 2019 อยู่ที่ 163,000 ล้านบาท เท่ากับว่ามีการเติบโตสูงถึง 81% นอกจากนี้ ยังพบว่าสินค้าที่เข้ามายังตลาด E-Marketplace (เฉพาะ Lazada Shopee JD Central) เพิ่มสูงขึ้นถึง 32% โดยในปี 2019 มี 174 ล้านชิ้น แต่ปี 2020 มีสินค้า 230 ล้านชิ้น เพิ่มขึ้นถึง 32% ในขณะที่มีร้านค้า (ผู้ขาย) เพิ่มสูงมากขึ้นถึง 50% โดยแบ่งเป็น ร้านค้าในประเทศมี 37% และร้านค้าจากต่างประเทศ 63% จึงทำให้เกิดการแข่งขันที่สูงมากขึ้น โดยเฉพาะร้านค้าจากประเทศจีน ที่มีราคาที่ค่อนข้างต่ำ ดังนั้นผู้ประกอบการไทยไม่อาจสู้ในเรื่อง “ราคา” ได้ คำแนะนำคือ ควรที่จะใส่ “ความคิดสร้างสรรค์” ลงไปในสินค้าด้วย จะเข้ามาช่วยตอบโจทย์ในการแข่งขันที่ E-Marketplace นี้ได้ ดังภาพที่ 12



ภาพ 12 มูลค่าตลาด E-Commerce ในสถานการณ์โควิด-19

ที่มา : <https://www.marketingoops.com/news/ecommerce/>

Trend 2 Direct to Consumer แบรินด์มุ่งสู่ออนไลน์มากขึ้น

ในอดีตแม้มีการจำหน่ายผ่านทางช่องทางออนไลน์อยู่บ้างแล้วในแบรินด์และธุรกิจต่างๆ แต่ไม่ได้มากอะไร แต่เมื่อเกิดวิกฤต COVID-19 เม็ดเงินจำนวนมากถูกซึบมายังช่องทางออนไลน์มากขึ้น อย่างทันทีในหลายช่องทาง ไม่ว่าจะเป็นการเปิดเว็บไซต์ออนไลน์ การใช้โซเชียลมีเดีย หรือ E-Marketplace ทั้งนี้พบว่า มีร้านค้าและแบรินด์มุ่งสู่ตลาดอีมาร์เก็ตเพิ่มขึ้น (นับบน Laz Mall และ Shopee Mall) ถึง 46% โดยในปี 2019 มี 3,400 ร้านค้า แต่ในปี 2020 มีสูงถึง 5,000 ร้านค้า ดังภาพที่ 13



ภาพ 13 จำนวนร้านค้า Brand Official Shop
ที่มา : <https://www.marketingoops.com/news/ecommerce/>

เหตุผลที่แบรินด์เริ่มทำการตลาดแบบ Direct to Consumer (D2C) บนออนไลน์มากขึ้นเพราะว่า

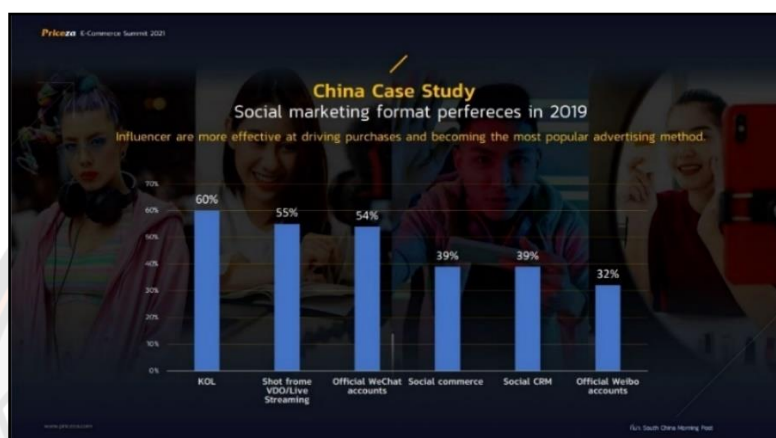
- แบรินด์ต้องการสร้างความสัมพันธ์โดยตรงกับลูกค้า
- ตลาด E-Commerce ได้รับการพิสูจน์ว่ามีแนวโน้มที่เติบโตสูงขึ้น โดยเฉพาะในช่วง COVID-19 และในปีนี้การกลับมาในระลอกใหม่คาดว่าจะการการค้าออนไลน์ของไทยเติบโตขึ้นอีกครั้ง
- ลูกค้าต้องการที่ซื้อตรงจากแบรินด์ในช่องทางออนไลน์ เพราะมั่นใจในเรื่องความเป็นของแท้

Trend 3 เปลี่ยนจากสินค้า “รอการค้นหา” เป็นสินค้า “เข้าหาลูกค้า”

จากอดีตที่ลูกค้าเข้ามาค้นหาสินค้า ได้ถูกเปลี่ยนเป็น “สินค้า” เข้าหาลูกค้าเองมากขึ้น โดยพบว่า ลูกค้าในเอเชียตะวันออกเฉียงใต้ ตัดสินใจซื้อสินค้าเพราะมีสินค้าและโฆษณาที่ต้องการเข้ามาถึง 53% ในขณะที่การเสิร์ชเพื่อค้นหาสินค้าและตัดสินใจซื้อมีเพียงแค่ 35% เท่านั้น นอกจากนี้ ยังพบว่า 40% นักชอปปออนไลน์คนไทย ซื้อสินค้าผ่าน Social Media ซึ่งมากกว่าในหลายๆ ประเทศ มากกว่าสหรัฐฯ อินโดนีเซีย และอินเดีย เป็นต้น

Trend 4 จาก Influencer Marketing สู่ Influencer Commerce

เป็นที่ทราบดีว่ากลยุทธ์ Influencer ได้รับความนิยมเพิ่มมากขึ้น แต่เพียงแค่การทำ การตลาดกับ Influencer หรือที่เรียกว่า Influencer Marketing ไม่อาจเพียงพออีกต่อไป แต่ Influencer ยังต้องช่วยปิดการขายได้ด้วย ซึ่งเรียกว่า “Influencer Commerce” ซึ่งประเทศที่ ประสบความสำเร็จในเรื่องอย่างมาก ได้แก่ ประเทศจีน โดยเฉพาะกลยุทธ์ KOL ที่ได้รับความนิยมสูง มากที่สุดในบรรดากลยุทธ์การตลาดออนไลน์ของจีน ดังภาพที่ 14



ภาพ 14 การตลาดกับ Influencer

ที่มา : <https://www.marketingoops.com/news/ecommerce/>

Trend 5 Convergence of Platforms

การที่แพลตฟอร์มเกิดการขยายตัวออกจากจุดเริ่มต้นเดิมของตัวเอง ไปสู่การเป็นแพลตฟอร์ม อื่นๆ ได้ เช่น การที่ Google ไม่ได้ทำหน้าที่เพียงแค่เสิร์ชเอนจิน แต่ยังสามารถทำทั้งแอดเวอร์ไทซิ่ง, อีคอมเมิร์ซ, เพย์เมนต์และโลจิสติกหรือแม้แต่ธนาคารไทยก้าวสู่การเป็นฟู้ดเดลิเวอรี่ ดังภาพที่ 15



ภาพ 15 Convergence of Platform in Thailand Digital Landscape

ที่มา : <https://www.marketingoops.com/news/ecommerce/>

บทบาทของ Online Marketplace

Online Marketplace มีความสำคัญกับธุรกิจในแง่ของการกระจายให้คนเห็น การหาภาพพิก จากลูกค้าใหม่ และฐานลูกค้าเดิม ให้มาซื้อสินค้าอยู่เรื่อย ๆ สร้างการรับรู้ให้กลุ่มเป้าหมาย ร้านค้า ออนไลน์บางร้านในปัจจุบันมองข้ามช่องทางนี้ไป ซึ่งถ้าร้านค้านี้มีช่องทางการจำหน่ายผ่าน Online Marketplace นอกจากจะได้รับการมองเห็น สร้างการรับรู้แล้วนั้น ยังเป็นการสร้างความน่าเชื่อถือให้กับ ร้านค้าได้อีกด้วย เพราะการขายสินค้าผ่าน Online Marketplace มีความโดดเด่นในด้านความโปร่งใส ลูกค้าเชื่อถือ อุ่นใจ สั่งสินค้าแล้วไม่โดนหลอกโอนเงินฟรี ๆ แน่นนอน เพราะร้านค้าสามารถตรวจสอบได้ มี ที่มาที่ไป มีเว็บไซต์ต้นสังกัดที่นำสินค้าเข้าไปขาย เป็นหลักประกันให้ลูกค้าอุ่นใจ เพราะอย่างที่ทราบกันดี อยู่แล้วว่า ในปัจจุบันมีข่าวคราวมากมายเกี่ยวกับการโกงเงินผ่านการสั่งซื้อสินค้าออนไลน์ บางคนโดนไป หลักร้อย ไปจนถึงหลักพันหลักล้าน ดังนั้นพฤติกรรมกรรมการสั่งซื้อสินค้าออนไลน์ของลูกค้าในปัจจุบัน จึงมี ความระมัดระวัง มีการคิดวิเคราะห์ที่สับสนที่มากขึ้นกว่าเดิม ร้านค้าออนไลน์จึงต้องสร้างความ น่าเชื่อถือและความโปร่งใส ช่องทางการชำระเงินมาก โอกาสขายเพิ่มขึ้น เมื่อเพิ่มช่องทางการขาย ออนไลน์ ควรเพิ่มช่องทางการชำระเงิน ซึ่งตอนนี้มีบริการ Prompt pay ซึ่งทำให้การโอนการจ่าย สะดวกมากยิ่งขึ้น เพราะในการโอนเงินผ่าน Prompt pay ไม่คิดค่าธรรมเนียมรายการเป็นแบบ ในเขต ชำมเขต หรือรายการในธนาคารเดียวหรือต่างธนาคารเหมือนกับการโอนเงินในรูปแบบเดิม แต่ อ่างอิงวงเงินในการโอนเงินแต่ละครั้ง ซึ่งค่าธรรมเนียมถูกกว่า คือ

- โอนเงินไม่เกิน 5,000 บาท ไม่คิดค่าธรรมเนียม
- โอนเงินมากกว่า 5,000 - 30,000 บาท ค่าธรรมเนียมรายการละไม่เกิน 2 บาท
- โอนเงินมากกว่า 30,000 - 100,000 บาท ค่าธรรมเนียมรายการละไม่เกิน 5 บาท
- โอนเงินมากกว่า 100,000 - วงเงินสูงสุดตามแต่ละธนาคารกำหนด ค่าธรรมเนียมรายการ

ละไม่เกิน 10 บาท โดยผู้ที่สนใจใช้บริการรับโอนเงิน Prompt pay ต้องลงทะเบียนกับธนาคารก่อน เพื่อผูกบัญชีเงินฝากกับเลขบัตรประจำตัวประชาชนหรือเบอร์โทรศัพท์มือถือ

ในยุคนี้สามารถลงขายสินค้าพร้อมกันได้หลาย ๆ E-Marketplace ทำให้คนเห็นสินค้า มากขึ้น และยังถ้ามีการรองรับการชำระเงินด้วยระบบ Prompt pay ที่ทำให้ลูกค้าไม่เสีย ค่าธรรมเนียมในการโอนเงิน ยิ่งเพิ่มโอกาสในการขายของบนโลกออนไลน์ ถึงแม้จะมีความคิดว่าถ้าลง ขายหลายที่ต้องจัดการหลายที่ แต่คนมีน้อยจะทำไหวหรือไม่ ไม่ต้องกังวล เพราะส่วนใหญ่ทุกเจ้า ที่เป็น E-Marketplace มีระบบจัดการที่ช่วยทำงานได้ง่าย ส่วน Prompt pay ถึงแม้ตอนนี้ยังไม่ได้มี การใช้ในวงกว้าง แต่ด้วยสิทธิประโยชน์และการสื่อสารที่มากขึ้นช่วยให้คนมาใช้บริการชำระเงินผ่าน ช่องทางนี้เป็นหลัก (บศกดี ลีวัลักษณ์, 2564)

ธุรกิจ e-Commerce ไทยอยู่ในมือของ e-Marketplace ต่างชาติ

หากพิจารณาถึงตลาดอิเล็กทรอนิกส์ หรือ e-Marketplace เจ้าใหญ่ในประเทศไทยทั้ง 3 แพลตฟอร์ม คือ Lazada, Shopee และ JD พบว่า Lazada ที่มี Alibaba Group เป็นบริษัทแม่ และ JD Central ที่เกิดขึ้นมาจากการร่วมทุนของ JD (Jingdong) กับกลุ่ม Central ของไทย ต่างเป็นแพลตฟอร์ม e-Commerce ที่มีผู้ลงทุนเป็นชาวจีน ส่วน Shopee เป็นของ Garena ที่จดทะเบียนในตลาดหลักทรัพย์ที่ประเทศสิงคโปร์ แต่ผู้ถือหุ้นใหญ่ คือ Tencent จากประเทศจีน เห็นได้อย่างชัดเจนว่าประเทศไทยกำลังเผชิญกับการไหลทะลักเข้ามาทำธุรกิจดิจิทัลและ e-Commerce ของนายทุนชาวจีนและประเทศเพื่อนบ้านที่มีความพร้อมด้านเม็ดเงินและทรัพยากรอื่น ๆ ซึ่งตลาด e-Commerce ในประเทศไทยมีแนวโน้มขยายตัวเพิ่มขึ้นอย่างก้าวกระโดดจากการคำนวณจำนวนทั้งหมดของสินค้าที่อยู่บน e-Marketplace 3 แพลตฟอร์มใหญ่ของไทย พบว่าในปี 2018 มีสินค้ารวมกันอยู่ที่ 74 ล้านชิ้น ซึ่งเมื่อเปรียบเทียบกับปี 2019 ที่มีมากถึง 174 ล้านชิ้น เห็นได้ว่ามีจำนวนเพิ่มมากกว่าถึง 2.4 เท่า โดยจำนวนสินค้าที่เพิ่มมากขึ้นมากกว่า 77% เป็นสินค้าที่มาจากต่างประเทศโดยเฉพาะประเทศจีน ส่งผลให้จำนวนสินค้าจากประเทศไทยเองกลับมีสัดส่วนอยู่บนแพลตฟอร์มน้อยกว่ามาก ในปี 2020 สินค้าที่อยู่บน e-Marketplace 3 แพลตฟอร์มใหญ่ของไทยมีสัดส่วนเพิ่มขึ้นมากกว่า 85% เป็นสินค้าที่มาจากประเทศจีน เห็นว่าตลาด e-Marketplace ไม่ได้แข่งในเกมที่ไทยเป็นผู้กำหนดกติกา แต่เป็นเกมของต่างชาติที่มาใช้ประเทศไทยเป็นสมรภูมิในการแข่งขัน โดยกลุ่ม e-Marketplace ที่ 3 ผู้ยิ่งใหญ่ Shopee-Lazada-JD กำลังมีอำนาจเหนือตลาดอย่างเบ็ดเสร็จแม้จะไม่ผูกขาด (นิยามคณะกรรมการแข่งขันทางการค้าว่า ถ้าผูกขาดต้องเป็นผู้เล่นรายเดียว) e-Marketplace platform ไม่ได้หวังเพียงแค่ขายของ แต่หวังผูกขาดธุรกิจต่อเนื่อง e-Marketplace platform กิจกรรมขั้นต้นคือ การให้บริการพื้นที่หรือสื่อกลางให้กับผู้ประกอบการและลูกค้าผ่านเว็บไซต์หรือแอปพลิเคชัน เพื่อเป็นช่องทางในการซื้อขายสินค้าหรือบริการ แต่ธุรกิจซื้อขาย e-Marketplace อยู่โดดเดี่ยวลำพังจะไม่โต ต้องมีธุรกิจ “ตัวกลาง” ต่อเชื่อมนับตั้งแต่การให้บริการชำระเงินออนไลน์แบบครบวงจร การให้บริการจัดส่ง การสร้างเครือข่ายระหว่างผู้ซื้อผู้ขาย และให้ข้อมูลข่าวสาร การอบรมเพื่อพัฒนาธุรกิจออนไลน์ ดังนั้น จะได้เห็นธุรกิจ “ตัวกลาง” เกิดมาช่วยภาคธุรกิจมากขึ้น โดยตัวกลางเหล่านี้จะคอยเชื่อมแพลตฟอร์มต่าง ๆ เข้ากับระบบหลังบ้านของธุรกิจ e-Marketplace และคอยรายงานยอดขาย จัดการส่งสินค้า รับชำระเงิน หรือแม้แต่จัดการด้าน warehouse ให้ได้ ธุรกิจ “ตัวกลาง” เจ้าของ e-Marketplace platform จะเอาต์ซอร์ซแต่เท่าที่เห็นจะเหมารวมทำเองหมด สิ่งที่ตามมาคือจะเห็นพ่อค้าแม่ค้าจากจีนไหลทะลักเข้ามาแข่งกับพ่อค้าแม่ค้าในไทยอย่างรวดเร็วผ่านแพลตฟอร์มทั้งสาม โดยกลุ่มที่มีความเสี่ยงสูงคือ กลุ่มสินค้ากีฬา นาฬิกา ยานยนต์ home entertainment, outdoor และกลุ่มสินค้าสำหรับสัตว์เลี้ยง เนื่องจากเป็นสินค้าที่จีนถนัดผลิต ประมาณการตัวเลขคร่าว ๆ ของสินค้าที่ขายบน e-Marketplace ในไทยตอนนี้ว่ามีประมาณ 50 ล้าน

รายการ ซึ่งในจำนวนนี้ 80% หรือ 40 ล้านรายการเป็นสินค้าที่มาจากต่างประเทศ จีนมีความน่ากลัว จึงอยู่ที่พฤติกรรมของผู้ซื้อชาวไทยที่กำลังเปลี่ยนแปลงไปอยู่บนโลกออนไลน์มากขึ้น เพราะเงินที่จับจ่ายใช้สอยบนโลกออนไลน์มีแนวโน้มจะไหลออกสู่ต่างประเทศโดยตรงผ่านแพลตฟอร์มผู้มีอำนาจเหนือตลาดทั้งสาม Shopee-Lazada-JD ซึ่งหากความเปลี่ยนแปลงเหล่านี้เกิดขึ้นจริง กลุ่มที่น่าเป็นห่วงที่สุดคือ ธุรกิจที่ไม่มีโมเดลการหารายได้ทางอื่นเตรียมไว้รองรับ เช่น ร้านค้าโชห่วยในต่างจังหวัดนั่นเอง (ภาวธ พงษ์วิทย์ภานุ, 2564)

จุดเด่น ข้อดี ที่ช่วยตอบโจทย์ให้ผู้ซื้อ-ผู้ขาย

สมัยนี้ไม่มีใครอยากขับรถออกไปซื้อของนอกบ้านแล้ว เพราะต้องปวดหัวกับปัญหาต่าง ๆ เช่น รถติด ฝนตก ที่จอดรถไม่มี ไหนจะปัญหาฝุ่นอีก ดังนั้นจุดเด่นที่เห็นได้ชัดของ Online Marketplace ออนไลน์คือไม่ต้องออกไปเผชิญปัญหาเหล่านั้น สามารถสั่งซื้อสินค้าที่ต้องการง่าย ๆ เพียงแค่ใช้ปลายนิ้ว อีกทั้งยังสามารถเปรียบเทียบราคา และคุณภาพสินค้าจากหลากหลายแห่งเพื่อเลือกซื้อสิ่งที่ดีที่สุด ที่สำคัญยังสะดวก และประหยัดเวลาอีกด้วย สำหรับตัวผู้ขายเอง สามารถตัดปัญหาเรื่องค่าใช้จ่ายเมื่อมีหน้าร้าน รวมถึงการหาสถานที่ขาย การจัดร้าน หรือค่าใช้จ่ายในเรื่องของค่าเช่าสถานที่และการจ้างพนักงานไปได้เลย อีกทั้ง Online Marketplace ยังมีช่องทางการชำระเงินที่หลากหลายเพื่อเป็นตัวเลือกให้กับผู้ซื้อ เช่น บัตรเครดิต, การเก็บเงินปลายทาง, Internet Banking, Wallet, PayPal เป็นต้น เพื่อดึงดูดลูกค้าใหม่ ๆ ให้เข้ามาซื้อ และร้านค้ายังสามารถจัดโปรโมชั่นหรือให้คูปองส่วนลดได้อีกด้วย และหาก Marketplace เป็นที่รู้จัก ผู้ซื้อจะเข้ามาเอง ดังนั้นหน้าที่การโปรโมท หากลูกค้าจึงเป็นของ Online Marketplace สิ่งที่ผู้ขายต้องทำคืออัปเดตสต็อกและเตรียมของให้พร้อมขายอยู่เสมอเท่านั้นเพื่อไม่ให้เสียโอกาสทางการขาย

ข้อดีของ Online Marketplace

1. เป็นตลาดกลางที่อำนวยความสะดวกแก่ผู้ขาย ไม่ว่าจะร้านจะมีขนาดเล็ก-ใหญ่สามารถขายของที่นี่ได้
2. เปิดร้านง่าย เพราะไม่ต้องลงมือสร้างร้านค้าออนไลน์ ไม่ต้องมีความรู้แบบเชิงลึก เพราะเป็นรูปแบบสำเร็จรูปที่ทาง Online Marketplace จัดการไว้ให้ครบหมดตามที่ร้านค้าควรมี
3. ร้านค้าไม่ต้องการหาลูกค้าเอง เพราะเป็นหน้าที่ของ Online Marketplace ที่ต้องสร้างตัวตน ประชาสัมพันธ์ให้คนรู้จักและสร้างความนิยมให้กับ “ตลาดออนไลน์” ของตนเอง
4. ลูกค้าสามารถเลือกสินค้าประเภทเดียวกันได้ โดยเปรียบเทียบราคา คุณภาพ ของหลายๆร้าน เหมือนการเลือก shopping ออนไลน์ จึงสะดวกและประหยัดเวลาในการเลือกซื้อสินค้า
5. ในแง่ของอำนาจของผู้ขาย และผู้ซื้อ ในอนาคต อาจเป็นผู้กำหนดราคามาตรฐานในการซื้อขายได้ ขึ้นอยู่กับความต้องการซื้อและขาย นั่นเอง

6. ภาระของร้านค้ามีเพียงจัดการกับสินค้าของตัวเอง ปรับปรุงสต็อกสินค้าให้ update อยู่เสมอ พร้อมขาย จัดส่งสินค้า ส่วนการรับโอนเงินขึ้นอยู่กับระบบจัดการของ Online Marketplace ว่าจะรับประกันการได้รับสินค้าแก่ลูกค้าในระดับไหน

7. ในตลาด Online Marketplace เปิดโอกาสให้ลูกค้าในการแจ้งผลตอบรับเรื่องสินค้าและบริการ คือการรีวิวจากลูกค้าตัวจริง เป็นการให้คำรับรองสินค้าและบริการที่เชื่อถือได้

ข้อเสียของ Online Marketplace

1. หาก Online Marketplace วางระบบการรับโอนเงินให้แก่ร้านค้าไม่ดี เปิดโอกาสให้มีการฉ้อโกงกันได้ เช่น จ่ายเงินแล้วไม่ได้รับของ เป็นต้น

2. เนื่องจาก Online Marketplace เป็นตลาดรวม จึงมีคู่แข่งทางการค้าที่ขายสินค้าประเภทเดียวกัน ผู้ขายจึงต้องมีกลยุทธ์ในการทำให้ร้านค้าปรากฏให้ลูกค้าเห็นก่อนร้านค้าอื่นในลักษณะเดียวกัน ซึ่งอาจต้องมีรายจ่ายเพื่อซื้อตำแหน่งร้าน

อนึ่ง แม้ Online Marketplace มีลูกค้าเข้ามา แต่ผู้ขายต้องเลือก Online Marketplace ที่เหมาะกับร้านของตนให้ตรงกับรสนิยมคนเข้า Online Marketplace แห่งนั้นด้วย

ข้อควรระวังของ Online Marketplace

การที่ลงขายง่าย ใคร ๆ ลงขายได้นั้นหมายถึง มีคู่แข่งทางการค้ามากขึ้นไปด้วย การแข่งขันที่สูงขึ้น ทำให้ผู้ขายจำเป็นต้องคิดหากกลยุทธ์ในการขายเอาไว้ดึงดูดลูกค้าเสมอและการซื้อขายสินค้าที่เป็นระบบออนไลน์ ผู้ซื้อจำเป็นต้องจ่ายเงินก่อนที่ทางร้านจัดส่งสินค้า ซึ่งอาจทำให้เกิดการฉ้อโกงได้ ทาง Online Marketplace ต้องมีมาตรการและระบบการจัดการที่ดีเพื่อรักษาผลประโยชน์ของทั้งผู้ซื้อและผู้ขาย และเมื่อ Online Marketplace มีลูกค้าเข้ามาใช้งานมาก ผู้ขายจึงจำเป็นต้องเลือก Online Marketplace ที่เหมาะสมกับสินค้าของตนและตรงความต้องการของผู้ซื้อใน Online Marketplace แห่งนั้น ๆ ด้วย (mycloudfulfillment, 2020)

งานวิจัยเกี่ยวกับตลาดซื้อขายสินค้าออนไลน์ (Online Marketplace)

(เปรมกมล หงส์ยนต์, 2562) ปัจจัยที่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชันออนไลน์ (ลาซาด้า) ของลูกค้ายุคดิจิทัลในกรุงเทพมหานคร ซึ่งการศึกษางานวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อศึกษาถึงปัจจัยที่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชัน (ลาซาด้า) ของลูกค้ายุคดิจิทัลในกรุงเทพมหานคร กลุ่มตัวอย่างที่เลือกใช้ในการศึกษาคั้งนี้ คือ ประชากรในกรุงเทพมหานคร จำนวน 400 คน โดยการสุ่มอย่างง่าย เครื่องมือที่เลือกใช้ในการเก็บรวบรวมข้อมูลคือเลือกใช้แบบสอบถาม สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ การแจกแจงความถี่ ค่าเฉลี่ย ค่าร้อยละ และส่วนเบี่ยงเบนมาตรฐาน สำหรับสถิติที่ใช้ทดสอบสมมติฐานในการวิเคราะห์การถดถอยแบบพหุคูณ (Multiple regressions analysis) ผลการศึกษาพบว่า ลูกค้าส่วนใหญ่เป็นเพศหญิง โดยมีอายุระหว่าง 20 – 30 ปี มีอาชีพส่วนใหญ่ ทำงานบริษัทเอกชน/รับจ้าง ระดับการศึกษา ระดับ

ปริญญาตรี ประชากรส่วนใหญ่ให้ระดับความสำคัญด้านปัจจัยส่วนประสมทางการตลาด 6P และพฤติกรรมการตัดสินใจซื้ออยู่ในระดับมาก ผลการทดสอบสมมติฐานพบว่า ปัจจัยส่วนประสมทางการตลาดออนไลน์ที่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชันออนไลน์ (ลาซาด้า) ด้วยกัน 2 ด้าน คือ ปัจจัยด้านการส่งเสริมการตลาด (.000) และปัจจัยด้านผลิตภัณฑ์ (.004*) โดยปัจจัยด้านราคาและปัจจัยด้านช่องทางการจัดจำหน่าย และปัจจัยด้านการรักษาความเป็นส่วนตัวไม่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชันออนไลน์ (ลาซาด้า) ของลูกค้ายุคดิจิทัลในกรุงเทพมหานคร

กิตติพัฒน์ พันเรือง และดชกรณัฏ์ ต้นเจริญ (2563) การวิเคราะห์ปัจจัยที่มีผลต่อการซื้อซ้ำของตลาดกลางพาณิชย์อิเล็กทรอนิกส์ในกรุงเทพฯและปริมณฑล เป็นการศึกษาเพื่อที่วิเคราะห์ปัจจัยที่มีผลต่อการซื้อซ้ำของตลาดกลางพาณิชย์อิเล็กทรอนิกส์ ในกรุงเทพฯและปริมณฑล วัตถุประสงค์ของงานวิจัยนี้เพื่อวิเคราะห์ปัจจัยเปรียบเทียบตลาดกลางพาณิชย์อิเล็กทรอนิกส์ หรือเว็บไซต์ขายสินค้าออนไลน์ที่เป็นที่นิยม 2 เว็บไซต์คือ Lazada และ Shopee โดยผ่านการสำรวจแบบสอบถามออนไลน์ในเขตกรุงเทพมหานครและปริมณฑล จำนวน 412ชุดแบ่งมาตรการวัดความพึงพอใจเป็น 7 ระดับ ซึ่งมีการวิเคราะห์ปัจจัยที่มีผลต่อการซื้อซ้ำทั้งหมด 8 ปัจจัย ประกอบด้วย ความคาดหวังในความพยายามใช้งานเว็บไซต์, ความคาดหวังในประสิทธิภาพเว็บไซต์, ความสามารถตนเองต่อการเข้าถึงเว็บไซต์, ความน่าเชื่อถือของเว็บไซต์ขายสินค้า, ราคา สินค้าและโปรโมชั่นในเว็บไซต์, คุณภาพของระบบเว็บไซต์ขายสินค้า, คุณภาพของข้อมูลสินค้าในเว็บไซต์ และคุณภาพด้านการบริการของเว็บไซต์ซึ่งผลการวิเคราะห์การกระจายของข้อมูลนั้น พบว่าร้อยละโดยรวมของความพึงพอใจ โดยมีความพึงพอใจที่มากที่สุดในด้าน ความสามารถตนเองต่อการเข้าถึงเว็บไซต์อยู่ที่ร้อยละ 81.67 ตามด้วยด้านความคาดหวังในความพยายามใช้งานเว็บไซต์ที่ ร้อยละ 80.99 และคุณภาพของข้อมูลสินค้าในเว็บไซต์ที่ร้อยละ 80.18 ตามลำดับ เมื่อแบ่งตามความพึงพอใจของแต่ละช่วง อายุพบว่าในช่วงอายุ 21 ถึง 30 ปีและ 31 ถึง 40 ปีซึ่งเป็นช่วงอายุที่มีจำนวนกลุ่มตัวอย่างที่มากที่สุด ให้ความพึงพอใจในด้านเดียวกัน คือ ด้านความสามารถตนเองต่อการเข้าถึงและผลของการเปรียบเทียบปัจจัยที่มีผลต่อความพึงพอใจระหว่าง Lazada และ Shopee เมื่อพิจารณาด้วยวิธีการวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอนุมาน ระดับนัยสำคัญทางสถิติที่ระดับ 0.05 พบว่ามี 4 ปัจจัยที่มีความแตกต่างกัน ได้แก่ ความคาดหวังในความพยายามใช้งานเว็บไซต์, ความสามารถตนเองต่อการ เข้าถึงเว็บไซต์, ราคาสินค้าและโปรโมชั่นในเว็บไซต์และ คุณภาพของระบบเว็บไซต์ขายสินค้า

จากการพิจารณาบทความและงานวิจัยนั้น นักวิจัยได้ใช้ข้อมูลจากแบบสอบถามที่กล่าวถึงปัจจัยส่วนประสมทางการตลาด 6P และพฤติกรรมการตัดสินใจซื้อคือ ปัจจัยด้านการส่งเสริมการตลาดและปัจจัยด้านผลิตภัณฑ์ โดยปัจจัยด้านราคาและปัจจัยด้านช่องทางการจัดจำหน่าย และปัจจัยด้านการรักษาความเป็นส่วนตัวไม่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชันออนไลน์ และมีการ

เปรียบเทียบปัจจัยที่มีผลต่อความพึงพอใจระหว่าง Lazada และ Shopee พบว่ามี 4 ปัจจัยที่มีความแตกต่างกัน ได้แก่ ความคาดหวังในความพยายามใช้งานเว็บไซต์, ความสามารถตนเองต่อการเข้าถึงเว็บไซต์, ราคาสินค้าและโปรโมชั่นในเว็บไซต์และ คุณภาพของระบบเว็บไซต์ขายสินค้า แต่ยังมีงานศึกษาน้อยและศึกษาแค่เพียงแพลตฟอร์มเดียวคือ Lazada ทำให้ผู้วิจัยมีความสนใจในส่วนของ การเข้าถึงการรีวิวสินค้า ซึ่งผู้รีวิวอาจไม่ใช่ลูกค้า แต่ถูกจ้างหรือได้รับผลประโยชน์จากการรีวิว สามารถเข้ามา รีวิวในตลาดซื้อขายสินค้าออนไลน์ได้อย่างอิสระและข้อความรีวิวนั้นส่งผลกระทบต่อความน่าเชื่อถือ และการตัดสินใจซื้อผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ได้

2.3 ช่องทางการให้บริการแบบดิจิทัล (Digital Touchpoint)

ช่องทางการให้บริการแบบดิจิทัล (Digital Touchpoint) สาเหตุหรือแรงขับเคลื่อนที่ทำให้เกิดแนวโน้มนี้มาจากความเป็นปัจเจกชนของผู้บริโภคที่ต้องการสินค้าเฉพาะตนและมีความเป็นส่วนตัวมากขึ้น (Individualization) ความก้าวหน้าของเทคโนโลยีการสื่อสารที่ทำให้การเข้าถึงข้อมูลลูกค้าเชิงลึกจาก On-line Platform ในรูปแบบต่าง ๆ เป็นเรื่องง่าย ทำให้ธุรกิจมีข้อมูลมากพอในการวิเคราะห์เพื่อออกแบบสินค้าและบริการที่ตรงความคาดหวังของผู้บริโภคอย่างรวดเร็ว (Customer Insight) ธุรกิจจะต้องค้นหาความต้องการในการตลาดใหม่ การสร้างประสบการณ์ใหม่ ๆ (New Experiences) รวมถึงช่องทางการเข้าถึงสินค้าและบริการโดยปราศจากข้อจำกัดในเรื่องของเวลาและสถานที่ (Accessibility) ทำให้ผู้ประกอบการสามารถติดต่อลูกค้าและขายสินค้าและบริการได้ตลอดเวลา โดยผู้นำทางด้านเทคโนโลยีโลกมีการคาดการณ์ว่า ภายในปี 2021 หรืออีก 2 ปีข้างหน้า ช่องทางการให้บริการแบบดิจิทัลนี้จะสร้างรูปแบบการดำเนินธุรกิจใหม่ ๆ เป็นจำนวนมาก และเป็นเครื่องมือที่ใช้วิเคราะห์ช่องทางที่กลุ่มเป้าหมายใช้งานในแต่ละวัน แต่ช่วงเวลา โดยไม่ได้เจาะจงว่าเป็นช่องทางที่ลูกค้าใช้เพื่อจะตัดสินใจซื้อ โดย Digital Touchpoint จะช่วยให้รู้ว่าช่องทางไหนที่กลุ่มเป้าหมายใช้งานบ่อยที่สุด ในช่วงเวลาใด นำไปใช้ต่อยอดในการเลือกช่องทาง และเวลา ในการนำเสนอคอนเทนต์ให้กลุ่มเป้าหมายพบเจอได้ง่ายมากขึ้น เพื่อเพิ่มโอกาสที่ลูกค้าจะซื้อสินค้ามากขึ้น ด้วย ตัวอย่างเช่น ลูกค้าอยากซื้อสินค้าสักชิ้น อาจจะค้นหาสินค้าผ่านช่องทาง online และเห็นร้านค้าผ่าน Website หรือ โฆษณาต่างๆ หรือ โทรเข้า Call Center กล่าวได้ว่าจุดสัมผัสของลูกค้ากับแบรนด์ของคุณสามารถเกิดขึ้นได้ทุกช่วง ตั้งแต่ช่วงก่อนจะเจอ ช่วงเจอ ช่วงระหว่างตัดสินใจ ช่วงตัดสินใจซื้อ และช่วงหลังการซื้อสินค้าและบริการ (Tirasak W.Warm, 2021)

ในการทำสินค้าหรือบริการ “ลูกค้า” คือคนสำคัญที่อยากให้เกิดความพึงพอใจ ไม่ว่าจะเป็นความพึงพอใจในกระบวนการก่อนที่จะซื้อสินค้าหรือบริการ ไม่ว่าจะเป็นความพึงพอใจในระหว่างซื้อสินค้าหรือบริการ หรือความพึงพอใจในบริการหลังการขายสินค้าหรือบริการ ดังนั้น Touchpoint หรือ จุดสัมผัสของลูกค้ากับสินค้าหรือบริการ เป็นจุดสำคัญที่จะเสริมสร้างความรู้สึกดี ความพึงพอใจ

ให้กับลูกค้าได้เป็นอย่างดี ต่อให้สินค้าหรือบริการจะดีแค่ไหน แต่ถ้าแต่ละจุดของ Touchpoint ไม่ดี ก็อาจทำให้ประสบการณ์โดยรวมของลูกค้าไม่ดี ส่งผลให้เกิดความรู้สึกไม่ดี หรือไม่พึงพอใจกับสินค้าหรือบริการ เพราะฉะนั้นการให้ความสำคัญกับจุดสัมผัสของลูกค้า หรือ Touchpoint จะทำให้มองเห็นแนวทางในการสร้างความพึงพอใจให้กับลูกค้า ให้ครบทุกจุดเพื่อให้เกิดประสบการณ์ที่ดีกับสินค้าหรือบริการ ตั้งแต่จุดเริ่มต้นจนถึงจุดสิ้นสุด ตัวอย่างเช่น ลูกค้าค้นหาธุรกิจผ่านออนไลน์ โฆษณา การเยี่ยมชมเว็บไซต์ การเห็นจากร้านค้า (Shop) หรือศูนย์บริการลูกค้า (Customer Service) ซึ่งสามารถกล่าวได้ว่าจุดสัมผัสของลูกค้ากับแบรนด์ สามารถเกิดขึ้นได้ตั้งแต่ช่วงก่อน ระหว่าง และหลังการซื้อขายสินค้าและบริการ โดยการที่จะสามารถระบุจุดสัมผัสต่างๆ ได้นั้น จำเป็นต้องสร้างแผนที่การเดินทางของลูกค้า (Customer Journey Map) เพื่อสร้างความมั่นใจว่าลูกค้าจะได้รับความพึงพอใจในทุกๆ จุดตลอดการเดินทาง ได้แก่ จุดสัมผัส 3 ช่วง ดังนี้ (ปรีดี นุกุลสมปรารถนา, 2563)

- จุดสัมผัสช่วงก่อนการซื้อ (Before Purchase)
 - เว็บไซต์
 - โซเชียล มีเดีย
 - การอ่านรีวิว หรือ Testimonial ของลูกค้า
 - ไตเร็คทอรีสินค้าต่างๆ
 - โฆษณาทั้งออนไลน์และออฟไลน์
 - โปรโมชั่นส่งเสริมการขาย
 - ประชาสัมพันธ์
 - การตลาดรูปแบบต่างๆ
 - ร้านค้า ห้างสรรพสินค้า
 - โฆษณาในร้านค้าหรือห้างสรรพสินค้า
 - การพูดแบบปากต่อปาก
 - อีเว้นท์ต่างๆที่จัดขึ้นในร้านค้าหรือห้างสรรพสินค้า
 - การทำคอนเท
 - นต์ผ่านสื่อต่างๆ
 - การมีส่วนร่วมในกิจกรรมตามชุมชนต่างๆ
- จุดสัมผัสช่วงระหว่างการซื้อ (During Purchase)
 - เว็บไซต์ (พวก E-Commerce)
 - รูปแบบโชว์รูม หรือร้านค้า
 - วิดีโอแสดงการสาธิตการใช้สินค้า
 - การมีส่วนร่วมของพนักงานขาย

- การส่งข้อความต่างๆ
- การรีวิวสินค้า
- แคมเปญส่งเสริมการขาย
- โฆษณา ณ จุดขาย
- จุดสัมผัสหลังการซื้อ (After Purchase)
 - การ์ดหรืออีเมลขอบคุณลูกค้า
 - การทำแบบสำรวจความเห็นและความพึงพอใจ
 - ความถูกต้องของใบเสร็จรับเงิน
 - รูปแบบบรรจุภัณฑ์สินค้า
 - ศูนย์บริการลูกค้าทั้งแบบร้านบริการ หรือแบบออนไลน์
 - ทีมสนับสนุนหลังการขาย
 - การนำเสนอโปรโมชั่นใหม่ๆ
 - Customer Loyalty Program
 - ระบบ CRM ในการบริหารความสัมพันธ์กับลูกค้า

การเชื่อมโยงแต่ละจุดสัมผัส

เมื่อสามารถระบุได้แล้วว่าลูกค้าติดต่อกับแบรนด์ผ่านจุดสัมผัสใดบ้าง ก็ต้องนำแต่ละจุดสัมผัสมาวางตามช่วงเวลาที่จะเกิดขึ้น ซึ่งต้องพิจารณาถึงกระบวนการและขั้นตอนของลูกค้าในกระบวนการซื้อสินค้าของแต่ละคน โดยใช้หลักการง่ายๆ ดังนี้

ขั้นที่ 1 การทำให้ลูกค้ารู้จักแบรนด์ของคุณ เพื่อให้ลูกค้ารู้สึกคุ้นเคยกับสินค้าและบริการ

ขั้นที่ 2 ทำให้ลูกค้าอยากที่จะมาหาซื้อสินค้าหรือสอบถามรายละเอียดสินค้าของคุณ

ขั้นที่ 3 เริ่มกระบวนการขาย เพื่อให้มั่นใจว่าลูกค้าจะซื้อสินค้าของคุณ

ขั้นที่ 4 ทำให้ลูกค้ากลับมาซื้อซ้ำอีกครั้งและบอกต่อให้คนอื่นๆรับรู้

การปรับปรุงแต่ละจุดสัมผัสให้ดีขึ้นตลอดเวลา การที่สามารถระบุจุดสัมผัสของลูกค้าแต่ละจุดได้แล้วนั้น ถือเป็นเพียงแค่จุดเริ่มต้นที่ทำให้ลูกค้ารู้จักกับแบรนด์ แต่การที่จะสร้างประสบการณ์ที่ดีจำเป็นต้องมีการปรับปรุงจุดสัมผัสแต่ละจุดอยู่ตลอดเวลา เพื่อหาโอกาสในการมอบสิ่งที่ดีที่สุดให้กับลูกค้า โดยการสร้างประสบการณ์ที่ดีกับแบรนด์ ตามแนวทางดังนี้

- ทำสิ่งง่ายๆ การสร้างประสบการณ์ที่ดี ไม่ควรสลับซับซ้อน ควรนำเสนอให้เรียบง่าย เข้าใจง่ายและตอบโจทย์ลูกค้ามากที่สุด
- มีความเหมาะสม การสร้างประสบการณ์ที่ดี ต้องเหมาะสมกับบริบทและช่องทางต่างๆที่ลูกค้าสนใจ

- มีความเกี่ยวข้อง การสร้างประสบการณ์ที่ดี ต้องเกี่ยวข้องกับความต้องการของลูกค้า หรือสิ่งที่ลูกค้าคาดหวัง
- มีความหมาย การสร้างประสบการณ์ที่ดี ต้องเสนอสิ่งที่มีคุณค่าและเป็นประโยชน์กับลูกค้า
- น่าดึงดูด การสร้างประสบการณ์ที่ดี ต้องพยายามเข้าไปอยู่ในใจของลูกค้า และสามารถมอบในสิ่งที่ลูกค้าต้องการได้

การทบทวนอยู่ตลอด ซึ่งจุดสัมผัสลูกค้าไม่ใช่สิ่งที่สามารถทำเสร็จได้ในครั้งเดียว แต่มันเป็นหนึ่งในกลยุทธ์ที่จำเป็นต้องมีการวางแผนและหาวิธีเติมเต็มแต่ละจุดสัมผัส ที่ต้องทบทวนและตรวจสอบอยู่เป็นระยะ เพราะในทุกๆวันก็อาจมีกลุ่มลูกค้าใหม่ๆและจุดสัมผัสใหม่ๆเกิดขึ้นได้อยู่ตลอดเวลา แต่สิ่งสำคัญที่สุดนั่นก็คือการปรับปรุงประสบการณ์ของลูกค้าที่มีต่อแบรนด์ให้ดียิ่งๆ ขึ้นไป

Touch Point ที่ตอบโจทย์

เมื่อเข้าใจ Customer Journey ของลูกค้าแล้ว และรู้ว่ามี TouchPoint อะไรบ้าง ยังไม่พอ เพราะการที่มี TouchPoint หรือ การมี TouchPoint เยอะๆ อาจจะได้ไม่ได้ช่วยสร้างประสบการณ์ที่ดีให้กับลูกค้าได้ ต้องทำการออกแบบให้แต่ละ TouchPoint สัมพันธ์กัน และสามารถตอบสนองวัตถุประสงค์หลักของแต่ละจุดของ Customer Journey ที่ TouchPoint นั้นอยู่ ได้แก่ สร้าง Touchpoint ให้ตอบรับกับ Journey เมื่อสามารถระบุได้แล้วว่าลูกค้าเจอสินค้าและบริการของคุณผ่าน Touchpoint อะไรบ้างในแต่ละ Customer Journey อาจจะมีมองตามช่วงเวลาที่จะเกิดขึ้น ดูจากกระบวนการและขั้นตอนของลูกค้าในกระบวนการซื้อสินค้าของแต่ละคน (Tirasak W.Warm, 2021)

- Awareness Touchpoint จะเป็นตัวที่ทำให้ลูกค้ารู้จักสินค้าและบริการ เพื่อให้ลูกค้ารู้สึกคุ้นเคยกับสินค้าและบริการให้เค้ารู้สึกอยากได้ ให้ลูกค้ารู้ว่าคืออะไร ในจุดนี้จะเป็นจุดเริ่มต้นเลยที่ลูกค้าอยากจะรู้จักเรา ถ้าเข้าใจว่าลูกค้าปกติใช้งานสื่อไหนเป็นหลักก็จะสามารถสร้าง Touchpoint ได้ตรงจุด
 - Consideration Touchpoint ทำให้ลูกค้าที่อยากที่จะมาหาซื้อสินค้าหรือบริการของคุณได้เข้าใจถึงรายละเอียดต่างๆ มองเป็นจุดดี จุดขาย สอบถามรายละเอียดสินค้า ไม่ว่าจะผ่านช่องทางไหนก็ตาม และควรเข้าถึงได้ง่ายและครบถ้วน
 - Decision/Purchase Touchpoint เริ่มกระบวนการขาย เพื่อให้มั่นใจว่าลูกค้าจะซื้อสินค้าอย่างมีประสบการณ์ที่ดี
 - Retention and Avocacy Touchpoint ทำให้ลูกค้ากลับมาซื้อซ้ำอีกครั้งและบอกต่อให้คนอื่นรับรู้
- ข้อเสนอแนะ สำหรับการสร้างหรือปรับปรุง Touchpoint**
- ทำสิ่งง่ายๆ การสร้างประสบการณ์ที่ดี ไม่ควรสลับซับซ้อน ควรนำเสนอให้เรียบง่าย เข้าใจง่ายและตอบโจทย์ลูกค้ามากที่สุด
 - มีความเหมาะสม การสร้างประสบการณ์ที่ดี ต้องเหมาะสมกับบริบทและช่องทางที่ลูกค้าสนใจ

- มีความเกี่ยวข้อง การสร้างประสบการณ์ที่ดี ต้องเกี่ยวข้องกับความต้องการของลูกค้า หรือสิ่งที่ลูกค้าคาดหวัง
- ความหมาย การสร้างประสบการณ์ที่ดี ต้องเสนอสิ่งที่มีคุณค่าและเป็นประโยชน์กับลูกค้า
- น่าดึงดูด การสร้างประสบการณ์ที่ดี ต้องพยายามเข้าไปอยู่ในใจของลูกค้า และสามารถมอบในสิ่งที่ลูกค้าต้องการได้

บทบาทของรีวิวลินค้ากับการทำการตลาดที่เปลี่ยนไป

การมีรีวิวลินค้าเป็นเรื่องที่ดีต่อผู้บริโภค (Sarunjade, 2020) เมื่อแบรนด์มีรีวิวลินค้าไปในทางบวก มีคนพูดถึงเยอะตาม Touchpoint ต่างๆ คนก็ยิ่งสนับสนุน ชื้อมาใช้และบอกต่อ แต่รีวิวลินค้าก็ไม่ใช่ว่าจะดีทุกอย่างที่ทำให้แบรนด์ประสบความสำเร็จ เพราะสินค้าขายได้ดีนั้นมันยังขึ้นอยู่กับตัวประสบการณ์การใช้งานของตัวผู้บริโภคเอง เช่นถ้ามีคนบอกว่าใช้สินค้าของแบรนด์นี้ดี แต่พอใช้แล้วกลับมีปัญหา แบบนี้การมีรีวิวลินค้าไปในทางบวกก็ไม่ช่วยให้คนนั้นเชื่อได้ว่าสินค้านั้นเป็นสินค้าที่ดี และยอดขายจะดีไม่ได้ถ้าตัวแบรนด์ไม่สื่อสารกับลูกค้า คือเสียงของตัวแบรนด์ยังมีความสำคัญอยู่ไม่น้อย หากบทบาทของการพูดปากต่อปากมากขึ้น บทบาทของประสบการณ์การใช้งานสินค้าของผู้บริโภค และเสียงของแบรนด์ก็จะมีบทบาทน้อยลง และเช่นกันหากตัวใดตัวหนึ่งมีบทบาทมากขึ้น ที่เหลือก็จะมีบทบาทน้อยลง แล้วจะรู้ได้อย่างไรว่าตอนไหนที่รีวิวลินค้าออนไลน์ของเรามีบทบาทต่อตัวแบรนด์ ยิ่งรีวิวลินค้ามีบทบาทต่อแบรนด์ของเรามากเท่าไร ประสบการณ์การใช้งานสินค้าและสิ่งที่แบรนด์กำลังพูดอยู่จะมีบทบาทน้อยลง นั่นทำให้แบรนด์คู่แข่งหน้าใหม่ๆ เข้ามาชิงส่วนแบ่งตลาดได้ง่าย เพราะไม่ต้องลงทุนอะไรมาก แค่ทำสินค้าให้ดี มีรีวิวลินค้า เยอะ ลูกค้าก็พร้อมที่จะเปลี่ยนมาใช้แบรนด์ของคุณคู่แข่งตลอดเวลา จะสังเกตปรากฏการณ์ในแพลตฟอร์มขายสินค้าออนไลน์อย่าง Lazada Shopee ที่ Brand Loyalty นั้นว่ามีน้อยลง หากแบรนด์เป็นแบรนด์ที่ไม่ได้ฟังฟังรีวิวลินค้ามากนัก แต่ฟังการพัฒนาคุณภาพของสินค้าให้ใช้งานได้ดีเยี่ยม และตัวแบรนด์เองก็สื่อสารและสร้างความสัมพันธ์กับลูกค้าอยู่ตลอดเวลา ความจงรักภักดีที่มีต่อแบรนด์ก็จะมีมากขึ้น คู่แข่งหน้าใหม่ก็มีโอกาสแย่งส่วนแบ่งตลาดได้ยากขึ้น เป็นไปได้หรือไม่ที่การกระตุ้นให้คนมารีวิวลินค้าจะทำให้การโฆษณาจากเหล่าเซเลบคนดังมีผลน้อยลงแบนเนอร์ที่เคยมีผลต่อการสร้างการรับรู้ของแบรนด์อาจได้ผลน้อยลงเช่นกัน เพราะถึงเวลาที่ลูกค้าซื้อของ สุดท้ายลูกค้าก็เชื่อรีวิวลิน ความจริงคือไม่เสมอไป ขึ้นอยู่กับสินค้าที่ขาย ถ้าเป็นสินค้าที่มีความสำคัญ ผู้บริโภคต้องใช้เวลาหาข้อมูลเยอะก่อนการตัดสินใจ แน่แน่นอนว่ารีวิวลินค้ามีความสำคัญ แต่ถ้าเราขายสินค้าที่ไม่จำเป็นต้องตัดสินใจนาน รีวิวลินค้าก็อาจมีบทบาทน้อยลง การตลาดแบบเดิมอย่างป้ายโฆษณา โปสเตอร์ แผ่นพับตามทางก็ยังได้ผลอยู่ นอกจากนี้รีวิวลินค้ายังมีผลในแง่ของการแบ่งประเภทของสินค้า เพราะสินค้าเดียวกันอาจจะฟังพารีวิวลินค้ามากเป็นพิเศษ

งานวิจัยการวิเคราะห์ช่องทางการให้บริการแบบดิจิทัล (Digital Touchpoint)

สลิลาทิพย์ ทิพย์ไกรศร (2561) เส้นทางของผู้บริโภค ในปัจจุบันที่เต็มไปด้วยความเปลี่ยนแปลงทางเทคโนโลยีทำให้ผู้บริโภคมีจุดสัมผัสกับ ตราสินค้าเปลี่ยนแปลงไป โดยเส้นทางของผู้บริโภคนั้นจะครอบคลุมตั้งแต่ก่อนการซื้อ ระหว่างการซื้อ และหลังการซื้อ สินค้าหรือบริการ ซึ่งมี 6 ขั้นตอนคือ การรับรู้ การประเมินทางเลือก ความชื่นชอบ การซื้อ ความจงรักภักดี การบอกต่อ ซึ่งเครื่องมือที่สำคัญที่จะทำให้ทราบถึงเส้นทางของผู้บริโภคนั้นคือ แผนที่เส้นทางของผู้บริโภคที่ธุรกิจจะต้องวิเคราะห์ ในเชิงลึกเกี่ยวกับการวิเคราะห์ความคิดของลูกค้า การวิเคราะห์ความต้องการของลูกค้า และการวิเคราะห์พฤติกรรมของลูกค้า แล้วจัดทำเป็นแผนที่เส้นทางของผู้บริโภคเพื่อให้ทราบว่า ผู้บริโภคมีจุดสัมผัสกับตราสินค้าอย่างไรเพื่อนำมาวางแผนกลยุทธ์ทางการตลาดให้ตอบโจทย์กลุ่มลูกค้าเป้าหมายให้เกิดความพึงพอใจและสามารถบริหารประสบการณ์ ของลูกค้าได้อย่างมีประสิทธิภาพมากขึ้น

2.4 รีวิวออนไลน์

ในปัจจุบันการขายของออนไลน์ทำไมต้องมีการรีวิว ซึ่งเป็นส่วนสำคัญที่ช่องทางออนไลน์ต้องมีการอำนวยความสะดวกให้การซื้อขายสินค้าหรือบริการง่ายขึ้น และยังเปิดโอกาสให้ลูกค้าแสดงความคิดเห็นได้ง่ายกว่าเดิมด้วยทุกครั้งที่ถูกซื้อของ ไม่ว่าจะร้านใดก็ตามประสบการณ์ในการ Shopping เป็นสิ่งสำคัญมาก หากมอบบริการที่ดึงดูดลูกค้าพึงพอใจ เสี่ยงตอบรับยอมออกมาดี แต่เมื่อใดก็ตามที่ การบริการบกพร่อง Comment ประสบการณ์ ด้านลบเป็นสิ่งที่เลี่ยงไม่ได้ สิ่งเหล่านี้พบเห็นได้ทั่วไป ในทุกร้านค้า หรือเรียกว่า “รีวิว” (Review) (วารสารกรมพัฒนาธุรกิจการค้า, 2562)

การมี “รีวิว” (Review) แสดงประกอบรายละเอียดสินค้าในเว็บไซต์ หรือร้านค้าบนแพลตฟอร์มออนไลน์ต่าง ๆ นั้นจะช่วยเสริมการพิจารณาของลูกค้าที่กำลังดูร้านค้าอยู่ เพราะสิ่งหนึ่งที่ลูกค้าในยุคนี้มองหาเมื่อสนใจสินค้าหรือบริการก่อนตัดสินใจซื้อนั่นคือรีวิวจากลูกค้าก่อนหน้านั้นนั่นเอง รีวิวมีส่วนสำคัญต่อชื่อเสียงของธุรกิจ หากคนที่มีแนวโน้มว่าจะใช้บริการได้อ่านรีวิวในแง่บวกมีโอกาสสูงกว่าคนคนนั้นใช้บริการธุรกิจนั้นๆ แต่หากมีรีวิวในแง่ลบ แม้ว່ไม่ใช่เรื่องจริง หรือรีวิวด้วยอคติเอนเอียง ผลลัพธ์อาจออกมาตรงกันข้าม และในฐานะเจ้าของธุรกิจย่อมไม่มีใครอยากให้เกิตรีวิวแย่ๆ กับตัวเอง นอกจากนี้ความสำคัญอื่นๆ ที่ได้รับจากการมีรีวิวสินค้าหรือรีวิวบริการ

- **ความน่าเชื่อถือ** เพราะก่อนเสียเงินไปซื้อของ หากดูแล้วร้านค้าที่กำลังสนใจมีรายละเอียดสินค้าที่ไม่น่าเชื่อถือ ไม่มีรีวิวจากลูกค้าว่าสั่งซื้อแล้วได้รับของมาสภาพไหน การจัดส่งไวหรือไม่ , ตอบคำถามลูกค้าดีไม่ดียังไง หากเป็นอย่างนี้ความเชื่อมั่นต่อร้านค้าก็ไม่เกิดขึ้น

- **ได้รับความคิดเห็นโดยตรงจากลูกค้า** บ่อยครั้งมีความคิดเห็นดี ๆ เพื่อนำไปพัฒนาปรับปรุงสินค้าและบริการ มาจากลูกค้าโดยตรง หลังจากที่ได้ทดลองใช้สินค้าแล้ว ซึ่งเป็นประสบการณ์จากผู้ใช้งานจริง ๆ ที่มีคุณค่า อาจช่วยให้พัฒนาสินค้าที่ตอบสนองความต้องการของลูกค้าได้มากที่สุด

- **ช่วยให้ค้นหาสินค้าเจอได้ง่ายขึ้น** เครื่องมือค้นหาเสิร์ชเอนจิน เช่น กูเกิ้ล (Google) ดึงเอาสินค้าที่มีคนรีวิวมาแสดงให้นักท่องเที่ยว เห็นด้วย โดยสินค้าได้รับการรีวิวหรือพูดถึงจากลูกค้ามีจำนวนมาก ยิ่งเป็นตัวชี้วัดถึงความน่าเชื่อถือและความน่าสนใจของสินค้ามากตามไปด้วย จนทำสนใจที่จะคลิกเข้าไปดูเว็บไซต์เพิ่มขึ้น

- **รีวิวด้านลบบางครั้งส่งผลดี** ความผิดพลาดเป็นสิ่งที่มีโอกาสเกิดขึ้นได้เสมอ หากเจ้าของร้านรับมือกับบางรีวิวที่ตำหนิการให้บริการได้ดีและแสดงออกถึงความจริงใจในการขออภัยลูกค้า รับผิดชอบในความผิดพลาดอย่างเหมาะสม ลูกค้าที่เข้ามาอ่านรีวิวจะรู้สึกถึงความน่าเชื่อถือ ช่วยเพิ่มการตัดสินใจในการใช้บริการมากขึ้น เพราะได้เห็นถึงการบริการหลังการขายอีกมุมหนึ่ง หากร้านค้ามีแต่คำชมอย่างเดียว ทำให้ความรู้สึกน่าสงสัยมากกว่าเกิดความไว้วางใจ

- **สินค้ามีรีวิว คนยิ่งสนใจ** เมื่อลูกค้ารายใหม่ ๆ เข้ามาเห็นรีวิวก่อนหน้านี้ จะเกิดแรงจูงใจและความกล้าที่แชร์ความเห็นเพิ่มเติมทิ้งไว้ให้คนอื่น ๆ ได้รับรู้ต่อไป ยิ่งสินค้าชิ้นไหนมีรีวิวใหม่ ๆ เพิ่มเติมสม่ำเสมอ คนให้ความสนใจเพิ่มตามไปด้วย "รีวิว" คือปัจจัยใหญ่ที่ช่วยให้สามารถขายสินค้าบนโลกออนไลน์ออกไปได้โดยง่าย หากหน้าเพจสินค้าไหนมีรีวิวแฉงบวกลบเยอะ ๆ แปลว่าสินค้าชิ้นนั้น ๆ บนเว็บสามารถซื้อได้อย่างปลอดภัยแน่นอน แต่ถ้าเป็นรีวิวแฉงลบ สินค้าชิ้นนั้น ๆ อาจดับจนไม่สามารถขายต่อได้เลยด้วยซ้ำ ซึ่งนั่นจึงทำให้บทความรีวิวหรือคอมเมนต์ที่มีมูลค่าขึ้นมา จนถึงบางครั้งต้องมีการจ้างเพื่อมาเขียนรีวิวปลอมกันเพื่อให้สามารถขายสินค้าได้

Amazon Fake Reviews Scam เปิดเผยในการละเมิดข้อมูล ตลาดออนไลน์ขนาดใหญ่ Amazon เป็นหนึ่งในการเผชิญกับการต่อสู้เพื่อจำกัดปัญหาที่แพร่หลายในขณะนี้ และต้องลดการหลอกลวง ผลการวิจัยเก็บข้อมูลล่าสุดของ SafetyDetectives กลุ่มนักวิจัยและผู้เชี่ยวชาญด้าน Cybersecurity ได้เปิดเผยข้อมูลใน ElasticSearch ฐานข้อมูลแบบเปิด พบว่ามีการเขียนรีวิวปลอมโดยคนขาย หรือคนที่ถูกว่าจ้างมาใน Amazon มากถึง 13,124,962 รายการ คิดเป็นปริมาณข้อมูลที่ใหญ่มากถึง 7 GB ด้วยกัน ได้แก่ ข้อมูลที่เกี่ยวข้องกับผู้ขาย ข้อมูลที่เกี่ยวข้องกับผู้รีวิว ซึ่งเกิดจากผู้ใช้งานมากกว่า 200,000 คนด้วยกัน โดยรีวิวปลอมเหล่านี้ คนที่เขียนและให้ดาวกับตัวสินค้าจะได้รับเงินค่าสินค้าผ่านทางระบบของ PayPal ทั้งนี้ข้อมูลต่าง ๆ สามารถตรวจสอบได้ เพราะคนเขียนรีวิวเหล่านี้มี E-Mail ที่ผูกกับ PayPal (มีปรากฏออกมา 232,664 รายการ), ลิงค์บัญชี Amazon และ Username (บันทึกไว้กว่า 75,000 รายการ) เก็บไว้ทั้งหมด ทีมความปลอดภัยทางไซเบอร์ของ

SafetyDetectives ค้นพบการละเมิดเมื่อวันที่ 1 มีนาคม 2021 ตรวจสอบสถานะของเซิร์ฟเวอร์ ElasticSearch ที่เปิดอยู่ในวันต่อมา และในวันที่ 6 มีนาคม 2021 ฐานข้อมูลที่ไม่มีการอ้างอิงสิทธิ์ได้รับการรักษาความปลอดภัย สิ่งที่ชัดเจนคือใครก็ตามที่เป็นเจ้าของเซิร์ฟเวอร์อาจถูกลงโทษตามกฎหมายคุ้มครองลูกค้า และใครที่จ่ายเงินสำหรับรีวิวลอมนี้อาจต้องเผชิญกับการคว่ำบาตรสำหรับการละเมิดข้อกำหนดในการให้บริการของ Amazon ซึ่งขนาดและผลกระทบของปัญหาหมายความว่าทุกคนควรทำเท่าที่ทำได้เพื่อระบุรีวิวลอมเมื่อใช้ตลาดออนไลน์ เช่น Amazon การระบุและการรายงานรีวิวที่สงสัยว่าเป็นรีวิวลอมจะช่วยป้องกันตนเอง การรายงานเหตุการณ์ดังกล่าวสามารถให้การสนับสนุนที่สำคัญแก่ตลาดกลางที่พยายามควบคุมปัญหา โดยวิธีสังเกตรีวิวลอมทางออนไลน์ มีดังนี้ (SafetyDetectives, 2021)

- สงสัยเกี่ยวกับความคิดเห็นที่รุนแรง ผลิตภัณฑ์ที่ 'สมบูรณ์แบบ' ไม่ค่อยมีอยู่จริง หากผลิตภัณฑ์มีรีวิวในเชิงบวกมากเกินไป (โดยเฉพาะอย่างยิ่งเมื่อเทียบกับผลิตภัณฑ์ที่คล้ายคลึงกัน) ควรตั้งคำถามถึงความถูกต้องกฎหมายของรีวิวเหล่านั้น นอกจากนี้ควรมองหารีวิวที่เป็นบวก 100% หรือลบ 100%
- มองหาภาษาที่น่าสงสัย รีวิวลอมใช้ภาษาที่ใช้อารมณ์น้อยกว่า และอาจอ่านยาก รีวิวลอมอาจดูเหมือนเป็นโฆษณา เป็นการดูหมิ่นคู่แข่งของผลิตภัณฑ์ในกระบวนการนี้
- มองหาข้อความทั่วไปเกี่ยวกับผลิตภัณฑ์ รีวิวระดับห้าดาวหลายรายการอาจเน้นจุดบวกที่เหมือนกัน หรือรีวิวโดยทั่วไปอาจขาดความแปรปรวน ไม่เปิดเผยอะไรเกี่ยวกับประสบการณ์เฉพาะของแต่ละคน รีวิวลอมอาจมีคำหลักทั่วไปจำนวนมาก ใช้คำซ้ำเยอะกว่าปกติ หรืออ้างอิงชื่อแบรนด์ไปจนถึงการอ้างสรรพคุณและสเปคของตัวสินค้าซ้ำ ๆ หลายครั้งอย่างผิดปกติ
- รีวิวลอมอาจสั้นลง หากรีวิวยาวเพียงไม่กี่คำ ผู้เขียนรีวิวอาจพยายามส่งผลกระทบต่อระดับดาวของผลิตภัณฑ์โดยเร็วที่สุด
- ระมัดระวังเป็นพิเศษเมื่อซื้อจากแบรนด์ที่ไม่รู้จัก การเริ่มต้นใหม่พยายามยกระดับสถานะด้วยรีวิวลอม ตรวจสอบรีวิวผลิตภัณฑ์ในเว็บไซต์อื่นๆ หากก่อนที่ซื้อต้องตรวจสอบว่ามีรายละเอียดสถานที่ติดต่อถูกต้องหรือไม่หากเกิดสิ่งผิดปกติ
- ตรวจสอบข้อมูลที่ไม่เกี่ยวข้อง 'การรวมรีวิว' เป็นเรื่องปกติสำหรับผู้ขายที่มีความผิด ซึ่งเผยแพร่รีวิวซ้ำจากผลิตภัณฑ์อื่น ๆ เป็นของตนเอง รีวิวลอมอาจมีตัวอย่างอื่นๆ ของข้อมูลที่เท็จด้วยเช่นกัน
- ตรวจสอบรีวิวระดับห้าดาวกับคำวิจารณ์ที่ไม่ดี รีวิวที่ไม่ดีอาจเน้นย้ำถึงปัญหาที่รีวิวระดับห้าดาวปลอมไม่รับรู้อย่างต่อเนื่อง รีวิวลอมอาจกล่าวได้ว่าคุณลักษณะนี้ของผลิตภัณฑ์เป็นไปในเชิงบวก

- ตรวจสอบบัญชีของผู้ตรวจสอบ หากแสดงความคิดเห็นในเชิงบวกเกี่ยวกับผลิตภัณฑ์ของผู้ขายรายเดียวกันจำนวนมาก อาจเป็นของปลอม และสามารถพูดได้เช่นเดียวกันหากออกความคิดเห็นเชิงลบ หากบัญชีไม่มีข้อมูลส่วนบุคคลและพฤติกรรมการซื้อขายเป็นแบบสุ่ม นั่นเป็นอีกสัญญาณหนึ่งของผู้ตรวจสอบปลอม

- ตรวจสอบรูปแบบ ความคิดเห็นเชิงลบอาจตามมาด้วยกลุ่มรีวิวดระดับห้าดาวปลอม นอกจากนี้ รีวิวจำนวนหนึ่งอาจฟังดูคล้ายคลึงกัน หรือผู้เขียนรีวิวปลอมอาจโพสต์รีวิวที่คล้ายกันในผลิตภัณฑ์หลายรายการ

- ตรวจสอบวันที่ของรีวิว หากมีการโพสต์รีวิวดระดับห้าดาวของผลิตภัณฑ์ก่อนที่มีการแสดงผลิตภัณฑ์หรือในช่วงเวลาสั้นๆ รีวิวเหล่านั้นอาจเป็นของปลอม

- ใช้ซอฟต์แวร์ มีเครื่องมือออนไลน์ดีๆ มากมายที่วิเคราะห์คำวิจารณ์ของผลิตภัณฑ์และบอกว่าดูเหมือนปลอมหรือไม่ใช้

การป้องกันการเปิดเผยข้อมูล

- หากกังวลเกี่ยวกับข้อมูลควรดำเนินการสองสามขั้นตอนในทันทีเพื่อลดความเสี่ยงที่ถูกเปิดเผยและลดผลกระทบจากอาชญากรรมทางอินเทอร์เน็ต

- ให้ข้อมูลแก่บริษัท/บุคคลที่รู้จักหรือเชื่อถือได้เท่านั้น

- ตรวจสอบให้แน่ใจว่าเว็บไซต์ที่เข้าชมนั้นปลอดภัย โดเมนเว็บไซต์ที่ปลอดภัยมี 'https' และ/หรือสัญลักษณ์ล็อคแบบปิดที่จุดเริ่มต้น

- อย่าให้ข้อมูลที่สามารถนำมาใช้ได้ง่าย (ควรเก็บหมายเลขประจำตัวประชาชนและความชอบส่วนบุคคล)

- ใช้ตัวอักษร ตัวเลข และสัญลักษณ์เพื่อสร้างรหัสผ่านที่มั่นคง

- อย่าคลิกลิงก์ในอีเมล (หรือที่ใดทางออนไลน์) ที่ไม่แน่ใจว่ามาจากแหล่งที่เชื่อถือได้

- ตรวจสอบให้แน่ใจว่าการตั้งค่าความเป็นส่วนตัวของบนเว็บไซต์โซเชียลมีเดียแสดงเฉพาะเนื้อหาและข้อมูลส่วนบุคคลต่อบุคคลที่เชื่อถือได้

- หลีกเลี่ยงการใช้บัตรเครดิตหรือพิมพ์รหัสผ่านบนเครือข่าย Wi-Fi ที่ไม่ปลอดภัย

- ให้ความรู้เกี่ยวกับการปกป้องข้อมูล อาชญากรรมทางอินเทอร์เน็ต และวิธีต่างๆ ที่สามารถหลีกเลี่ยงการโจมตีแบบฟิชชิ่งและแรนซัมแวร์ได้

ความเห็นออนไลน์เป็นช่องทางแรกของคนสมัยนี้โดยเฉพาะเวลาที่กำลังมองหาร้านอาหารและโรงแรม ดังนั้นจึงกลายเป็นส่วนสำคัญของธุรกิจจำนวนมาก เพราะคนเหล่านี้จะกลายมาเป็นบันทึกที่มีชีวิตที่อยู่ไปอีกนาน ในการบอกต่อคุณภาพของบริการและผลิตภัณฑ์ และธุรกิจเหล่านี้ต้อง

ต้นตัวอยู่ตลอดเวลาเพื่อเฝ้าดูความคิดเห็นที่ไม่เป็นธรรมหรือรีวิวปลอมต่างๆ ที่ถูกสร้างขึ้นโดยคู่แข่งที่เล่นสกปรก หรือลูกค้าที่ไม่พอใจในสินค้าและบริการ แต่เร็วๆ นี้จะมีภัยคุกคามใหม่ที่สำคัญต่อโลกของการเขียนรีวิว นั่นคือการเขียนรีวิวปลอมโดยใช้ ปัญญาประดิษฐ์ (AI) ถูกพัฒนาเพิ่มขึ้นอย่างมากโดยไม่ได้ผ่านการตรวจสอบ สามารถทำให้เว็บไซต์รีวิวทั้งหลายเสียความน่าเชื่อถือลงไปเยอะ เทคโนโลยีมีขอบเขตที่กว้างขึ้นและน่าเป็นห่วงขึ้น ซึ่งอาจส่งผลกระทบต่อสังคม ความน่าเชื่อถือ และข่าวปลอมที่อาจเกิดมากขึ้นตามมา Ben Y. Zhao ศาสตราจารย์ด้านวิทยาการคอมพิวเตอร์ แห่งมหาวิทยาลัยชิคาโก กล่าวว่า “นี่เป็นภัยคุกคามที่กำลังใหญ่ขึ้นเรื่อย ๆ และกระจายตัวไปในสังคมที่กลุ่มผู้ใช้มีขนาดใหญ่และไม่ค่อยสนใจอะไร และจะสั่นคลอนความเชื่อมั่นในการตัดสินว่าอะไรจริงอะไรไม่จริง” เว็บไซต์รีวิวต่างๆ มีหน้าที่ในการระบุชี้และดำเนินการป้องกันบุคคลผู้ที่พยายามส่งรีวิวปลอมเข้ามา รับผิดชอบต่อเรื่องนี้อย่างเคร่งครัดที่สุด และเป็นเช่นนี้มานานกว่าหนึ่งทศวรรษแล้วที่ Tripadvisor ได้ลงทุนทุ่มเทในด้านเทคโนโลยีใหม่และทีมผู้ตรวจสอบที่เชี่ยวชาญเพื่อตรวจจับรีวิวปลอม

รีวิวปลอม

หมายถึงรีวิวใดๆ ที่ส่งเข้ามาโดยผู้ที่มีอคติหรือไม่เคยมีประสบการณ์ใช้บริการธุรกิจที่ตนเขียนรีวิว โปรดทราบว่ารีวิวปลอมไม่เหมือนกับรีวิวที่มีการโต้แย้ง (เป็นรีวิวที่เขียนบอกเล่าประสบการณ์โดยลูกค้าที่ไม่ได้มีความเกี่ยวข้องกับธุรกิจนั้นซึ่งถูกธุรกิจนั้นโต้แย้ง) รีวิวปลอมมีคุณสมบัติอยู่ในหมวดหมู่หนึ่งใน 3 หมวดหมู่ต่อไปนี้ รีวิวเชิงบวกที่ลำเอียง รีวิวเชิงลบที่ไม่เป็นธรรม และรีวิวที่ถูกจ้างให้เขียน(Tripadvisor, 2020)

1. **รีวิวเชิงบวกที่ลำเอียง** คือ เมื่อมีบางคนมีความเกี่ยวข้องกับธุรกิจ เช่น เป็นเจ้าของธุรกิจ พนักงาน เป็นเพื่อนหรือญาติ พยายามที่จะโพสต์รีวิวเชิงบวกให้แก่ธุรกิจนั้น ยังเรียกการทำเช่นนี้ว่า การส่งเสริมโดยใช้รีวิว รีวิวในลักษณะนี้ไม่ได้ให้ความเห็นที่เป็นกลางเกี่ยวกับประสบการณ์ที่ลูกค้าได้รับจริง เนื่องจากผู้เขียนรีวิวมีความเกี่ยวข้องส่วนตัวหรือได้รับผลประโยชน์จากธุรกิจนั้น รีวิวเชิงบวกที่ลำเอียงยังสามารถเกิดขึ้นได้เมื่อธุรกิจเสนอสิ่งจูงใจให้ลูกค้า เช่น อาหารฟรีหรือส่วนลด เพื่อแลกกับการเขียนรีวิวเชิงบวก มีการตรวจจับรีวิวเชิงบวกที่ลำเอียงด้วยระบบการติดตาม ซึ่งตรวจจับรูปแบบและตัวชี้วัดทางออนไลน์ที่อาจระบุชี้ว่ามีเพื่อน ครอบครัว หรือพนักงานคนใดกำลังพยายามส่งรีวิวให้ธุรกิจหรือไม่ สนับสนุนให้ชุมชนแจ้งให้ทราบเมื่อพบเห็นธุรกิจที่เสนอสิ่งจูงใจเพื่อแลกกับการเขียนรีวิว เมื่อทราบจะทำการตรวจสอบรายงานเหล่านี้ และดำเนินการแก่ธุรกิจใดๆ ที่ถูกจับได้ว่าพยายามรับรีวิวเชิงบวกที่ส่งเข้ามาด้วยวิธีนี้

2. **รีวิวเชิงลบที่ไม่เป็นธรรม** คือ เมื่อมีบางคนส่งรีวิวเชิงโจมตีสถานที่ให้บริการโดยเจตนาเพื่อพยายามลดการจัดอันดับของธุรกิจนั้นลงอย่างไม่เป็นธรรม หรือเพื่อทำให้สถานที่ให้บริการนั้นเสียชื่อเสียงในทางใดทางหนึ่งอย่างไม่เหมาะสม ยังเรียกการทำเช่นนี้ว่า การก่อกวนโดยใช้รีวิว รีวิวเชิงลบที่ไม่เป็นธรรมส่วนใหญ่มาจากหนึ่งในสองแหล่ง คือ มาจากบางคนที่มีความเกี่ยวข้องกับธุรกิจคู่แข่ง

หรือมาจากบางคนที่พยายามขู่กรรโชกธุรกิจโดยข่มขู่ว่าจะส่งรีวิวเชิงลบที่ไม่เป็นจริง มีหลายวิธีที่ตรวจจับรีวิวประเภทนี้ วิธีการคล้ายกับการตรวจจับรีวิวเชิงบวกที่ลำเอียง ระบบการติดตามสามารถตรวจจับลักษณะของผู้เขียนรีวิวที่อาจจะบ่งชี้ได้ว่าผู้เขียนรีวิวมีความเกี่ยวข้องกับธุรกิจคู่แข่งหรือไม่ แม้ว่าผู้เขียนรีวิวพยายามปกปิดร่องรอยของตน แต่รีวิวจะไม่เป็นไปตามรูปแบบที่เจอจากตัวอย่างของรีวิวที่มาจากลูกค้าจริงๆ ระบบสามารถตรวจจับรีวิวประเภทนี้ได้และส่งเข้าสู่กระบวนการตรวจสอบเหนือสิ่งอื่นใด หากมีเครื่องมือที่เจ้าของธุรกิจสามารถใช้เพื่อรายงานเหตุการณ์เมื่อมีคนขู่กรรโชกเจ้าของธุรกิจว่าจะเขียนรีวิวแย่ๆ หากไม่ได้รับส่วนลดหรือของกำนัล

3. **รีวิวที่ถูกจ้างให้เขียน** หมายถึง เมื่อธุรกิจจ้างหรือใช้บริการบุคคลหรือบริษัทให้เขียนรีวิวเชิงบวกเพื่อเพิ่มตำแหน่งการจัดอันดับของตนบน Tripadvisor ทั้งโดยเจตนาหรือรู้เท่าไม่ถึงการณ์ เรียกวิธีการทำเช่นนี้ว่า การเพิ่มประสิทธิภาพโดยใช้รีวิว

วิธีลบรีวิวเชิงลบปลอมบน Google ธุรกิจส่วนใหญ่ที่อ่านทางออนไลน์ หากธุรกิจมีรีวิวระดับสี่หรือห้าดาวในปริมาณที่เหมาะสม มีแนวโน้มลองใช้มากขึ้น แต่รีวิวระดับหนึ่งหรือสองดาวเพียงสองสามดวงอาจทำให้สงสัยได้ การจัดการชื่อเสียงเป็นศิลปะที่มีอิทธิพลและควบคุมชื่อเสียงของบุคคลหรือธุรกิจ และในปี 2020 มีความสำคัญมากกว่าที่เคย สถิติบางส่วนสำหรับรีวิวออนไลน์ มีดังนี้ (google, 2021)

- 86% ของลูกค้าอ่านรีวิว เกี่ยวกับธุรกิจในห้องถิ่น (รวมถึง 95% ของผู้ที่มีอายุ 18-34 ปี)
- ข้อมูลของ Moz สัญญาณการตรวจสอบคิดเป็น 15% ของปัจจัยการจัดอันดับแพ็คเกจในพื้นที่ของ Google
- ลูกค้าอ่านรีวิวออนไลน์โดยเฉลี่ย 10 รายการ ก่อนรู้สึกว่าจะสามารถไว้วางใจธุรกิจในห้องถิ่นได้
- 57% ของลูกค้า ใช้เฉพาะบริษัทที่มีดาวรีวิวสี่ดาวขึ้นไปเท่านั้น

แนวโน้มที่จะเชื่อถือรีวิวนี้เป็นเพียงธรรมชาติของมนุษย์ และเมื่อพูดถึงรีวิวธุรกิจออนไลน์ โพสต์เชิงลบสองสามโพสต์สามารถสร้างความเสียหายอย่างมหาศาลให้กับบริษัทได้

รีวิวเชิงลบปลอมคืออะไร บางครั้งบริษัทได้รับรีวิวเชิงลบที่ดูเหมือนเป็นเท็จ สิ่งเหล่านี้ถูกโพสต์ว่าเป็นสแปมหรือบางครั้งเป็นคู่แข่งที่ผิดจรรยาบรรณโดยทั่วไปแล้ว รีวิวปลอมเป็นหนึ่งในดาวและหรือมีการร้องเรียนที่ไม่มีมูลในความคิดเห็นน่าเสียดาย แม้ว่ารู้ว่ารีวิวนั้นเป็นของปลอม แต่ผู้มีโอกาสเป็นลูกค้าก็ไม่วางใจ การวิจัยล่าสุดแสดงให้เห็นว่า 74% ของลูกค้าอ่านรีวิวปลอมในปีที่ผ่านมาบอกได้ อย่างไรก็ตาม รีวิวของ Google เป็นของปลอม ลำดับแรกของธุรกิจในการจัดการกับรีวิวปลอมของ Google คือการค้นหาวารีวิวที่นั้นเป็นของปลอมหรือไม่ ถ้าผลการตรวจมาจากลูกค้าจริงและไม่มีความสุขวิธีที่ได้รับการตรวจสอบออกไม่ได้และดีที่สุดที่ตอบสนองอย่างรวดเร็ว

รีวิวปลอมมองเห็นได้ง่าย สิ่งที่ต้องดูเมื่อต้องระบุรีวิวที่เป็นรีวิวปลอม

- ดูชื่อผู้เขียนรีวิวและตรวจสอบเพื่อดูว่าอยู่ที่ใดในซอฟต์แวร์ CRM / จุดขาย หากไม่พบบันทึกใดๆ แสดงว่ารีวิวอาจเป็นของปลอม
- คลิกที่โปรไฟล์ของผู้เขียนรีวิวและตรวจสอบว่าได้เขียนรีวิวที่คล้ายคลึงกันกับธุรกิจอื่นๆ หรือไม่
- บางครั้งผู้ตรวจสอบอาจเข้าใจผิดว่าธุรกิจอาจเป็นธุรกิจอื่นและเขียนรีวิวให้บริษัท หากรีวิวกล่าวถึงผลิตภัณฑ์หรือบริการที่ไม่ได้นำเสนอ ผู้ตรวจทานอาจเข้าใจผิดว่าเป็นแบรนด์อื่น
- หากพบรีวิวเชิงลบหลายครั้งในระยะเวลาอันสั้น อาจเป็นนักส่งสแปมหรือคู่แข่งที่กำหนดเป้าหมายธุรกิจ

วิธีการรายงานและลบรีวิว Google ปลอม

ขั้นตอนเริ่มต้นในการรายงานรีวิวต่อ Google นั้นค่อนข้างตรงไปตรงมาและทำตามขั้นตอนง่ายๆ ไม่กี่ขั้นตอน ปัญหาที่แท้จริงอยู่ที่ว่า Google จะตอบสนองต่อคำขอในเวลาที่เหมาะสมหรือไม่

- ไปที่รีวิวที่พิจารณาแล้วว่าปลอมและค้นหาจุดแนวตั้งสามจุดทางด้านขวา
- การคลิกที่จุดจะแสดงตัวเลือกในการติตตรงว่าไม่เหมาะสม หากแน่ใจว่ารีวิวนั้นเป็นของปลอม ให้ตั้งค่าสถานะรีวิวนั้น Google อาจลงโทษบริษัทที่รายงานรีวิวที่พบว่าถูกต้องตามกฎหมาย
- จากนั้นจะถูกย้ายไปยังหน้าการละเมิดนโยบาย ซึ่งระบบขอให้ป้อนอีเมลและเลือกการละเมิดที่เชื่อว่าการตรวจสอบมีความผิด

หลักเกณฑ์และการละเมิดรีวิวของ Google

- สแปมและเนื้อหาปลอม
- นอกเรื่อง
- เนื้อหาที่ถูกจำกัด
- เนื้อหาที่ผิดกฎหมาย
- เนื้อหาเกี่ยวกับเรื่องเพศอย่างโจ่งแจ้ง
- เนื้อหาที่ไม่เหมาะสม
- เนื้อหาที่เป็นอันตรายและเสื่อมเสีย
- การแอบอ้างบุคคลอื่น
- ขัดผลประโยชน์

ส่งคำขอผ่าน Google My Business

- ไปที่ โพรไฟล์ Google My Business
- คลิกการสนับสนุนด้านล่างของหน้า
- กล้องความช่วยเหลือปรากฏขึ้นพร้อมตัวเลือก "ต้องการความช่วยเหลือเพิ่มเติม" "รีวิว และรูปภาพของลูกค้า" "จัดการรีวิวของลูกค้า" และสุดท้ายคือ "การสนับสนุนทางอีเมล" หากมีหลักฐานเพียงพอที่จะสนับสนุนกับการตรวจสอบ ส่งอีเมลไปยังฝ่ายสนับสนุนลูกค้าเพื่อสรุปสถานการณ์

หาก Google ปฏิเสธคำขอ แม้ว่าแน่ใจว่ารีวิวนั้นต้องเป็นของปลอม Google อาจไม่ลบรีวิวนั้นออก โดยเฉพาะอย่างยิ่ง Google จะลบรีวิวต่อเมื่อเห็นว่าเป็นการละเมิดนโยบายโดยตรงเท่านั้น ในทางกลับกันอาจตัดสินใจไม่ได้ให้หลักฐานเพียงพอ หากเป็นกรณีนี้สามารถติดตามข้อมูลเพิ่มเติมได้ หากท้ายที่สุดแล้ว Google ตัดสินใจที่ไม่นำรีวิวออก อย่างน้อยควรตอบกลับรีวิวปลอมและตั้งเป้าที่จะแก้ไขปัญหาที่กล่าวถึง คำตอบนี้แสดงให้เห็นในอนาคตว่าธุรกิจมีความกระตือรือร้นในการตอบสนองต่อคำวิจารณ์อย่างมีอารยธรรม ตัวอย่างเช่น “ขอบคุณที่สละเวลาเขียนรีวิว และให้ความสำคัญกับความคิดเห็นทั้งหมด แต่ไม่พบบันทึกของลูกค้าที่ตรงกับชื่อและข้อมูลในระบบ ยินดีที่จะหารือเกี่ยวกับข้อกังวลนี้เพิ่มเติมและช่วยเหลือในทุกวิถีทางที่สามารถทำได้”

งานวิจัยรีวิวออนไลน์ (Online Reviews)

แพรวพอลิน มณีขัติย์ (2563) อิทธิพลของรีวิวออนไลน์ที่ส่งผลต่อพฤติกรรมการวางแผนท่องเที่ยวของนักท่องเที่ยว ชาวไทยกลุ่มเจนเนเรชันวาย การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาพฤติกรรมการท่องเที่ยวของกลุ่มเจนเนเรชันวาย ปัจจัยที่เกี่ยวข้องกับ การรับชมรีวิวการท่องเที่ยวออนไลน์ ศึกษาอิทธิพลของรูปแบบการรีวิวการท่องเที่ยว ออนไลน์ และความ พึงพอใจในการเปิดรับรีวิวการท่องเที่ยวผ่านทางสื่อออนไลน์ที่ส่งผลต่อพฤติกรรมการวางแผนการท่องเที่ยว โดยเป็นการวิจัยเชิงปริมาณโดยใช้แบบสอบถามเก็บข้อมูลจากกลุ่มตัวอย่าง นักท่องเที่ยวชาวไทยเจนเนเรชันวาย จำนวน 385 คน เครื่องมือที่ใช้ในการวิจัยคือ แบบสอบถาม สถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ การหา ค่าเฉลี่ย (\bar{X}) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) วิเคราะห์ความแตกต่างค่าเฉลี่ยสองกลุ่ม (Independent-samples t-test) การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA) และ ChiSquare Test สำหรับการหาค่าความเป็นอิสระต่อกันระหว่างตัวแปร ผลการวิจัยพบว่า ผู้ตอบแบบสอบถามส่วนใหญ่เป็นเพศหญิง อายุ 24-29 ปี ศึกษาในระดับปริญญาตรี หรือเทียบเท่า ประกอบอาชีพพนักงานบริษัทเอกชน มีรายได้เฉลี่ยต่อเดือน 20,001-30,000 บาท ส่วนใหญ่ มีลักษณะการท่องเที่ยวกับครอบครัวและเลือกที่ท่องเที่ยวด้วยตนเอง รวมถึงใช้สื่อสังคมออนไลน์ Social Network เช่น Facebook เป็นจำนวนมากและบ่อยที่สุด โดยเฉลี่ย 1-2 ครั้งต่อสัปดาห์ และ

ใช้สื่อ 1-2 ประเภทในการค้นหา รีวิวเป็นเวลา 1-2 ชั่วโมง เหตุผลในการใช้เพื่อการค้นหาการรีวิว สถานที่ท่องเที่ยว ส่วนใหญ่ ค้นหาการรีวิวเพื่อดูรูปภาพและวิดีโอสถานที่ท่องเที่ยวมากที่สุดเพื่อ ประเมินทางเลือก ผลการทดสอบสมมติฐานการวิจัย พบว่า เมื่อพิจารณาข้อมูลทั่วไปและพฤติกรรม การรับรู้ข้อมูลการรีวิว ผ่านช่องทางสื่อออนไลน์โดยรวมมีผลไม่แตกต่างกันในพฤติกรรมการวางแผนการท่องเที่ยวของนักท่องเที่ยว กลุ่มเจนเนเรชันวาย แต่เมื่อพิจารณาเฉพาะด้านพบว่า มีข้อมูลบาง ด้านคือ เพศ อายุ และความถี่ในการใช้สื่อ สังคมออนไลน์มีผลแตกต่างกันในการวางแผนการ ท่องเที่ยวของนักท่องเที่ยวกลุ่มเจนเนเรชันวาย ส่วนด้าน อิทธิพลของแหล่งข้อมูลรีวิวผ่านทางสื่อสังคม ออนไลน์และความพึงพอใจในการเปิดรับรีวิวผ่านทางสื่อ ออนไลน์นั้นพบว่า มีผลต่อการวางแผนการ ท่องเที่ยวของนักท่องเที่ยวกลุ่มเจนเนเรชันวายในทุกด้าน

Daive Proserpio et al (2020) พบว่าลูกค้าจะทำการรีวิวปลอม ซึ่งในระหว่างเดือน เมษายน ถึง มิถุนายน 2020 ตลาดอีคอมเมิร์ซในสหรัฐอเมริกามีการเติบโตมากเป็นประวัติการณ์ถึง 44.4% และมีแนวโน้มที่เติบโตอย่างต่อเนื่องเนื่องจากธุรกิจขนาดใหญ่และขนาดเล็กเปลี่ยนไปขายออนไลน์ เมื่อเผชิญกับการแพร่ระบาดของ Covid-19 เพื่อให้อีคอมเมิร์ซทำงานได้แพลตฟอร์มอย่าง Amazon จำเป็นต้องมีระบบการให้คะแนนและการตรวจสอบที่ช่วยให้ลูกค้าตัดสินใจซื้อทางออนไลน์ได้อย่าง มั่นใจ แต่เนื่องจากโดยทั่วไปรีวิวเหล่านี้เป็นปัจจัยสำคัญในอัลกอริทึมอันดับการค้นหา ดังนั้นจึงมี ผลกระทบอย่างมากต่อการแสดงผลิตภัณฑ์และการขายระบบเหล่านี้จึงสร้างแรงจูงใจที่มีประสิทธิภาพ สำหรับผู้ขายในการจัดการการอันดับผลิตภัณฑ์ผ่านรีวิวปลอมที่มีขนาดใหญ่และเฟื่องฟู กลไกที่พบ บ่อยที่สุดอย่างหนึ่งพบในการจัดการรีวิวเหล่านี้คือผ่านทางกลุ่ม Facebook ส่วนตัว ผู้ขายใช้กลุ่ม เหล่านี้ในการรับสมัครผู้คนเพื่อซื้อผลิตภัณฑ์ของตนและออกจากรีวิวระดับห้าดาวที่ฟังดูน่าเชื่อถือ จากนั้นจึงชดเชยค่าใช้จ่ายผ่าน PayPal ของสินค้าภาษีและค่าธรรมเนียมใด ๆ และในบางกรณีมีค่า คอมมิชชั่น \$ 5-10 นอกจากนี้ยังพบว่าบางครั้งกลุ่มเหล่านี้จะหายไปจากนั้นกลุ่มใหม่ที่คล้ายกันจะถูก แทนที่ในเกือบทันที การค้นพบที่สำคัญประการแรกของคือแม้ว่าปรากฏการณ์นี้อาจไม่เป็นที่รู้จัก แต่ เป็นที่แพร่หลายมาก จากการสังเกตคาดว่า มีผู้ขายมากถึง 4.5 ล้านคนที่มา รีวิวปลอมผ่านทางกลุ่ม Facebook เหล่านี้ในปีที่ผ่านมา ประการที่สอง รีวิวปลอมดูเหมือนจะพบได้บ่อยที่สุดสำหรับผลิตภัณฑ์ บางประเภท เช่น ผลิตภัณฑ์เหล่านี้มีราคาใกล้เคียงกับคู่แข่งโดยทั่วไปอยู่ในช่วงราคา \$ 15 - \$ 40 และโดยทั่วไปแล้วมีคะแนนสูงอยู่แล้วโดยมีคะแนนเฉลี่ย 4.4 และมีรีวิวเฉลี่ย 183 รายการ (ซึ่งบ่งบอก ว่ารีวิวที่มีอยู่แล้วจำนวนมากเหล่านี้ อาจเป็นของปลอม) แต่การเปลี่ยนแปลงล่าสุดในนโยบายของ Amazon เพื่อส่งเสริมให้ผู้ขายทั่วโลกเพิ่มขึ้นส่งผลให้ผู้ผลิตชาวจีนที่ขายสินค้าบนแพลตฟอร์มโดยตรง เพิ่มขึ้นอย่างมาก ผู้ขายรายใหม่เหล่านี้มีอัตรากำไรที่ไม่สูงนักและมีชื่อเสียงเพียงเล็กน้อยที่จะรักษาไว้ ได้ทำให้เกิดปัญหาด้านการควบคุมคุณภาพมากมาย เห็นได้ชัดว่า รีวิวปลอมก่อให้เกิดปัญหาที่แท้จริง สำหรับผู้ซื้อ แต่ไม่ใช่แค่ลูกค้าเท่านั้นที่ต้องทนทุกข์ทรมานเมื่อผู้ขายใช้ รีวิวปลอม เนื่องจาก รีวิวปลอม

ทำลายความไว้วางใจของลูกค้าในระบบการตรวจสอบแพลตฟอร์ม การค้นหาวิธีในการปราบปรามนี้ จะช่วยลดสิ่งที่ผิดกฎหมายบนแพลตฟอร์มโซเชียลมีเดียเพิ่มความไว้วางใจในแพลตฟอร์มอีคอมเมิร์ซ และช่วยให้ลูกค้าได้รับประสบการณ์เชิงบวกและเชื่อถือได้มากขึ้น

อังกฤษสั่งตรวจสอบแอมะซอน กูเกิล ปมรีวิวปลอมระบาด หน่วยงานกำกับดูแลด้านการแข่งขันและตลาด (CMA) ของอังกฤษมีคำสั่งให้สอบสวนแอมะซอนและกูเกิลอย่างเป็นทางการ ซึ่งทั้งสองบริษัทไม่สามารถจัดการกับรีวิวปลอมของสินค้าและบริการต่างๆ ได้ทำให้ทั้งสองบริษัทต้องรับผิดชอบและพร้อมดำเนินการในด้านต่างๆ หากผลตรวจสอบพบว่าทั้งสองบริษัทยังมีความบกพร่องในด้านการดำเนินการ” รีวิวปลอมของสินค้าและบริการต่างๆ ที่หลอกลวงลูกค้า” ซึ่งเป็นปัญหาสำคัญสำหรับธุรกิจอีคอมเมิร์ซ โดยแอมะซอนนั้นเป็นเป้าหมายหลักของผู้ขายที่ต้องการจะสร้างกระแสความสนใจทางออนไลน์ให้กับสินค้าของตน และแอมะซอนได้เรียกร้องให้บริษัทโซเชียลมีเดียต่างๆ ดำเนินการตรวจสอบรีวิวปลอมเหล่านี้ด้วย CMA มีการตรวจสอบปัญหารีวิวปลอมมาตั้งแต่เดือนพ.ค. 2563 โดยในช่วงต้นปี 2563 นั้น CMA ได้สั่งให้เฟซบุ๊กและอีเบย์ระงับการใช้งานบัญชีผู้ใช้และปิดกลุ่มที่มีการซื้อและขายรีวิวปลอมของสินค้าและบริการต่างๆ (สำนักข่าวอินโฟเควสท์, 2021)

2.5 การเรียนรู้ของเครื่อง (Machine Learning)

การเรียนรู้ของเครื่อง หรือ machine learning นั้น เป็นรูปแบบหนึ่งของการวิเคราะห์ข้อมูลที่ดำเนินการวิเคราะห์ด้วยแบบจำลองอย่างเป็นอัตโนมัติ ซึ่งเป็นสาขาหนึ่งของเทคโนโลยีด้าน AI (artificial intelligence) ซึ่งตั้งอยู่บนรากฐานแนวคิดที่ว่า ระบบต่าง ๆ นั้น สามารถที่เรียนรู้และมีปฏิสัมพันธ์กับชุดข้อมูลต่าง ๆ รวมถึงสามารถระบุ และทราบรูปแบบต่าง ๆ ที่เกิดขึ้น และนำไปสู่การตัดสินใจได้เองโดยไม่จำเป็นต้องพึ่งพามนุษย์อีกต่อไป (SAS Institute Inc, 2021)

วิวัฒนาการของ Machine learning

วิทยาการด้านการคำนวณและคอมพิวเตอร์นั้นมีพัฒนาการอย่างต่อเนื่องอยู่เสมอ ซึ่งส่งผลให้ความก้าวหน้าของเทคโนโลยี Machine learning ในปัจจุบันนี้รู้หน้ากว่าในอดีตอย่างมาก วิทยาการด้านการเรียนรู้ของเครื่องนี้เริ่มต้นขึ้นจากการวิเคราะห์รูปแบบหรือ pattern recognition และทฤษฎีที่ว่า คอมพิวเตอร์อาจสามารถเรียนรู้และค่อย ๆ พัฒนาการทำงานต่าง ๆ ด้วยตนเองได้ โดยมีต้องอาศัยการตั้งโปรแกรมอยู่ตลอดเวลา ซึ่งนี่เป็นต้นกำเนิดที่ทำให้นักวิจัยด้าน AI เกิดความสนใจว่า คอมพิวเตอร์และระบบนั้น สามารถที่เรียนรู้จากการมีปฏิสัมพันธ์กับข้อมูลได้หรือไม่ ทั้งนี้ คุณลักษณะสำคัญของวิทยาการนี้คือการทำงานแบบซ้ำ (iterative process) ซึ่งเป็นหัวใจของการพัฒนาขีดความสามารถ เนื่องจากระบบและแบบจำลองการวิเคราะห์มีโอกาสปรับตัวตามชุดข้อมูลที่ได้พบ และนำไปสู่การพัฒนาการทำงานด้วยตนเอง ระบบจะเรียนรู้และปรับแก้ไขความคลาดเคลื่อนจากการทำงานที่ผ่าน ๆ มา จนกระทั่งสามารถสร้างผลลัพธ์ที่เชื่อถือได้และทำซ้ำได้อย่างสม่ำเสมอ เป็นที่

น่าสนใจว่า นี่ไม่ใช่วิทยาการที่เพิ่งเกิดขึ้นมาใหม่ หากแต่เป็นวิทยาการที่กำลังเกิดการพัฒนาย่างก้าวกระโดด และมีพลวัตอย่างต่อเนื่อง แม้ว่าอัลกอริทึมสำหรับการเรียนรู้ของเครื่องจำนวนมากมีอยู่มานานแล้ว แต่พัฒนาการที่สำคัญยิ่งในปัจจุบันคือการเกิดขึ้นของขีดความสามารถในการนำการคำนวณทางคณิตศาสตร์ที่มีความซับซ้อนสูงยิ่งมาใช้งานสำหรับ big data ในการทำงานแบบซ้ำที่ต่อเนื่อง ซึ่งตัวอย่างของการนำเทคโนโลยี machine learning มาใช้งานที่โดดเด่นและแพร่หลายซึ่งคุ้นเคยมีดังนี้

- กระแสความตื่นตัวเกี่ยวกับรถยนต์ขับเคลื่อนอัตโนมัติที่พัฒนาขึ้นโดย Google ซึ่งนั่นคือหนึ่งในนวัตกรรมสำคัญที่มีรากฐานบนเทคโนโลยี machine learning นี้
- การแนะนำวิดีโอหรือภาพยนตร์ที่เหมาะสมบน Amazon หรือ Netflix นั่นคืออีกหนึ่งตัวอย่างของการใช้งาน machine learning ที่พบเห็นได้ในชีวิตประจำวัน
- กระแสความรู้สึกลูกค้าที่กล่าวถึงใน Twitter คือผลจากการผสมผสาน machine learning เข้ากับการวิเคราะห์ทางภาษาหรือ linguistic rule creation
- การตรวจจับการทุจริตเป็นรูปแบบการใช้ประโยชน์จากการเรียนรู้ของเครื่องที่เป็นที่นิยม

ความเชื่อมโยงระหว่าง Machine Learning และ Artificial Intelligence

ข้อแตกต่างสำคัญประการหนึ่งของเทคโนโลยีทั้งสอง คือ AI นั้น เป็นวิทยาการในเชิงกว้างที่มุ่งเน้นการให้คอมพิวเตอร์พัฒนาขีดความสามารถในการทำงานเลียนแบบมนุษย์ ในขณะที่ machine learning คือการใช้ขีดความสามารถเฉพาะบางด้านของเทคโนโลยี AI เพื่อฝึกทักษะการเรียนรู้แก่ระบบและอุปกรณ์ ขอเชิญรับชมวิดีโอนี้เพื่อช่วยให้เข้าใจถึงความสัมพันธ์ของเทคโนโลยีทั้งสองนี้ได้ดียิ่งขึ้น วิดีโอได้ทำการอธิบายรายละเอียดในภาพรวมของทั้งสองสาขาผ่านตัวอย่างและเกร็ดความรู้ที่สนุกสนาน

ความสำคัญของระบบการเรียนรู้ของเครื่อง

กระแสความสนใจเกี่ยวกับ Machine learning ในทุกวันนี้เกิดขึ้นได้จากปัจจัยเดียวกันกับที่กระตุ้นให้การทำเหมืองข้อมูลและการวิเคราะห์แบบ Bayesian Analysis ได้รับความนิยมมากอย่างไม่เคยมีมาก่อน เกิดปรากฏการณ์ต่าง ๆ เช่นการขยายตัวของทั้งปริมาณและความหลากหลายของข้อมูล หรือขีดความสามารถทางเทคโนโลยีการประมวลผลที่ทั้งมีต้นทุนต่ำลงและทรงพลังกว่าที่เคยมี รวมถึงระบบการจัดเก็บข้อมูลที่ประหยัดและเปี่ยมประสิทธิภาพ สิ่งเหล่านี้เป็นปัจจัยที่ทำให้การสร้างและใช้งานแบบจำลองเป็นไปได้ด้วยความรวดเร็วอย่างอัตโนมัติ เพื่อการทำงานวิเคราะห์ แบบจำลองสำหรับข้อมูลที่มีปริมาณมหาศาลและมีความซับซ้อนสูงยิ่ง ด้วยผลการทำงานที่แม่นยำและรวดเร็วแม้เมื่อรับมือกับงานขนาดใหญ่ และองค์กรต่าง ๆ ได้รับประโยชน์จากการพัฒนาและใช้งานแบบจำลองที่มีความแม่นยำ โดยเพิ่มโอกาสการค้นพบช่องทางสร้างกำไรอันงดงามทางธุรกิจ และป้องกันความเสี่ยงที่ซ่อนอยู่

สิ่งจำเป็นในการสร้างระบบเพื่อการเรียนรู้ของเครื่องที่มีประสิทธิภาพ

- ขีดความสามารถในการจัดการและจัดเตรียมข้อมูล
- อัลกอริทึมเพื่อการทำงาน ทั้งระดับพื้นฐานและระดับสูง
- กระบวนการทำงานแบบเป็นอัตโนมัติและการทำงานซ้ำแบบต่อเนื่อง
- ความสามารถในการยกระดับการทำงาน
- การใช้เทคนิค Ensemble Modeling

Machine Learning ในโลกปัจจุบัน

ในปัจจุบันนี้องค์กรและธุรกิจต่าง ๆ สามารถนำอัลกอริทึมมาใช้ในการสร้างแบบจำลองสำหรับหาความสัมพันธ์ของตัวแปรต่าง ๆ ซึ่งช่วยให้ได้ผลลัพธ์ข้อมูลสำหรับการตัดสินใจเชิงกลยุทธ์

ผู้ใช้งานหลัก ภาคส่วนธุรกิจและอุตสาหกรรมต่าง ๆ ที่ต้องทำงานกับข้อมูลปริมาณมหาศาล ล้วนแต่ตระหนักถึงความสำคัญของเทคโนโลยีด้านการเรียนรู้ของเครื่องหรือ machine learning นี้ โดยหากองค์กรใดมีขีดความสามารถในการสกัดข้อมูลเชิงลึกและนัยยะทางยุทธศาสตร์ออกจากกระแสของข้อมูลที่เกิดขึ้นได้นั้น ย่อมนำมาซึ่งประสิทธิภาพในการทำงานที่สูงขึ้นและส่งผลถึงขีดความสามารถในการแข่งขันที่ได้เปรียบคู่แข่งโดยปริยาย ภาคส่วนสำคัญที่เป็นผู้ใช้งานหลัก ดังนี้

- ธุรกิจเกี่ยวกับสุขภาพในภาคส่วนด้านการแพทย์ สุขภาพ และการสาธารณสุขนั้น machine learning กำลังเป็นที่สนใจเนื่องมาจากการพัฒนาเทคโนโลยีอุปกรณ์ด้านสุขภาพแบบสวมใส่ติดตัว (wearable devices) ซึ่งเก็บข้อมูลด้านสุขภาพร่างกายของคนไข้แบบเรียลไทม์ ซึ่งเทคโนโลยีดังกล่าวนี้ช่วยให้แพทย์และบุคลากรทางการแพทย์สามารถตรวจสอบสุขภาพของคนไข้และตรวจจับสัญญาณเตือนทางสุขภาพต่าง ๆ เพื่อนำไปสู่การตรวจวินิจฉัยโรคและแนวทางการรักษาที่ทันท่วงทีต่อไป

- หน่วยงานภาครัฐต่าง ๆ เช่นหน่วยงานที่รับผิดชอบด้านสาธารณสุขประโยชน์และการป้องกันสาธารณสุขภัยต่าง ๆ คือกลุ่มที่ได้รับประโยชน์จากการใช้งาน machine learning อย่างมาก เนื่องจากองค์กรลักษณะดังกล่าวมีทั้งโอกาสการเข้าถึงข้อมูลปริมาณมหาศาล และความจำเป็นในการใช้งานข้อมูลเหล่านั้นเพื่อการสกัดนัยยะเชิงนโยบาย

- ภาคส่วนด้านการตลาด อาจเคยเห็นการแนะนำสินค้าหรือโฆษณาต่าง ๆ ที่เว็บไซต์ประเมินว่า อาจเป็นที่สนใจ โดยการแนะนำโฆษณาเหล่านี้เกิดขึ้นจากการใช้งาน machine learning เพื่อประเมินความสนใจจากสินค้าที่ได้เคยซื้อหรือสืบค้นข้อมูลก่อนหน้านี้ ซึ่งทำให้ระบบสามารถประเมินและแสดงผลสินค้าอื่น ๆ ที่สนใจซื้อเพิ่มเติมต่อไปได้ ซึ่งความสามารถในการเก็บรวบรวมข้อมูล ไปจนถึงวิเคราะห์และสร้างรูปแบบสำหรับลูกค้าแต่ละราย (personalized shopping experience) ตามรสนิยมและความสนใจนั้น อาจกล่าวได้ว่าเป็นแนวโน้มสำหรับการค้าปลีกในยุคอนาคตนับจากนี้

วิธีการ Machine Learning

วิธีการเกี่ยวกับการเรียนรู้ของเครื่องที่เป็นที่แพร่หลายที่สุดสองกระบวนการได้แก่ วิธี supervised learning และ unsupervised learning แต่ยังมีวิธีการและเทคนิคอื่น ๆ อีกหลากหลายรูปแบบ และสิ่งนี้คือภาพรวมของเทคนิคที่ได้รับความนิยมมากที่สุดในปัจจุบัน

การเรียนรู้ของเครื่อง (Machine Learning) เป็นที่ยอมรับในด้านการประยุกต์ใช้งานกับแก้ไขปัญหาหลายด้าน หรือใช้สำหรับโปรแกรมประยุกต์มากมาย โดยแต่ละกระบวนการเรียนรู้มีความเหมาะสมกับข้อมูลในรูปแบบเฉพาะ ปัจจุบันมีการใช้เทคนิค Machine Learning (ML) ในอุตสาหกรรมต่างๆ พร้อมกับโครงการและความซับซ้อนที่เพิ่มขึ้น สิ่งนี้ทำให้เกิดความจำเป็นในการกำกับดูแลที่มากขึ้น กล่าวคือ ความสามารถในการประสานและควบคุมการพัฒนาและปรับใช้ตลอดวงจรชีวิตของ ML ทั้งหมด (การประมวลผลล่วงหน้า การฝึกโมเดล การทดสอบ การปรับใช้) ในทางกลับกัน ความจำเป็นในการปรับขยายได้ เช่น สามารถทำซ้ำทุกส่วนของกระบวนการได้อย่างมีประสิทธิภาพ เพื่อจัดการแบบจำลอง ML หลายแบบ โดยวงจรของ Machine Learning Life Cycle แบ่งขั้นตอนของการดำเนินงานออกเป็น 7 ขั้นตอน ได้แก่

1) การได้มาซึ่งข้อมูล (Data Acquisition) หมายถึง ข้อมูลซึ่งเป็นองค์ประกอบสำคัญอย่างหนึ่งในวงจรชีวิตของการเรียนรู้ของเครื่อง ต้องได้รับข้อมูลที่เหมาะสมเพื่อแก้ปัญหาเฉพาะอย่างใดอย่างหนึ่ง

2) การเตรียมข้อมูล (Data Preparation) หมายถึง กระบวนการใด ๆ ที่จำเป็นต้องทำกับข้อมูลดิบ (raw data) ที่ได้รับมา เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสม ที่นำไปโหลดเข้าฐานข้อมูลหรือนำไปวิเคราะห์ต่อไป

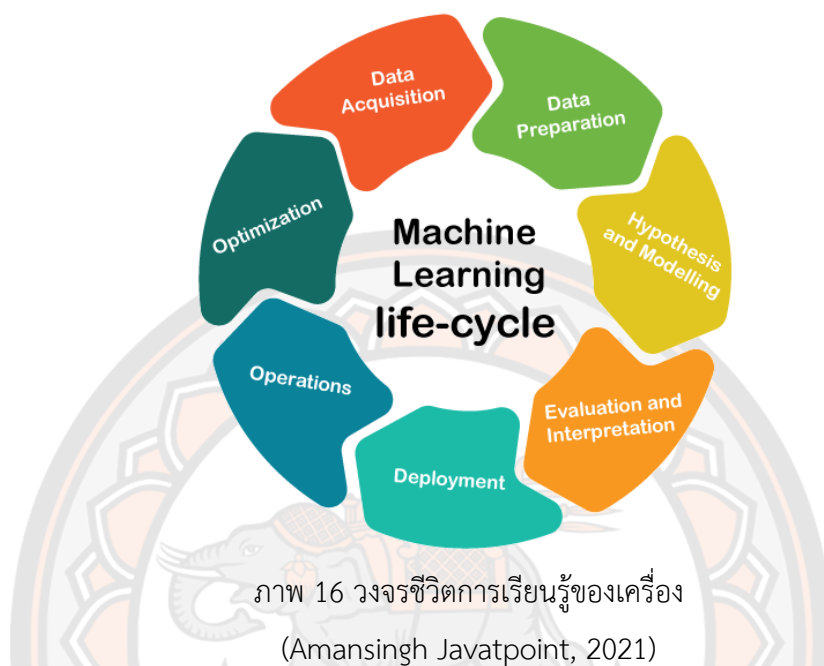
3) สมมติฐานและการสร้างแบบจำลอง (Hypothesis and Modeling) หมายถึง การหาวิธีการเพื่อสร้าง Model นั้นๆ ด้วย Training set (ข้อมูลทดลอง) เพื่อให้คาดเดาผลลัพธ์ผ่านกระบวนการเขียน ดำเนินการ และปรับแต่งโปรแกรมเพื่อวิเคราะห์และรับข้อมูลเชิงลึกทางธุรกิจที่มีความหมายจากข้อมูล โปรแกรมเหล่านี้สามารถเขียนเป็นภาษาต่างๆ เช่น Python, R, MATLAB หรือ Perl ในขั้นตอนนี้ใช้ข้อมูลของการฝึกอัลกอริธึมการเรียนรู้ของเครื่อง และเลือกอัลกอริธึมที่ให้ผลลัพธ์ที่ดีที่สุด

4) การประเมินผลและการตีความ (Evaluation and Interpretation) หลังจากฝึกโมเดลเกี่ยวกับข้อมูลแล้วประเมินผลลัพธ์สุดท้ายเพื่อดูว่าทำงานได้ดีเพียงใดหรือเชื่อถือได้เพียงใด

5) การปรับใช้ (Deployment) สามารถกำหนดเป็นแอปพลิเคชันของแบบจำลองที่ทำให้การคาดการณ์โดยใช้ข้อมูลใหม่ได้

6) ปฏิบัติการ (Operations) ในขั้นตอนนี้ได้มีการจัดทำแผนเพื่อกำกับดูแลข้อมูล ตรวจสอบประสิทธิภาพของโมเดลในแง่ของการอัปเดตและดาวน์โหลดได้

7) การเพิ่มประสิทธิภาพ (Optimization) เป็นขั้นตอนสุดท้าย หากโมเดลแมชชีนเลิร์นนิงมีแหล่งข้อมูลใหม่เข้ามาหรือทำตามขั้นตอนที่จำเป็นสำหรับการอัปเดตประสิทธิภาพของโมเดลแมชชีนเลิร์นนิง ดังภาพที่ 16



ภาพ 16 วงจรชีวิตการเรียนรู้ของเครื่อง
(Amansingh Javatpoint, 2021)

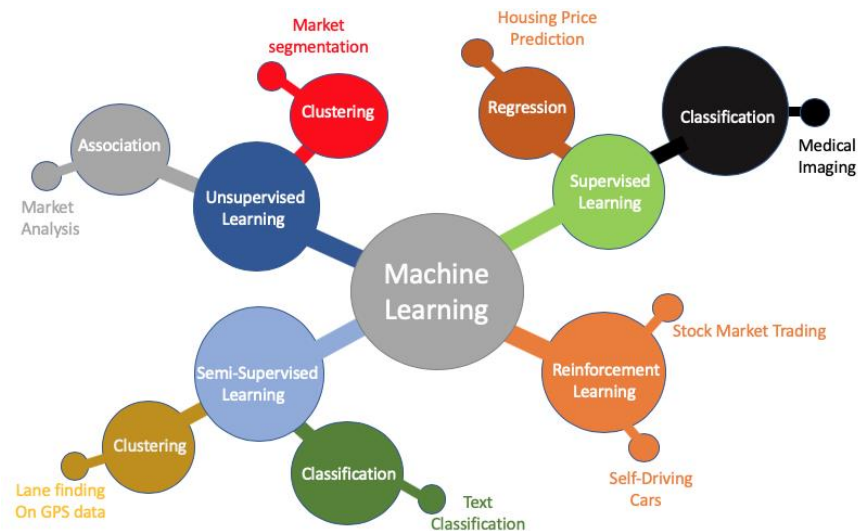
วงจรชีวิตการเรียนรู้ของเครื่อง แบ่งเป็นการเรียนรู้ได้ 4 วิธีตามลักษณะการทำงาน ได้แก่

1) Supervised Learning โมเดลที่เรียนรู้จากข้อมูลตัวอย่างที่ป้อนเข้าไป โดยข้อมูลตัวอย่างมีการระบุคำตอบอย่างชัดเจน (Labeled, Structured Data) โมเดลในกลุ่มนี้ที่พบบ่อย ได้แก่ การทำนายผล, การระบุประเภท (Classification) และ Regression

2) Unsupervised Learning โมเดลที่รับข้อมูลเข้าไปเพื่อค้นหาแพทเทิร์นและความสัมพันธ์ภายในชุดข้อมูลนั้นจะถูกใช้ในงานประเภทจัดกลุ่ม (Clustering)

3) Semi-supervised Learning โมเดลที่สามารถเรียนรู้จากข้อมูลที่มีผลลัพธ์ระบุ (Labeled Data) ในจำนวนน้อย และสามารถขยายผลไปทำงานกับข้อมูลอื่นๆได้ ส่วนมากเป็นการนำเทคนิค Unsupervised Learning มาเสริมการทำงานให้กับโมเดลแบบ Supervised Learning และพบบ่อยในงานประเภทการวิเคราะห์ภาพถ่าย หรือการวิเคราะห์ข้อความหรือคำพูด

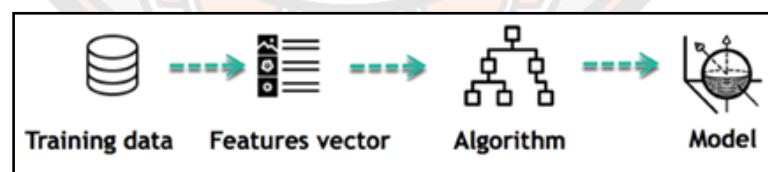
4) Reinforcement Learning โมเดลที่สามารถเรียนรู้ที่ทำงานผ่านการลองผิดลองถูก โดยตัวอย่างของเทคนิคที่ใช้ในการสร้างแบบจำลองการจำแนกข้อมูลที่มีความนิยมใช้กัน ได้แก่ Naïve Bayes, Decision Tree, Logistic Regression, K-Nearest Neighbor (KNN), Neural Network, Random Forest และ Support Vector Machine (SVM) ดังภาพที่ 17



ภาพ 17 วิธีการเรียนรู้ของเครื่อง
(Varun Achary, 2021)

วัตถุประสงค์หลักของ Machine Learning คือ การเรียนรู้ (Learning) และการอนุมาน (Inference) สิ่งแรกคือ Machine จะเรียนรู้ผ่านการค้นพบรูปแบบหรือแบบแผนซ้ำๆ เดิม ๆ การค้นพบการมีอยู่ของข้อมูล (Data) เป็นส่วนที่สำคัญของนักวิทยาศาสตร์ข้อมูล (Data Scientist) เป็นการเลือกว่าข้อมูลไหนที่มากการสนับสนุนของเครื่อง คุณสมบัติต่างๆ ที่ถูกใช้เพื่อแก้ปัญหาถูกเรียกว่า Feature Vector เป็นซั้บเซตย่อยของข้อมูลทั้งหมดที่ถูกใช้ในการแก้ปัญหาด้วยแบบจำลอง (Model) ดังนั้น ขั้นตอนการเรียนรู้ (Learning Stage) ถูกใช้เพื่ออธิบายเกี่ยวกับข้อมูลและรวบรวมให้กลายเป็นแบบจำลอง ดังภาพที่ 18

Learning Phase

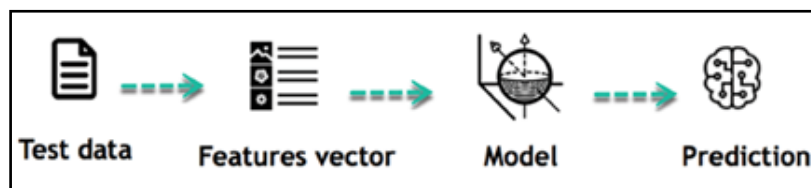


ภาพ 18 การเรียนรู้แบบจำลอง

ที่มา : <https://www.thaiprogrammer.org/>

การอนุมาน (Inference) เมื่อแบบจำลองถูกสร้างขึ้นมา ต้องถูกทดสอบตามสมรรถนะของข้อมูลที่ไม่เคยพบมาก่อน โดยข้อมูลที่มีความหลากหลายถูกแปลงให้เป็นการสร้างเวกเตอร์คุณลักษณะ (Feature Vector) และถูกรวบรวมให้กลายเป็นแบบจำลอง (Model) และทำการพยากรณ์ออกมาด้วย Machine Learning ซึ่งไม่ต้องเพิ่มเติมกฎหรือ Train เพิ่มเติมเข้าไปในแบบจำลอง สามารถใช้แบบจำลองก่อนหน้านี้ Train เพื่อทำการอนุมานข้อมูลใหม่ ๆ ต่อไปได้ ดังภาพที่ 19

Inference from แบบจำลอง



ภาพ 19 การอนุมาน

ที่มา : <https://www.thaiprogrammer.org/>

การเรียนรู้เชิงลึก (Deep Learning) คือ ศาสตร์แขนงหนึ่งของ Machine Learning ที่เลียนแบบระบบเซลล์ประสาทในสมองของมนุษย์ (Neural Network) โดยเพิ่มความสามารถของโครงข่ายประสาทเทียม (Artificial Neural Networks – ANN) คือ การสร้างคอมพิวเตอร์ที่จำลองเอาวิธีการทำงานของสมองมนุษย์หรือทำให้คอมพิวเตอร์รู้จักคิดและจดจำในแนวเดียวกับ โครงข่ายประสาทของมนุษย์เพื่อช่วยให้คอมพิวเตอร์ฟังภาษามนุษย์ได้เข้าใจ อ่านออก และรู้จำได้ซึ่งอาจเรียกได้ว่าเป็น “สมองกล” เป็นหนึ่งในเทคนิคของการทำเหมืองข้อมูล (Data Mining) คือ โมเดลทางคณิตศาสตร์สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนกชันนิสต์ (Connectionist) เพื่อจำลองการทำงานของโครงข่ายประสาทในสมองของมนุษย์ด้วยวัตถุประสงคที่สร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำรูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ๆ (Knowledge Extraction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์

Supervised learning นั้น คือเมื่ออัลกอริทึมได้รับการ "ฝึกฝน" ด้วยการใช้ตัวอย่างแบบมีเป้าหมายหรือ labeled example เช่น การให้ข้อมูลขาเข้า (inputs) โดยระบุถึงผลลัพธ์ขาออก (outputs) ที่ต้องการให้ทราบ ตัวอย่างเช่น อุปกรณ์ชิ้นหนึ่งอาจเรียนรู้โดยกำหนดผลลัพธ์เป็นสองทางคือ "F = failed" และ "R = runs" โดยอัลกอริทึมการฝึกฝนดังกล่าวจะได้รับข้อมูลขาเข้าจำนวนหนึ่งพร้อมกับผลลัพธ์ขาออกที่ถูกต้อง และให้ระบบเรียนรู้ด้วยการเปรียบเทียบผลลัพธ์ที่ตนประเมินออกมาได้ เปรียบเทียบกับ "เฉลย" เพื่อหาความแตกต่างและจุดที่ผิดพลาด จากนั้น ระบบจึงปรับแก้แบบจำลองการวิเคราะห์ของตนเองตามข้อมูลนั้น โอโยอาศัยเทคนิคหลากหลายประเภท ไม่ว่าจะเป็นการจำแนกแยกแยะข้อมูล การทำสมการถดถอย การคาดการณ์ หรือเทคนิค gradient boosting ซึ่งเป็นการที่กระบวนการ supervised learning นี้ ใช้รูปแบบต่าง ๆ ที่ตนพบเพื่อคาดการณ์ "ค่า" ของตัวแปรหรือ labels ต่าง ๆ เมื่อใดที่พบกับข้อมูลที่ยังไม่มีการเฉลยชุดต่อ ๆ ไปในอนาคต โดยกระบวนการ supervised learning นี้ เป็นที่นิยมใช้กันในสถานการณ์ที่ข้อมูลในอดีตสามารถเป็นเครื่องทำนายเหตุการณ์ในอนาคตได้อย่างมีประสิทธิภาพ เช่น ระบบที่ฝึกฝนอย่างดีแล้วจะสามารถ

คาดการณ์ล่วงหน้าได้ว่าธุรกรรมทางบัตรเครดิตใดบ้างที่มีแนวโน้มจะเป็นการทุจริต หรือคาดการณ์ได้ว่าผู้เอาประกันภัยรายใดมีโอกาสที่จะเบิกค่าประกันภัยมากกว่ารายอื่น ๆ เป็นต้น

Unsupervised learning ใช้กับข้อมูลที่ไม่มี label สำหรับข้อมูลในอดีตมาก่อน ระบบจะไม่ได้รับการบอก "เฉลย" และต้องทำการเรียนรู้ หาคำตอบด้วยตัวเอง เป้าหมายของการทำเช่นนี้ คือการสำรวจชุดข้อมูล และค้นหารูปแบบหรือโครงสร้างบางอย่างที่ซ่อนอยู่ภายใน การเรียนรู้แบบ unsupervised learning นั้น ทำงานได้ดีสำหรับการจัดการข้อมูลการทำธุรกรรม ตัวอย่างเช่น ระบบสามารถทำการระบุกลุ่มลูกค้าที่มีพฤติกรรมหรือคุณลักษณะคล้าย ๆ กัน ซึ่งจะนำไปสู่การตอบสนองที่คล้ายคลึงกันสำหรับลูกค้าที่มีลักษณะใกล้เคียงกัน เมื่อมีการทำโครงการด้านการตลาด เป็นต้น หรือระบบสามารถทำการค้นหาลักษณะสำคัญที่เป็นตัวแบ่งแยกลูกค้าแต่ละกลุ่มออกจากกันได้เช่นกัน เทคนิคที่เป็นที่นิยมใช้ในการเรียนรู้ประเภทนี้ เช่น การทำ self-organizing maps, nearest-neighbor mapping, k-means clustering และเทคนิค singular value decomposition เป็นต้น โดยอัลกอริทึมเหล่านี้ยังถูกนำมาใช้เพื่อแบ่งแยกหัวเรื่องของข้อความต่าง ๆ ทำการแนะนำรายการหรือสิ่งต่าง ๆ และตรวจหาข้อมูลที่ผิดปกติได้

Semisupervised learning นั้น ใช้ในสถานการณ์เดียวกันกับ supervised learning แต่ความแตกต่างคือ มีการใช้ข้อมูลทั้งแบบ labeled data และ unlabeled data เพื่อการฝึกฝนระบบ โดยทั่วไปแล้วใช้ข้อมูลแบบ labeled จำนวนน้อย ผสมกับข้อมูลแบบ unlabeled จำนวนมาก (ทั้งนี้เนื่องจากข้อมูลแบบ unlabeled มีต้นทุนถูกกว่าและได้มาด้วยความง่ายตายมากกว่า) การเรียนรู้ของเครื่องในรูปแบบดังกล่าวนี้ สามารถใช้ร่วมกับเทคนิคอื่น ๆ เช่น การจำแนกประเภทข้อมูล การใช้งานสมการถดถอย หรือการคาดการณ์และพยากรณ์ โดย semisupervised learning นั้น มีประโยชน์เมื่อต้นทุนของการทำ labeling ให้แก่ข้อมูลนั้นสูงเกินกว่าที่จะทำการฝึกฝนระบบด้วยกระบวนการแบบ labeled เต็มรูปแบบ ตัวอย่างแรก ๆ ของการใช้วิธีนี้ เช่น การระบุและจดจำใบหน้าของบุคคลด้วยกล้อง webcam

Reinforcement learning นั้น นำมาใช้กับด้านหุ่นยนต์ การเล่นเกมต่าง ๆ และการนำทางและการเคลื่อนที่ ด้วยการเรียนรู้ชนิดนี้ อัลกอริทึมจะลองผิดลองถูกและเรียนรู้ว่าเส้นทางการทำงานแบบใดที่ให้ผลตอบแทนที่ดีที่สุด การเรียนรู้ประเภทนี้ แบ่งองค์ประกอบออกเป็นสามส่วน คือ agent (ผู้เรียนรู้หรือตัดสินใจ นั่นคือระบบ), environment (ทุกสิ่งที่ agent มีปฏิสัมพันธ์ด้วย) และ actions (สิ่งที่ agent สามารถเลือกปฏิบัติได้) เป้าหมายของกระบวนการนี้ คือการให้ agent เลือกหนทางการปฏิบัติที่ให้รางวัลหรือผลตอบแทนสูงที่สุดในเวลาที่กำหนด โดย agent หรือตัวระบบนั้นสามารถบรรลุเป้าหมายได้รวดเร็วขึ้นมากหากมีแนวทางการทำงานที่มีประสิทธิภาพ

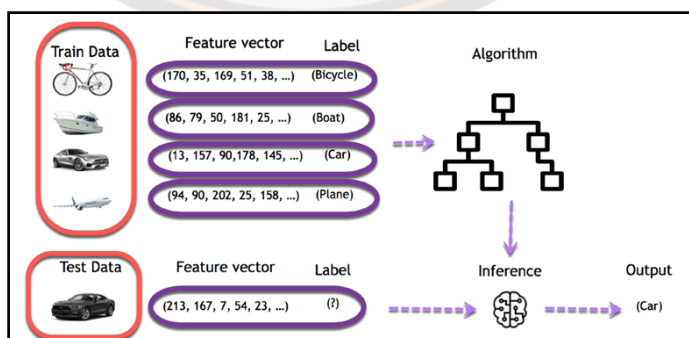
กระบวนการของ Machine Learning

กระบวนการของ Machine Learning สมมติว่าต้องการสร้างโปรแกรมที่มีการจดจำวัตถุขึ้นมา เพื่อที่จะฝึก (train) model ต้องใช้ตัวแบ่งประเภท (classifier) classifier ใช้คุณลักษณะ (feature) ของวัตถุเพื่อพยายามหาประเภทของวัตถุนั้น ในตัวอย่าง classifier ถูก train เพื่อตรวจจับ (detect) ถ้ารูปเป็น จักรยาน เรือ รถยนต์ เครื่องบิน 4 วัตถุนี้คือประเภทของวัตถุที่แตกต่างกัน classifier จำเป็นต้องจดจำ เพื่อที่สร้าง classifier ขึ้น จำเป็นต้องมีข้อมูลจำนวนหนึ่งเป็นข้อมูลขาเข้าและกำหนดให้เลยว่าข้อมูลแต่ละอันเป็นประเภทอะไร อัลกอริทึมนี้จะนำข้อมูลไปหารูปแบบ (pattern) แล้วแบ่งประเภทของข้อมูลตามประเภทต่าง ๆ การกระทำรูปแบบนี้เรียกว่า การเรียนรู้แบบมีผู้สอน (Supervised Learning) ใน Supervised Learning, ข้อมูลสำหรับการฝึก (training data) จะส่งต่อไปในอัลกอริทึมพร้อมกับทำสัญลักษณ์เพื่อบอกผลลัพธ์ไว้แล้ว (Nessessence, 2018)

การ train อัลกอริทึมจำเป็นต้องทำตามวิธีการดังนี้

- เก็บข้อมูล
- Train classifier
- ทำการทำนาย

ขั้นตอนแรกเป็นขั้นตอนที่จำเป็นมาก การเลือกข้อมูลที่ถูกต้องจะนำมาซึ่งอัลกอริทึมที่ประสบความสำเร็จหรือล้มเหลว ข้อมูลซึ่งเลือกมาเพื่อ train ถูกเรียกว่า คุณลักษณะ (feature) ในตัวอย่างของวัตถุ, feature คือพิกเซล (pixel) ของรูป แต่ละรูปในแถวของข้อมูลขณะที่แต่ละ pixel เป็นหลักแทน ถ้ารูปมีขนาด 28×28 ชุดข้อมูล มีขนาด 784 หลัก ในรูปภาพด้านล่าง แต่ละรูปถูกแปลงเป็นเวกเตอร์ของคุณลักษณะ (feature vector) การทำสัญลักษณ์ไว้ (label) เป็นการบอกคอมพิวเตอร์ว่าอะไรอยู่ในรูปภาพนั้น ดังภาพที่ 20



ภาพ 20 Training Data

ที่มา : <https://www.thaiprogrammer.org/>

วัตถุประสงค์หลักเป็นการใช้ training data เพื่อแบ่งประเภทของชนิดของวัตถุ ในขั้นตอนแรกประกอบไปด้วยการสร้าง feature เป็นหลัก แล้วต่อมา, ขั้นตอนที่ 2 เกี่ยวข้องกับการเลือกอัลกอริทึมเพื่อ train model นั้น เมื่อ train เสร็จแล้ว model ดังกล่าว จะทำนายว่ามีสิ่งใดอยู่ในรูปภาพบ้าง หลังจากนั้นง่ายมากที่ใช้ model นั้นไปทำนายรูปภาพอื่น ๆ ต่อไป สำหรับการนำรูปภาพใหม่ ๆ เข้าไปสู่ model นั้น machine นั้น ทำนายประเภทของวัตถุนั้นว่าอยู่ประเภทไหน ยกตัวอย่างเช่น มีรูปภาพใหม่ทั้งหมดอยู่โดยปราศจากการ label ไว้ จึงนำไปใช้กับ model ดังกล่าว สำหรับมนุษย์เป็นสิ่งที่ง่ายตายมากตอบว่ารูปภาพนั้นมีอะไรอยู่แต่ machine ใช้ความรู้ที่เพิ่งกล่าวมาทำได้มากที่สุดแค่ทำนายว่ามีอะไรอยู่ในรูปภาพนั้นเท่านั้นเอง

งานวิจัยการเรียนรู้ของเครื่อง (Machine Learning)

รัชกฤต อารีราษฎร์ (2561) การจำแนกข้อความเข้าข่ายหมิ่นประมาทบนสื่อสังคมออนไลน์ การสื่อสารผ่านสื่อสังคมออนไลน์ในปัจจุบันเป็นที่นิยมกันอย่างแพร่หลาย การแสดง ความคิดเห็นหรือแบ่งปันข้อมูลที่มีเนื้อหาก้าวร้าว โจมตี หรือดูหมิ่นผู้ใช้งานคนอื่นบนสื่อสังคม ออนไลน์ อาจส่งผลกระทบต่อสังคมในด้านลบ โดยเนื้อหาดังกล่าวอาจผิดกฎหมายอาญาหมวด 3 ความผิดฐานหมิ่นประมาท มาตรา 326 วิทยานิพนธ์นี้เสนอคุณลักษณะเพื่อใช้ในการจำแนกข้อมูล เข้าข่ายหมิ่นประมาทบนสื่อสังคมออนไลน์ด้วยขั้นตอนวิธีเพอเซ็ปตรอนหลายชั้น ซับพอร์ทเวกเตอร์ แมชชีน และการถดถอยโลจิสติกส์ โดยเปรียบเทียบประสิทธิภาพแต่ละขั้นตอนวิธี ซึ่งการทดลอง พบว่าเอ็น-แกรม คลังคำศัพท์จากศาสตร์ และโครงสร้างไวยากรณ์แบบขึ้นต่อกันเป็นคุณลักษณะที่สามารถใช้ในการจำแนกข้อความหมิ่นประมาทได้โดยได้ค่าความเที่ยงสูง แต่ค่าเรียกคืนต่ำ แต่เมื่อมีการจัดการข้อมูลที่ไม่สมดุลด้วยแล้ว พบว่าการจำแนกมีประสิทธิภาพดีขึ้นโดยที่ขั้นตอนวิธีเพอเซ็ปตรอนหลายชั้นมีความสามารถในการจำแนกได้ดีที่สุดโดยมีค่าความเที่ยงเป็น 0.93 ค่าเรียก คืนเป็น 0.98 และค่าเอฟวันเป็น 0.95 นอกจากนี้จำนวนมิติของเอ็น-แกรมมีผลต่อประสิทธิภาพ ของการจำแนกข้อความโดยจำนวนมิติที่เหมาะสมของเอ็น-แกรมแต่ละชนิดขึ้นอยู่กับขั้นตอนวิธีที่ใช้

Nutmos (2018) Google ใช้ machine learning ร่วมกับมนุษย์ กำจัดรีวิวขยะบน Play Store ได้นับล้านในสัปดาห์เดียว ระบบรีวิวบน Google Play นั้นมีผลต่อการตัดสินใจติดตั้งแอปพลิเคชันของผู้ใช้มาก แต่ทุกวันนี้มีปัญหารีวิวขยะเกิดขึ้นมากเนื่องจากนักพัฒนาบางรายใช้วิธีซื้อรีวิวเพื่อทำให้แอปพลิเคชันให้ดูดีขึ้น ซึ่ง Google รับทราบปัญหานี้ดี และพยายามที่จะหาวิธีกำจัดเนื่องจากเป็นปัญหาที่กระทบต่อความเชื่อมั่นของผู้ใช้ เรตติ้งและรีวิวที่ถูกจัดว่าผิดกฎของ Google Play มีลักษณะสามอย่างหลัก ๆ คือ คอนเทนต์ขยะ รีวิวประเภทหยาบคาย, ใช้คำแสดงความเกลียดชังหรือนอกเรื่อง เรตติ้งปลอมเป็นเรตติ้งและรีวิวที่มีจุดประสงค์ต้องการเปลี่ยนเรตติ้งโดยเฉลี่ยหรือที่ปรีวิวของแอปพลิเคชัน การปั่นเรตติ้งโดยใช้ที่เป็นมนุษย์จริง ๆ มีการปั่นเรตติ้งและรีวิวเพื่อแลกกับเงินหรือของบางอย่าง ทีมงาน Google Play Trust & Safety ได้พัฒนาระบบที่ใช้มนุษย์และ machine

learning ตรวจสอบและจัดการเรตติ้งและรีวิวขยะใน Play Store โดยมีทีมวิศวกรและนักวิเคราะห์คอยมอนิเตอร์เพื่อตรวจสอบเหตุการณ์ไม่ปกติในเรตติ้งและรีวิวอยู่เสมอ ๆ เพื่อปรับปรุงความแม่นยำของโมเดล รวมถึงสอบถามนักรีวิวที่มีทักษะด้วย เพื่อตรวจสอบการตัดสินใจของโมเดล Google ระบุว่า เป็นงานใหญ่มาก และจากการใช้เครื่องมือในสัปดาห์ที่ผ่านมา Google ได้ตรวจสอบพบรีวิวและเรตติ้งขยะนับล้าน โดยลบออกจาก Play Store เป็นที่เรียบร้อยแล้ว ส่วนแอปขยะที่ตรวจจับได้จากการมีพฤติกรรมรีวิวและเรตติ้งที่น่าสงสัยตรวจพบได้นับพันแอป สำหรับนักพัฒนา Google แนะนำให้ทำตามนโยบายของ Play Store คืออย่าซื้อเรตติ้ง หรือทำแคมเปญปั่นเรตติ้งอย่างเช่นกดห้าดาวแล้วจะให้ไอเทมในแอปพลิเคชัน ส่วนผู้ใช้ไม่ควรรับเงินหรือของขวัญเพื่อแลกเปลี่ยนกับการรีวิว รวมถึงการเขียนรีวิวนอกเรื่อง, เขียนไร้สาระ, ใช้คำแสดงความเกลียดชัง, หยาบคาย หรือเรื่องเพศที่ไม่ได้รับอนุญาตด้วยเช่นกัน สุดท้าย Google ระบุว่าถ้าพบเรตติ้งและรีวิวขยะ ขอความร่วมมือให้ผู้ใช้กดปุ่มบอกว่าคอมเมนต์เป็นสแปมด้วย เพื่อให้ Google ตรวจสอบและจัดการต่อไป

M.Pussapol (2020) Facebook ใช้ Machine Learning ตรวจสอบบัญชีปลอม ในปี 2019 โดยเฉลี่ยแล้ว Facebook ปิดบัญชีปลอม 2 พันล้านบัญชีต่อไตรมาส บรรดาผู้ฉ้อโกง หรือ Fraudster ใช้บัญชีปลอมเหล่านี้เพื่อกระจายสแปม ลิงค์ฟิชซิง หรือมัลแวร์ เป็นธุรกิจประเภททำนาบนหลังคนสามารถทำลายล้างผู้ใช้ได้เพียงไม่กี่คลิกที่หลงติดกับดักอย่างง่ายดาย เมื่อไม่นานมานี้ Facebook ได้เปิดเผยรายละเอียดเกี่ยวกับระบบ Machine Learning ที่ใช้เพื่อรับมือกับความท้าทายนี้ ยักษ์ใหญ่ด้านเทคโนโลยีแยกความแตกต่างระหว่างบัญชีปลอม 2 ประเภท อันดับแรกมี "บัญชีที่ผู้ใช้อาจจำแนกประเภทสับสน" โปรไฟล์ส่วนบุคคลสำหรับธุรกิจหรือสัตว์เลี้ยงที่ตีความเป็น 2 นับได้ว่าเป็น Page หรือไม่ สิ่งเหล่านี้จัดการได้ค่อนข้างตรงไปตรงมา คือ Convert ทั้งหมดเป็น Page อีกส่วนหนึ่งที่ดูซีเรียสกว่า คือ "การละเมิดบัญชี" เหล่านี้เป็นโปรไฟล์ส่วนบุคคลที่มีส่วนร่วมในการหลอกลวง สแปมหรือละเมิดข้อกำหนดการให้บริการของแพลตฟอร์ม บัญชีที่ละเมิดต้องถูกลบออกโดยเร็วที่สุด โดยไม่ขยายวงในการทำลายบัญชีจึงกว้างเกินไป ในการทำเช่นนี้ Facebook ใช้การถอดรหัสและ Machine Learning เพื่อบล็อกบัญชีปลอมทั้งก่อนที่จะถูกสร้างขึ้น หรือก่อนที่เปิดใช้งาน กล่าวอีกนัยหนึ่งคือ ก่อนที่จะเป็นอันตรายต่อผู้ใช้งานจริง ขั้นตอนสุดท้าย คือ หลังจากที่บัญชีปลอมได้เผยแพร่แล้ว สิ่งนี้เกิดขึ้นเมื่อการตรวจจับซับซ้อนมากขึ้น มีระบบ Machine Learning ใหม่ที่เรียกว่า Deep Entity Classification (DEC) เข้ามาล้วงข้อมูลลึกมากขึ้น DEC เรียนรู้ที่จะแยกแยะผู้ใช้ปลอมกับผู้ใช้จริงโดยใช้รูปแบบการเชื่อมต่อผ่านเครือข่าย สิ่งนี้เรียกว่า "Deep Features" ซึ่งรวมข้อมูลต่างๆ เช่น อายุเฉลี่ย หรือการกระจายเพศของเพื่อนของผู้ใช้ Facebook ใช้พีเจอร์ที่ลึกกว่า 20,000 รายการในการกำหนดลักษณะของแต่ละบัญชี โดยแสดงภาพรวมของการทำงานของแต่ละโปรไฟล์ เพื่อให้ผู้ต้องการโจมตีระบบแม้เปลี่ยนกลยุทธ์ไปเรื่อย ๆ ระบบเริ่มต้นด้วยการใช้ฉลากที่สร้างจากเครื่องจักรที่มีความแม่นยำต่ำจำนวนมาก สิ่งเหล่านี้สร้างขึ้นจากการผสมผสานของกฎและโมเดล Machine Learning อื่นๆ ที่ประเมินว่าผู้ใช้เป็นความจริง

หรือตัวปลอม เมื่อข้อมูลนั้นถูกนำมาใช้เพื่อสร้างปัญญาประดิษฐ์แล้ว แบบจำลองจะถูกปรับแต่งด้วยชุดข้อมูลขนาดเล็กที่มีความแม่นยำสูง ซึ่งสร้างขึ้นโดยผู้คนที่มีความเข้าใจในวัฒนธรรมท้องถิ่น ระบบการจำแนกขั้นสุดท้ายสามารถระบุ 1 ใน 4 ประเภทของโปรไฟล์ปลอม ประกอบด้วย บัญชีผิดกฎหมายไม่ได้เป็นตัวแทนของบุคคล บัญชีที่ถูกบงกชของผู้ใช้จริงที่ถูกโจมตีโดยผู้โจมตี ผู้ส่งอีเมลขยะที่ส่งข้อความสร้างรายได้ซ้ำๆ หลอกหลวงผู้ใช้งาน เปิดเผยข้อมูลส่วนบุคคล ระบบช่วยให้เหลือบัญชีปลอมบนแพลตฟอร์มประมาณ 5% ของ Active User รายเดือน ทั้งนี้ รายละเอียดของความพยายามในการสะสางภายในของ Facebook เกิดขึ้นท่ามกลางความกังวลเกี่ยวกับการจัดการเรื่องการเลือกตั้งประธานาธิบดีสหรัฐฯ โดยเฉพาะที่เกี่ยวกับข้อมูลปลอมเพื่อการเลือกตั้งบริสุทธิ์ยุติธรรม ทีม Facebook กล่าวว่า “การละเมิดทั่วไป ไม่ได้มีเป้าหมายเฉพาะเจาะจงเกี่ยวกับหัวข้อการเลือกตั้ง ” แต่ DEC (Deep Entity Classification) จะทำให้สมบูรณ์ขึ้น โดยเสริมศักยภาพอื่นๆ ของแพลตฟอร์มที่ลดการโฆษณาชวนเชื่อของการเลือกตั้ง

Aviv Ovadya ผู้ก่อตั้ง Thoughtful Technology Project องค์กรไม่แสวงหาผลกำไร เป็นผู้ศึกษาการออกแบบและกำกับดูแลแพลตฟอร์มด้านธรรมาภิบาล กล่าวว่า ความพยายามของ Facebook ในการเพิ่มความโปร่งใสให้กับกระบวนการทำงานนั้นน่ายกย่อง “มีประโยชน์ และมีพลังมากหากพูดคุยกันอย่างพิถีพิถันเกี่ยวกับการตัดสินใจเชิงสถาปัตยกรรม และวิธีการทำงานของระบบรักษาความปลอดภัยที่บริษัทอื่นๆ สามารถเลียนแบบได้” “เนื่องจากบริษัทอย่าง Facebook มีทรัพยากรในการลงทุนมากกว่าบริษัทขนาดเล็กอย่างมีนัยสำคัญ จึงเป็นประโยชน์อย่างยิ่งที่จะทำเรื่อง Knowledge Sharing” แต่ความพยายามบนเส้นทางเพิ่มความโปร่งใสให้กับกระบวนการทำงานยังอีกยาวไกล ด้วยผู้ใช้งานรายเดือน 2.5 พันล้านราย 5% เป็นบัญชีปลอม ซึ่งเท่ากับ 125 ล้านบัญชี Machine Learning จะก้าวหน้าไปได้เร็วแค่ไหน ไม่ว่าจะได้รับข้อมูลมากเท่าไร ย่อมไม่สามารถจับผิดทุกบัญชีปลอมด้วยความแม่นยำสมบูรณ์แบบท้ายที่สุดแล้ว แพลตฟอร์มคงต้องหันไปใช้การผสมผสานระหว่างมนุษย์กับเครื่องจักรเพื่อแก้ไขปัญหานี้

FEED NEWS (2021) นักวิจัยพบการเขียนรีวิวปลอมบน AMAZON เพื่อแลกเงินและสินค้า มีบันทึกกว่า 13 ล้านรายการ Safety Detectives กลุ่มนักวิจัยและผู้เชี่ยวชาญ cybersecurity ออกรายงานเปิดเผยฐานข้อมูล Elasticsearch แบบเปิด พบว่ามีการใช้รีวิวปลอมใน Amazon อันผิดจรรยาบรรณการค้าขาย พบฐานข้อมูลการรีวิวระหว่างผู้ขายของ Amazon และลูกค้าที่เต็มใจเขียนรีวิวปลอมเพื่อแลกกับสินค้าฟรี โดยรวมแล้วมีบันทึกไว้ 13,124,962 รายการ (ข้อมูล 7 GB) กระทั่งต่อข้อมูลมากกว่า 200,000 คน เมื่อผู้รีวิวเขียนรีวิวให้ 5 ดาวแล้ว จะได้เงินคืนค่าสินค้าผ่าน PayPal โดยข้อมูลที่ถูกเปิดเผยมาด้วยยังมีอีเมลที่ผูกกับ PayPal, ลิงค์บัญชี Amazon, ชื่อ username หนึ่งในแบนด์ที่กระตุ้นให้คนมาเขียนรีวิวปลอมคือ Mpow แบนด์ขายอุปกรณ์ USB charger จากจีนดำเนินการโดย ByteDance และ Patozon บริษัทผลิตภัณฑ์สำหรับลูกค้าที่ได้รับ

การสนับสนุนจาก Xiaomi นอกจากนี้ยังมีแบรนด์ Aukey ผู้ทำ accessories และอุปกรณ์ไฟฟ้า จากจีนเช่นกัน ซึ่งหน้าร้านทั้งสองถูกลบออกจาก Amazon ไปแล้ว ตัวแทนของ Aukey ระบุว่าไม่รู้รายละเอียดของสิ่งที่เกิดขึ้นเช่นกันขณะนี้กำลังหาทางแก้ไข บริษัทยังไม่ได้รับผลกระทบอย่างมีนัยสำคัญในขณะนี้การทำรีวิวปลอมบน Amazon เป็นสิ่งที่กำจัดออกไม่หมด ตอนนี้มีเว็บไซต์ขายรีวิวสินค้าที่ขายบน Amazon แบบปลอมๆ รวม 10 แห่ง ซึ่งการค้นพบของ SafetyDetectives ขยายผลให้เห็นภาพเพิ่มเติมว่าปริมาณการรีวิวปลอมนั้นมากมายขนาดไหน

2.6 การวิเคราะห์เหมืองข้อมูล (Data Mining)

การวิเคราะห์เหมืองข้อมูล (Nessessence, 2018) คือ การทำเหมืองข้อมูลเป็นกระบวนการในการค้นหาความผิดปกติ รูปแบบ และความสัมพันธ์ภายในชุดข้อมูลขนาดใหญ่เพื่อคาดการณ์ผลลัพธ์ด้วยการใช้เทคนิคที่หลากหลาย สามารถใช้ข้อมูลนี้เพื่อเพิ่มรายได้ ลดต้นทุน พัฒนาความสัมพันธ์กับลูกค้า ลดความเสี่ยง และอื่นๆ อีกมากมาย กระบวนการขุดผ่านข้อมูลเพื่อค้นหาความเชื่อมโยงที่ซ่อนอยู่และคาดการณ์แนวโน้มที่เกิดขึ้นในอนาคตมีประวัติศาสตร์อันยาวนาน บางครั้งเรียกว่า "การค้นพบองค์ความรู้ในฐานข้อมูล" ส่วนคำศัพท์ "การทำเหมืองข้อมูล" นั้นยังไม่ได้มีการใช้งานจนถึงช่วงทศวรรษที่ 1990 ทั้งนี้ รากฐานของสิ่งนี้ประกอบขึ้นจากระเบียบวิธีทางวิทยาศาสตร์ 3 อย่างที่เกี่ยวข้องกัน ได้แก่ สถิติ (การศึกษาเชิงตัวเลขของความสัมพันธ์ของข้อมูล) ปัญญาประดิษฐ์ (ปัญญาเสมือนมนุษย์ที่เกิดขึ้นจากซอฟต์แวร์และ/หรือเครื่องจักร) และการเรียนรู้ของเครื่อง (อัลกอริทึมที่สามารถเรียนรู้จากข้อมูลเพื่อสร้างการคาดการณ์) ทว่าไม่ใช่สิ่งใหม่แต่อย่างใด เนื่องจากเทคโนโลยีการทำข้อมูลมีการพัฒนาอย่างต่อเนื่องเพื่อให้ทันกับศักยภาพที่ไร้ขีดจำกัดของ Big Data และพลังการประมวลผลที่มีต้นทุนต่ำลง

ในปัจจุบันความก้าวหน้าด้านกำลังและความเร็วในการประมวลผลทำให้สามารถก้าวข้ามการปฏิบัติในแบบแมนนวลที่น่าเบื่อและใช้เวลานานไปสู่การวิเคราะห์ข้อมูลที่รวดเร็ว ง่ายตาย และเป็นแบบอัตโนมัติ ยิ่งชุดข้อมูลที่รวบรวมมีความซับซ้อนมากเท่าใดยิ่งมีโอกาสค้นพบข้อมูลเชิงลึกที่เกี่ยวข้องมากขึ้นเท่านั้น ร้านค้าปลีก ธนาคาร ผู้ผลิต ผู้ให้บริการ โทรคมนาคมและ บริษัทประกัน และอื่นๆ มีการใช้การวิเคราะห์เหมืองข้อมูลเพื่อค้นหาความสัมพันธ์ระหว่างทุกอย่างนับตั้งแต่ การเพิ่มประสิทธิภาพด้านราคา โปรโมชัน และข้อมูลประชากรไปจนถึงเศรษฐกิจ ความเสี่ยง การแข่งขัน และสื่อสังคมมีผลกระทบต่อรูปแบบการทำธุรกิจ รายได้ การดำเนินงาน และความสัมพันธ์กับลูกค้าของของพวกอย่างไร

ความสำคัญของการทำเหมืองข้อมูล

เนื่องจากจำนวนที่ผันแปร ปริมาณของข้อมูลที่ทวีคูณขึ้นทุกสองปี โดยข้อมูลที่ไม่มีโครงสร้างเพียงอย่างเดียวเป็นส่วนประกอบถึง 90% ของจักรวาลดิจิทัล แล้วแต่ปริมาณข้อมูลที่มาขายนี้นี้ไม่ได้แปลว่ามีองค์ความรู้มากขึ้น การทำเหมืองข้อมูลช่วยให้กลั่นกรองข้อมูลที่ยุ่งเหยิงและซ้ำซ้อนในข้อมูลทำความเข้าใจว่าสิ่งใดที่เกี่ยวข้องและใช้ประโยชน์จากข้อมูลนั้นเพื่อประเมินผลลัพธ์ที่เป็นไปได้ เร่งความเร็วให้กับการตัดสินใจที่ชาญฉลาด การทำเหมืองข้อมูลในยุคปัจจุบันเป็นเสมือนเสาหลักของการวิเคราะห์ ซึ่งช่วยให้พัฒนารูปแบบที่สามารถเผยแพร่เชื่อมต่อกภายในข้อมูลนับล้านๆ หรือพันล้านรายการ ซึ่งการทำเหมืองข้อมูลนำมาใช้ในธุรกิจอย่างหลากหลายสามารถทำให้การบริหารจัดการเป็นไปอย่างง่ายดาย และสามารถวางแผนวิเคราะห์การตลาดเพื่อให้ได้ทราบสิ่งที่เป็นประโยชน์ อาทิเช่น

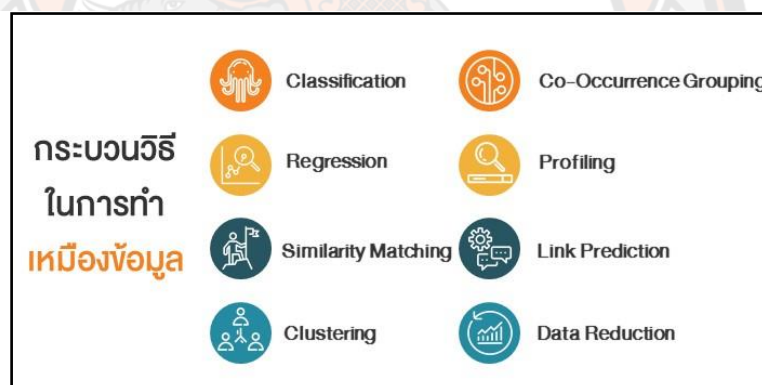
- ธุรกิจค้าปลีกสามารถใช้งาน Data Mining ในการพิจารณาหากกลยุทธ์ให้เป็นที่สนใจกับลูกค้าในรูปแบบต่าง ๆ เช่น ที่ว่างในชั้นวางของจะจัดการอย่างไรถึงจะเพิ่มยอดขายได้ เช่น ที่ Midas ซึ่งเป็นผู้แทนจำหน่ายอะไหล่สำหรับอุตสาหกรรมรถยนต์ งานที่ต้องทำคือการจัดการกับข้อมูลที่ได้รับจากสาขาทั้งหมด ซึ่งต้องทำการรวบรวมและวิเคราะห์อย่างทันท่วงที
- กิจการโทรคมนาคม เช่นที่ Bouygues Telecom ได้นำมาใช้ตรวจสอบการโกงโดยวิเคราะห์รูปแบบการใช้งานของสมาชิกลูกค้าในการใช้งานโทรศัพท์ เช่น คาบเวลาที่ใช้จ่ายหมายปลายทาง ความถี่ที่ใช้ ฯลฯ และคาดการณ์ข้อบกพร่องที่เป็นไปได้ในการชำระเงิน เทคนิคนี้ยังได้ถูกนำมาใช้กับลูกค้าโทรศัพท์เคลื่อนที่ซึ่งระบบสามารถตรวจสอบได้ว่าที่ใดที่เสี่ยงที่สูญเสยลูกค้าสูงในการแข่งขัน France Telecom ได้ค้นหาวิธีรวมกลุ่มผู้ใช้ให้เป็นหนึ่งเดียวด้วยการสร้างแรงดึงดูดในเรื่องค่าใช้จ่ายและพัฒนาเรื่องความจงรักภักดีต่อสินค้า
- การวิเคราะห์ผลิตภัณฑ์ เก็บรวบรวมลักษณะและราคาของผลิตภัณฑ์ทั้งหมดสร้างโมเดลด้วยเทคนิค Data Mining และใช้โมเดลในการทำนายราคาผลิตภัณฑ์ตัวอื่น ๆ
- การวิเคราะห์บัตรเครดิต ช่วยบริษัทเครดิตการ์ดตัดสินใจในการที่ให้เครดิตการ์ดกับลูกค้าหรือไม่ แบ่งประเภทของลูกค้าว่ามีความเสี่ยงในเรื่องเครดิต ต่ำ ปานกลางหรือสูงป้องกันปัญหาเรื่องการทุจริตบัตรเครดิต (SAS Institute Inc, 2020)
- การวิเคราะห์ลูกค้า ช่วยแบ่งกลุ่มและวิเคราะห์ลูกค้าเพื่อที่จะผลิตและเสนอสินค้าได้ตรงตามกลุ่มเป้าหมายแต่ละกลุ่ม ทำนายว่าลูกค้าคนใดจะเลิกใช้บริการจากบริษัทภายใน 6 เดือนหน้า
- การวิเคราะห์การขายพบว่า 70 % ของลูกค้าที่ซื้อโทรศัพท์แล้วซื้อวิดีโอตามมา ดังนั้นผู้จัดการจึงควรมุ่งไป ลูกค้าที่ซื้อโทรศัพท์แล้วจึงส่งเมลไปยังลูกค้าเหล่านั้นเพื่อที่เชิญชวน หรือให้ข้อเสนอที่ดี เพื่อให้ลูกค้ามาซื้อวิดีโอในครั้งต่อไป ช่วยในการโฆษณาสินค้าได้อย่างเหมาะสมและตรงตามเป้าหมาย ช่วยในการจัดวางสินค้าได้อย่างเหมาะสม

- Text Mining เป็นการปรับใช้ Data Mining มาอยู่ในรูปของข้อมูลตัวอักษรซึ่งเป็นรูปแบบของภาษาเครื่อง SDP Infoware ตัวอย่างของงานคือใช้เป็นเครื่องมือตรวจระดับความพึงพอใจของผู้ที่เข้าชมนิทรรศการโดยผ่านการประมวลผลจากแบบสอบถาม
- E-Commerce ช่วยให้เข้าใจพฤติกรรมของลูกค้า เช่น ลูกค้าเข้าไปที่ web ใดตามลำดับก่อนหลัง ช่วยในการปรับปรุง web site เช่น พิจารณาว่าส่วนใดของ web ที่ควรปรับปรุงหรือควรเรียงลำดับการเชื่อมโยงในแต่ละหน้าอย่างไรเพื่อให้สะดวกกับผู้ใช้เยี่ยมชม

เทคนิคเหมืองข้อมูล (Data mining)

เทคนิคเหมืองข้อมูล (ภานุวัฒน์ เมฆะ, 2564) เป็นกระบวนการ (Process) ที่กระทำกับข้อมูลขนาดใหญ่เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นโดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์เพื่อให้ได้สารสนเทศที่ไม่รู้ออกมา โดยสารสนเทศที่ได้มีเหตุผลและสามารถนำไปใช้ประโยชน์ได้ ซึ่งมีเทคนิคต่าง ๆ ดังภาพที่ 21

กระบวนการวิธีในการทำเหมืองข้อมูล (Data Mining)



ภาพ 21 กระบวนการวิธีในการทำเหมืองข้อมูล

ที่มา : <https://www.g-able.com/digital-review/>

1. Classification คือ การทำนาย (Prediction) ว่าข้อมูลแต่ละตัว (Individual) ในประชากร (Population) ควรจัดให้อยู่ (Classify) ในกลุ่มใดบ้าง โดยแต่ละกลุ่มมีการกำหนดชื่อไว้ล่วงหน้า (Pre-Defined) อาทิ หากมีข้อมูลของดอกกุหลาบ เช่น สี ขนาดกลีบดอก จำนวนกลีบสามารถทำนายได้ว่ากุหลาบนั้นน่าจะเป็นสายพันธุ์ใด

2. Regression หมายถึง การประมาณ (Value Estimation) ว่าข้อมูลแต่ละตัวควรมีค่าเชิงตัวเลข (Numerical Value) เป็นเท่าใด เช่น จากยอดขายของร้านค้าแห่งหนึ่งในระยะเวลา 5 ปีที่ผ่านมา มีอัตราการเติบโตอยู่ที่ 15 % ต่อปี ดังนั้นอาจทำนายการเติบโตของยอดขายในปีถัดไปว่าน่าจะเป็น 15% ด้วย

3. Similarity Matching หมายถึง การหาอัตลักษณ์ที่เหมือนกัน (Similar Identifying) บนพื้นฐานมิติต่างๆ ระหว่างข้อมูลแต่ละตัวในประชากรทั้งหมด เช่น Amazon, Lazada เมื่อลูกค้าสั่งซื้อสินค้า A ระบบจะแนะนำสินค้า B ให้ด้วย เพราะลูกค้าคนอื่นส่วนใหญ่ที่ซื้อสินค้า A นิยมซื้อสินค้า B ด้วย ซึ่งนิยมเรียกว่า ระบบการแนะนำสินค้า

4. Clustering การค้นหาการเกาะกลุ่มหรือกระจุกตัวของข้อมูล ซึ่งแตกต่างจาก Classification ในลักษณะที่ Clustering ไม่มีการกำหนดจำนวนของกลุ่มล่วงหน้า จำนวนกลุ่มที่ได้เป็นผลลัพธ์ที่ได้จากการประมวลผลทั้งสิ้น เช่น การจำแนกกลุ่มลูกค้า (Customer Segmentation) จากฐานลูกค้าที่มีอยู่

5. Co-Occurrence Grouping (Association Rule Discovery / Market-Basket Analysis) การค้นหาความสัมพันธ์ระหว่างสินค้าตั้งแต่สองชนิดขึ้นไปจากบันทึกการซื้อขาย (Transaction Records) คือสินค้าใดบ้างที่ลูกค้านิยมซื้อพร้อมกัน และทำไมถึงเป็นอย่างนั้น เช่น ร้านค้าพบว่าสินค้า A และ B นิยมซื้อพร้อมกัน ทำการส่งเสริมการขาย โดยการลดสินค้า A หรือ B อย่างหนึ่งอย่างใดหรือทั้งสองอย่างพร้อมกันเพื่อกระตุ้นยอดขาย ข้อแตกต่างกับ Similarity Matching คือเป็นระบบ e-Commerce ที่มี ข้อมูลลูกค้าและประวัติการสั่งซื้อ ส่วน Co-Occurrence ใช้ในธุรกิจค้าปลีกที่การเก็บประวัติการซื้อสินค้าที่เป็นไปได้ยาก

6. Profiling (Anomaly Detection) เป็น การวิเคราะห์ คุณลักษณะ (Characteristics) ที่เกี่ยวข้องกับพฤติกรรม (Behavior) ในการทำกิจกรรมอย่างใดอย่างหนึ่ง อาทิ ระบบตรวจจับการใช้บัตรเครดิตปลอม (Fraud Detection) เช่น หากเกิดการใช้จ่ายบัตรเครดิตในสถานที่ที่ห่างไกลจากพื้นที่ที่มีการใช้จ่ายเป็นประจำ (Frequent Location) หรือการซื้อสินค้าบางชนิดที่มีราคาสูงผิดปกติจากธรรมดา เจ้าหน้าที่จะโทรศัพท์สอบถามไปที่เจ้าของบัตรเพื่อสอบถามเกี่ยวกับการซื้อสินค้านั้นว่าเกิดขึ้นจริงหรือไม่

7. Link Prediction คือ การทำนายความสัมพันธ์เชื่อมโยง (Link) ระหว่างข้อมูลแต่ละตัวว่าควรมีความสัมพันธ์กันหรือไม่ และความแข็งแรง (Strength) ของความสัมพันธ์ควรเป็นเท่าไร เช่น Facebook มีการแนะนำคนที่น่าจะรู้จัก (May Know) เมื่อมีเพื่อนคนเดียวกันหลายคนให้สามารถขอเป็นเพื่อนด้วย

8. Data Reduction เป็นการลดขนาดของข้อมูลที่มีขนาดใหญ่มาก (Large Data) ให้มีขนาดเล็กลง (Smaller Set) แต่คงสาระสำคัญของข้อมูลขนาดใหญ่ไว้ เช่น การกำจัดข้อมูลที่ไม่ครบถ้วน สูญหาย ออกไปจากฐานข้อมูลก่อนนำไปประมวลผลนอกจากนี้ยังมีองค์ประกอบอื่นๆ ที่เกี่ยวข้องกับวิทยาศาสตร์ข้อมูล (Data Science)

ส่วนประกอบที่สำคัญของระบบการทำเหมืองข้อมูล

1. Database & Data Warehouse เป็นแหล่งที่เก็บข้อมูลอื่น เช่น แฟ้มข้อมูล สเปรดชีท ซึ่งเป็นแหล่งที่เก็บข้อมูลดิบสำหรับการทำเหมืองข้อมูล
2. Database/Data Warehouse server ทำหน้าที่นำเข้าข้อมูลตามคำขอของผู้ใช้
3. Knowledge base ได้แก่ ความรู้ในงานที่ทำจะเป็นประโยชน์ต่อการชี้แนะทางการสืบค้นหรือการประเมินความน่าสนใจของรูปแบบผลลัพธ์ที่ได้ โดยเป็นความรู้เฉพาะด้านในงานที่ทำ เช่น ความรู้เฉพาะทางการแพทย์ หรือดาราศาสตร์ เป็นต้น
4. Data Mining Engine เป็นส่วนประกอบหลัก ประกอบด้วยโมดูลซึ่งรับผิดชอบงานทำเหมืองประเภทต่างๆ ได้แก่ การหาความสัมพันธ์ การจำแนกประเภท การจัดกลุ่ม เป็นต้น
5. Pattern Evaluation Module ทำงานร่วมกับ Data Mining Engine โดยใช้ค่าขีดแบ่งมาตรฐานวัดความน่าสนใจในการกลั่นกรองรูปแบบผลลัพธ์ที่ได้ ซึ่งใช้ประเมินความน่าสนใจของรูปแบบการทำเหมืองที่ได้
6. Graphical User Interface ส่วนติดต่อประสานระหว่างผู้ใช้กับระบบทำเหมือง ซึ่งช่วยให้ผู้ใช้สามารถระบุงานทำเหมืองที่ต้องการ

กลยุทธ์ในการทำเหมืองข้อมูล

1. แบบจำลองในการทำนาย (Predictive/ Supervised Modeling) เป็นผลลัพธ์ที่สร้างจากการอนุมาน (Inference) ชุดข้อมูลปัจจุบัน เพื่อใช้ในการทำนายประเภทตัวอย่างในอนาคต
2. แบบจำลองในการบรรยาย (Descriptive/ Unsupervised Modeling) ในที่นี้ อาจเป็นการหาความสัมพันธ์ต่างๆ (Association) หรือทำการจัดกลุ่มข้อมูล (Clustering) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย

กระบวนการชุดข้อมูลทำงานอย่างไร

แม้ว่าการชุดข้อมูลแตกต่างกันไปตามขนาดของข้อมูลที่เข้าถึงและกระบวนการที่จำเป็นสำหรับสิ่งนี้ โดยทั่วไปแล้วเกิดขึ้น ดังนี้

- ขั้นแรก ตรวจสอบพบสแต็กข้อมูลและมั่นใจในความปลอดภัยของสแต็กนี้
- ข้อมูลที่ไร้ประโยชน์และไร้ความหมายจะถูกล้าง
- ข้อมูลที่เหลือจะถูกรวมและแปลง
- นักชุดข้อมูลจัดกลุ่มข้อมูลด้วยวิธีต่างๆ เช่น การจัดกลุ่ม โครงสร้างสนับสนุนการตัดสินใจ การจัดประเภท ฯลฯ ซึ่งเหมาะสำหรับข้อมูลในมือ
- ผลลัพธ์ที่ได้รับจะถูกทดสอบและประเมินผล

องค์ประกอบการทำเหมืองข้อมูล (Data Mining)

สำหรับขั้นตอนในการทำเหมืองข้อมูลหรือการทำ Data Mining มีองค์ประกอบที่เกี่ยวข้องทั้งหมดดังต่อไปนี้ (G-Able, 2022) ดังภาพที่ 22



ภาพ 22 องค์ประกอบการทำเหมืองข้อมูล

ที่มา : <https://slideplayer.in.th/>

1. การทำความเข้าใจเกี่ยวกับธุรกิจ (Business Understanding) เป็นขั้นตอนแรกที่สำคัญมาก เพราะต้องทำความเข้าใจว่าปัญหาคืออะไร ต้องการคำตอบของปัญหาในทิศทางหรือลักษณะใด หากไม่เข้าใจปัญหาอย่างถ่องแท้ทำให้ขั้นตอนต่อไปดำเนินไปในทิศทางที่ไม่ถูกต้อง ซึ่งนอกจากไม่ได้คำตอบที่ต้องการแล้ว ยังสูญเสียเวลาและทรัพยากรไปโดยไร้ประโยชน์ด้วย

2. การทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding) เป็นการทำความเข้าใจว่าข้อมูลที่น่ามาใช้ควรมีลักษณะอย่างไร แหล่งข้อมูลอยู่ที่ใดและที่สำคัญที่สุดคือ ค่าใช้จ่ายหรือต้นทุน (Costs of Data) ที่ได้มาซึ่งข้อมูลเหล่านั้นมีค่าใช้จ่ายเท่าไร รวมทั้งควรต้องประเมินมูลค่าของประโยชน์ที่ได้รับจากการนำเอาข้อมูลดังกล่าวมาใช้

3. การเตรียมข้อมูล (Data Preparation) โดยปกติระบบประมวลผลข้อมูล นำเข้าข้อมูล อยู่ในรูปแบบที่จำกัด (Fixed Known Format) แต่ในความเป็นจริงข้อมูลส่วนใหญ่ไม่ได้จัดเก็บในลักษณะดังกล่าวจึงต้องการกระบวนการแปลงข้อมูล (Data Transformation) หรือเปลี่ยนชนิดข้อมูล (Data Conversion) เพื่อให้ข้อมูลอยู่ในลักษณะหรือรูปแบบที่ง่ายต่อการนำไปประมวลผลหรือวิเคราะห์ต่อไป

4. การสร้างแบบจำลอง (Modeling) คือ การสร้างรูปแบบความสัมพันธ์ (Relational Pattern) อยู่ในรูปของแบบจำลองบนซอฟต์แวร์ (Computer Model) หรือสมการความสัมพันธ์ (Equation) ก็ได้

5. การประเมินผล (Evaluation) หลังจากที่ได้แบบจำลองแล้ว ต้องทำการประเมินผลว่าแบบจำลองนั้นมีความถูกต้องแม่นยำมากน้อยเพียงใด โดยอาจทดลองในระบบเสมือน (Simulation) นำไปประมวลผลกับข้อมูลจริงที่มีอยู่เพื่อเปรียบเทียบผล ของการวิเคราะห์ว่าถูกต้องร้อยละเท่าใด

6. การนำไปใช้งาน (Deployment) หลังจากที่ได้แบบจำลองที่มีคุณภาพและความถูกต้องแม่นยำตามที่ต้องการสามารถนำไปใช้งานจริง โดยอาจต้องมีการปรับแต่งเพื่อความเหมาะสมในสภาวะจริง อีกทั้งยังต้องติดตั้งร่วมกับระบบอื่นๆ เช่น ระบบช่วยการตัดสินใจ (Decision Support System) หลังจากการติดตั้งแล้วควรมีการปรับปรุงแบบจำลองเป็นระยะๆ (Periodic Update) เพราะแท้จริงแล้วการทำเหมืองข้อมูลไม่มีที่สิ้นสุด

ลักษณะเฉพาะของข้อมูลที่มีการใช้ทำ Data Mining

- ข้อมูลขนาดใหญ่ เกินกว่าจะพิจารณาความสัมพันธ์ที่ซ่อนอยู่ภายในข้อมูลได้ด้วยตาเปล่า หรือโดยการใช้ Database Management System (DBMS) ในการจัดการฐานข้อมูล
- ข้อมูลที่มาจากหลายแหล่ง โดยอาจรวบรวมมาจากหลายระบบปฏิบัติการหรือหลาย DBMS เช่น Oracle , DB2 , MS SQL , MS Access เป็นต้น
- ข้อมูลที่ไม่มีการเปลี่ยนแปลงตลอดเวลาที่ทำการ Mining หากข้อมูลที่มีอยู่นั้นเป็นข้อมูลที่เปลี่ยนแปลงตลอดเวลาต้องแก้ปัญหานี้ก่อน โดยบันทึกฐานข้อมูลนั้นไว้และนำฐานข้อมูลที่บันทึกไว้มาทำ Mining แต่เนื่องจากข้อมูลนั้นมีการเปลี่ยนแปลงอยู่ตลอดเวลา จึงทำให้ผลลัพธ์ที่ได้จากการทำ Mining สมเหตุสมผลในช่วงเวลาหนึ่งเท่านั้น ดังนั้นเพื่อให้ได้ผลลัพธ์ที่มีความถูกต้องเหมาะสมอยู่ตลอดเวลาจึงต้องทำ Mining ใหม่ทุกครั้งในช่วงเวลาที่เหมาะสม
- ข้อมูลที่มีโครงสร้างซับซ้อน เช่น ข้อมูลรูปภาพ ข้อมูลมัลติมีเดีย ข้อมูลเหล่านี้สามารถนำมาทำ Mining ได้เช่นกันแต่ต้องใช้เทคนิคการทำ Data Mining ขั้นสูง

คุณลักษณะของ Data Mining

- การค้นหาข้อมูลโดยอาศัยเทคโนโลยีการทำเหมืองข้อมูล ภายใตฐานข้อมูลขนาดใหญ่ หรือคลังข้อมูล ซึ่งข้อมูลอาจถูกสะสมมานานหลายปี
- ผู้ใช้งานระบบสารสนเทศไม่จำเป็นต้องมีทักษะในการเขียนโปรแกรม เนื่องจากมีเครื่องมือช่วยค้นหาข้อมูลจากคลังข้อมูลได้อย่างรวดเร็ว

- ผู้ใช้ต้องกำหนดขอบเขตการค้นหาข้อมูลให้ชัดเจนเพื่อความรวดเร็วอาจมีการประมวลผลข้อมูลแบบขนาน เพื่อเพิ่มประสิทธิภาพในการค้นหาและวิเคราะห์ข้อมูล
- เครื่องมือสำหรับทำเหมืองข้อมูลสามารถใช้งานร่วมกับโปรแกรม Spreadsheet และเครื่องมือพัฒนาต่าง ๆ ได้เป็นอย่างดี

งานวิจัยการวิเคราะห์เหมืองข้อมูล (Data Mining)

นุชนาฏ ปิ่นเมือง และจारी ทองคำ (2561) การจำแนกความคิดเห็นของคนไทยเกี่ยวกับสื่อออนไลน์โดยใช้การทำเหมืองข้อความ เหมืองข้อความเป็นการวิเคราะห์ข้อมูลตัวอักษรเพื่อสกัดข้อมูลที่เป็นประโยชน์จากแหล่งข้อมูล ปัจจุบันเทคนิคในการ จำแนกเหมืองข้อความมีหลายวิธี งานวิจัยนี้มีวัตถุประสงค์เพื่อค้นหาเทคนิคการจำแนก จาก 5 เทคนิคที่มีประสิทธิภาพ คือ เทคนิค Naïve Bayes เทคนิค Support Vector Machine (SVM) เทคนิค K-Nearest Neighbor เทคนิค ต้นไม้ตัดสินใจ (Decision tree) และเทคนิค C4.5 ในการเก็บรวบรวมข้อมูลความคิดเห็นต่อการใช้บริการพร้อมเพย์บนสื่อออนไลน์จำนวนทั้งหมด 1,570 ข้อความ ในกระบวนการคัดเลือกค่าบ่งชี้เพื่อใช้ในการแยกคุณลักษณะได้เลือกใช้คำวิเศษณ์เพื่อทำการแยกคุณลักษณะ เชิงบวกและเชิงลบ คณะผู้วิจัยได้ใช้หลักการ 10-fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูล ทดสอบ และวัดประสิทธิภาพการจำแนกของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่า ความระลึก (Recall) เมื่อทำการทดสอบและวัดประสิทธิภาพของโมเดลพบว่า เทคนิค Naïve Bayes ให้ผลดีที่สุดในการจำแนก ข้อความความคิดเห็น โดยให้ค่าความถูกต้อง 93.88% ค่าความแม่นยำ 94.02% และค่าความระลึก 93.54%

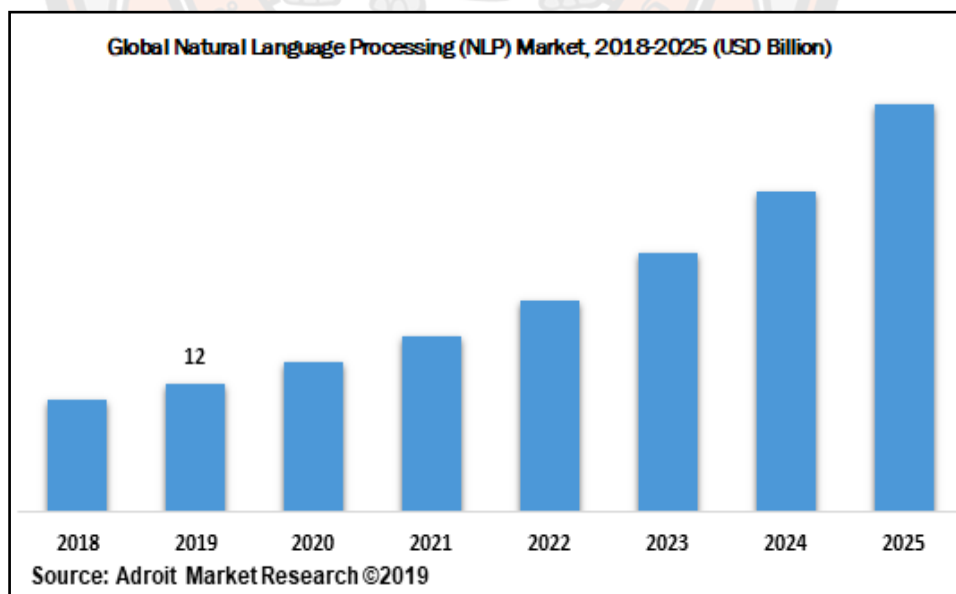
สมศักดิ์ ศรีสุวรรณ และสมัย ศรีสวย (2563) การวิเคราะห์เหมืองความคิดเห็นโดยใช้เทคนิคการสกัดคำ ในงานวิจัยนี้ได้นำข้อมูลจากบทวิจารณ์ออนไลน์ผ่านเครือข่ายเฟซบุ๊กสาธารณะของนักท่องเที่ยวมาสกัดคำแยกความคิดเห็นเชิงบวก เชิงลบ และได้ทำการเปรียบเทียบประสิทธิภาพจากค่าความถูกต้องด้วยอัลกอริทึมนาอิวเบย์ อัลกอริทึม การหาเพื่อนบ้านใกล้ที่สุด และอัลกอริทึมการเรียนรู้ของต้นไม้ตัดสินใจผลการศึกษา พบว่า อัลกอริทึมนาอิวเบย์ให้ค่าความถูกต้อง 87.97% อัลกอริทึมการหาเพื่อนบ้านใกล้ที่สุดให้ค่าความถูกต้อง 83.80% และ อัลกอริทึมการเรียนรู้ของต้นไม้ตัดสินใจให้ค่าความถูกต้อง 79.89% ผู้วิจัยได้เลือกวิธีนาอิวเบย์ มาใช้ในการพัฒนาระบบวิเคราะห์ความคิดเห็นจากเครือข่ายเฟซบุ๊กโรงแรมและที่พักในจังหวัดลำปางมาทดสอบระบบ ผลพบว่าระบบสามารถสกัดคำแยกความคิดเห็นเชิงลบและเชิงบวกตามเวลาจริงทำงานได้อย่างมีประสิทธิภาพ

จากการพิจารณาทฤษฎี บทความและงานวิจัยนั้น นักวิจัยได้ใช้เทคนิคการจำแนกเหมืองข้อความในกระบวนการคัดเลือกค่าบ่งชี้เพื่อใช้ในการแยกคุณลักษณะได้เลือกใช้คำวิเศษณ์เพื่อทำการแยกคุณลักษณะ เชิงบวกและเชิงลบ และการสกัดคำจากเหมืองความคิดเห็น พบว่า เทคนิค Naïve

Bayes มีการจำแนกข้อความคิดเห็นโดยให้ค่าความถูกต้อง ค่าความแม่นยำและค่าความระลึกที่ให้ผลดีที่สุด ทำให้ผู้วิจัยมีความสนใจในส่วนของ การใช้เทคนิค Naïve Bayes มาใช้ในการวิจัย สำหรับเทคนิค Support Vector Machine (SVM) เทคนิคต้นไม้ตัดสินใจ (Decision tree) เป็นเทคนิคที่เหมาะสมสำหรับการรวบรวมข้อมูลความคิดเห็นและนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลการระบุวิธีที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์

2.7 การประมวลผลภาษาธรรมชาติ (Natural Language Processing)

Natural Language Processing หรือ NLP เป็นหนึ่งในเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence หรือที่เรียกกันว่า AI) ซึ่งเป้าหมายของการใช้ NLP นั้นคือ การทำให้คอมพิวเตอร์เข้าใจภาษาของมนุษย์ เพื่อให้เกิดการสื่อสารข้อมูล รวมทั้งการวิเคราะห์ข้อมูลต่าง ๆ ออกมาได้อย่างที่มนุษย์สื่อสารกัน เช่น การใช้ระบบคอมพิวเตอร์ NLP แปลภาษา จากภาษาหนึ่งเป็นอีกภาษาหนึ่ง เป็นต้น เวลาที่คอมพิวเตอร์ประมวลผล โปรแกรมถูกออกแบบให้เข้าใจรหัสต่าง ๆ และตัวเลขที่มีความซับซ้อน ซึ่งไม่ใช่ภาษาที่เป็นประโยคแบบที่มนุษย์ใช้กัน ดังนั้น การใช้เทคโนโลยี AI อย่าง NLP เข้ามาเป็นตัวกลางในการประมวลผลข้อมูล และแสดงผลออกมาในรูปแบบการสื่อสารของมนุษย์มีส่วนช่วยให้การทำงานของผู้ใช้งานเป็นไปได้อย่างรวดเร็วขึ้น และสามารถนำไปใช้ประโยชน์ได้กับวงการอุตสาหกรรมต่าง ๆ (STEPS Academy, 2021) ดังภาพที่ 23



ภาพ 23 การเติบโตของตลาดที่ใช้งานเทคโนโลยี NLP ทั่วโลก

ที่มา : <https://stepstraining.co/foundation/natural-language-processing>

ประโยชน์ของการใช้ NLP ที่นำไปใช้ในชีวิตประจำวัน

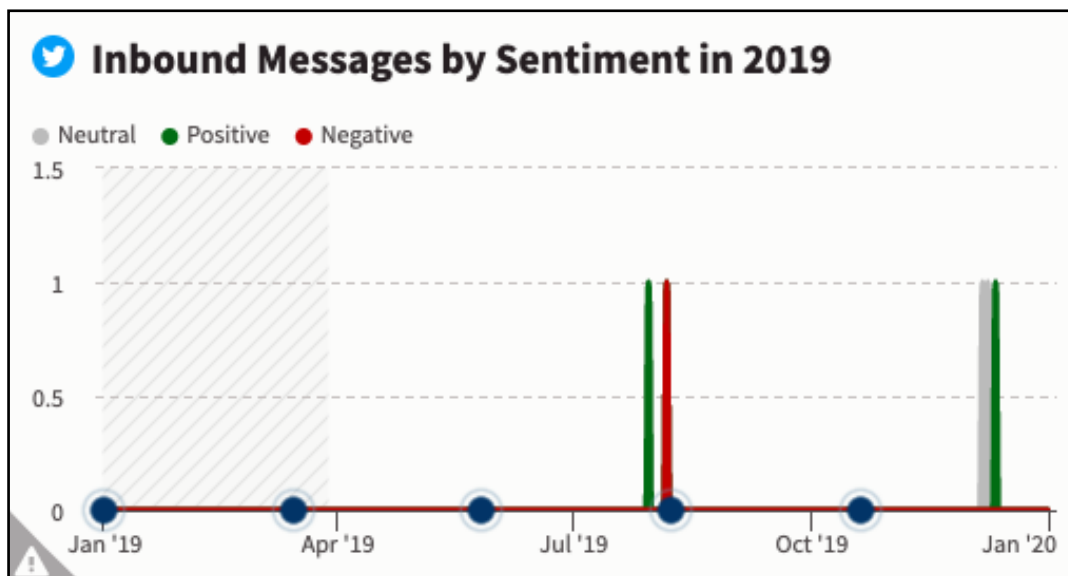
- เครื่องมือแปลภาษา อย่าง Google Translate ที่นอกจากการแปลภาษาให้ถูกต้องตามแต่ละประโยค หรือคำศัพท์ ยังมีเรื่องของการจัดเรียงคำให้สละสลวยเหมือนกับที่มนุษย์เขียน
- การปรับแก้ไวยากรณ์ หรือการประมวลผลคำที่ถูกต้องบนโปรแกรม Microsoft Word หรือ Google Doc
- การสร้างระบบอัลกอริทึมที่เข้าใจภาษามนุษย์ได้ BERT ของ Google ด้วยการจับ Keyword
- การค้นหาเอกสารบน Google Doc. ที่รวดเร็วและแม่นยำ
- การใช้ Voice Search ที่มีการสื่อสารระหว่างมนุษย์และ ระบบคอมพิวเตอร์อย่าง Siri โดยที่โปรแกรมจะมีการโต้ตอบกับผู้ใช้งานได้อย่างเป็นธรรมชาติ
- การสำรวจและช่วยมนุษย์วิเคราะห์ เพื่อนำไปใช้ในอุตสาหกรรมการผลิตและการก่อสร้าง

การทำงานของ NLP ที่มีผลต่อ Digital Marketing

ในมุมมองของธุรกิจและการทำการตลาดออนไลน์แล้ว NLP มีส่วนช่วยให้กลุ่มเป้าหมายเกิดความเข้าใจในสิ่งที่แบรนด์ต้องการสื่อสารออกไป ทั้งในรูปแบบของคอนเทนต์ การทำโฆษณา หรือการวิเคราะห์ข้อมูลจากการวัดผลทางการตลาด ตัวอย่างเช่น อะไรเป็นแรงจูงใจที่ทำให้กลุ่มเป้าหมายคลิกโฆษณา สาเหตุของการเกิดยอด Bounce Rate บนหน้า Landing Page คืออะไร สาเหตุที่ลูกค้าไม่ซื้อสินค้าหรือออกจากหน้าชำระสินค้าก่อนจ่ายเงินจริง เป็นเพราะอะไร นอกจากนี้ NLP มีส่วนช่วยนักการตลาดในการใช้คำค้นหาด้วยเสียง หรือ Voice Search เป็นอย่างมาก เนื่องจากระบบอัลกอริทึม การแปลงสารจากคอมพิวเตอร์ให้กลายเป็นภาษามนุษย์ ล้วนแล้วแต่ใช้ NLP ทั้งสิ้น

แนวทางการใช้ Natural Language Processing เพื่อการตลาดดิจิทัล มีดังนี้

- ทำให้แบรนด์เข้าใจลูกค้า ไม่ว่าแบรนด์ไม่ใช่องค์กรใหญ่ หรือเป็นบริษัท Start Up ที่เพิ่งเริ่มต้นทำธุรกิจ สามารถนำเทคโนโลยีมาใช้กับธุรกิจได้ เพื่อให้เกิดประสิทธิภาพสูงสุดในการทำการตลาด ซึ่งแบรนด์สามารถใช้ NLP ในการค้นหารีวิว หรือคำติชมจากลูกค้าที่ใช้สินค้า และบริการได้บนโลกออนไลน์ ระบบ NLP จะช่วยวิเคราะห์บทความต่าง ๆ ที่มีเนื้อหาเกี่ยวข้องกับแบรนด์ ทำให้ทราบว่าลูกค้าหรือคนบนโลกออนไลน์มีความคิดเห็นอย่างไรกับแบรนด์ เช่น จากกราฟ แสดงให้เห็นถึงข้อความจากลูกค้าทั้งในเชิงบวกและเชิงลบที่อยู่บนแพลตฟอร์ม Twitter สามารถใช้เครื่องมือในการเก็บข้อมูล เพื่อค้นหารีวิวต่าง ๆ ผ่านแพลตฟอร์มออนไลน์ได้ ดังภาพที่ 24



ภาพ 24 การวิเคราะห์ความรู้สึกเชิงบวกและเชิงลบ

ที่มา : <https://stepstraining.co/foundation/natural-language-processing>

- สร้างโปรแกรม Chatbots เพื่อบริการลูกค้าและสร้าง Lead Generation ในปัจจุบันผู้ประกอบการไม่จำเป็นต้องนั่งตอบคำถามลูกค้าแบบ 24 ชั่วโมงอีกต่อไป เพียงแค่ติดตั้งระบบ Chatbots เพื่อให้ลูกค้าได้เข้าไปสอบถามรายละเอียดเบื้องต้น และช่วยให้ลูกค้าไม่ต้องเสียเวลารอนาน นอกจากนี้การลงทุนกับระบบ Chatbot เพียงครั้งเดียวสามารถใช้ในการทำธุรกิจได้ระยะยาว ไม่ต้องเสียงบประมาณในการจ้างพนักงาน การใช้ Chatbot นั้นสามารถตอบสนองความต้องการของลูกค้าได้หลากหลายประการ ซึ่งตัวระบบได้มีการถูกคาดการณ์เอาไว้หลายประการในการใช้งานให้เกิดประโยชน์
- สามารถวิเคราะห์เทรนด์การใช้งานที่จะมาถึงในอนาคต คุณอาจเคยใช้เครื่องมืออื่น ๆ ในการรวบรวมข่าวสาร หรือ ฟีด RSS* เพื่อค้นหาข้อมูลทั่วไป ที่เกี่ยวข้องกับสินค้า และ บริการจากแบรนด์แบบเฉพาะเจาะจง ซึ่ง NLP มีวิธีการรวบรวมข้อมูลเหล่านั้นมาสรุปประเด็นหลัก ๆ ได้เช่นเดียวกัน แลผมยังสามารถทำได้แบบแยกส่วนเพื่อให้ง่ายต่อการจัดประเภทอีกด้วย (RSS* ย่อมาจาก Really Simple Syndication เป็นข้อมูลที่อยู่ในรูปแบบเอ็กซ์เอ็มแอล (XML) สามารถนำมาใช้เพื่อการกระจายข่าวทางเว็บไซต์ และ Blog ซึ่ง RSS ช่วยให้เราได้รับข่าวสารจากเว็บไซต์ต่าง ๆ ที่คอยให้บริการข้อมูลได้เร็วขึ้น)
- การสร้าง Copywriting ในอดีตใครจะรู้ว่าคอมพิวเตอร์สามารถสร้างคอนเทนต์ได้แบบเนียนเหมือนที่มนุษย์เขียนออกมาซึ่งดูเป็นไปได้ในยุคดิจิทัล เมื่อ NLP สามารถสร้าง Copywriting ออกมาเพื่อใช้เป็นคำอธิบายและแนะนำสินค้าต่าง ๆ บนหน้าเว็บไซต์ได้อีกด้วย ซึ่งแบรนด์ขนาดใหญ่

อย่าง Alibaba มีระบบ AI ในการเขียนคอนเทนต์อธิบายรายละเอียดผลิตภัณฑ์ต่าง ๆ บนเว็บไซต์ ทำให้มนุษย์ประหยัดเวลาในการเขียน Copywriting ที่หลายร้อยหน้าอีกด้วย

- ช่วยในการใช้คำค้นหาด้วยเสียง Voice Search ที่ใช้เทคโนโลยี NLP เข้ามาช่วยให้เกิดประโยชน์ ในปี 2020 ที่ผ่านมา ผู้บริโภคชาวอเมริกันมีลำโพงอัจฉริยะติดบ้านเอาไว้มากกว่า 75 % โดยผู้ใช้งานลำโพงอัจฉริยะได้ใช้ประโยชน์ เพื่อค้นหาสิ่งต่าง ๆ ผ่านการค้นหาด้วยเสียง เช่นการถามพยากรณ์อากาศ การสอบถามเส้นทางจราจร ซึ่งถือได้ว่า แปรนตร์ต่าง ๆ ที่ผลิตลำโพงอัจฉริยะออกมาประสบความสำเร็จเป็นอย่างมาก

NLP ใช้ machine learning algorithms กับข้อมูลประเภท text, speech หรือแม้แต่ภาพ ซึ่งสามารถประยุกต์ใช้ในงานทางด้าน NLP (Natratanon Kanraweekultana, 2019) เช่น

- speech recognition แปลงเสียงพูด (Audio File) เป็นข้อความตัวอักษร (Text) โดยสามารถแจกแจงคำพูดต่างๆ

- document summarization การสรุปความในภาษา ที่เป็นการลดความซับซ้อนและขนาดของเอกสารข้อความโดยไม่ทำให้ความหมายหรือสาระสำคัญของข้อมูลเอกสารสูญหายไป

- machine translation เครื่องมือที่ใช้สำหรับการแปลข้อความจำนวนมาก ๆ จากภาษาหนึ่งไปยังอีกภาษาหนึ่ง เช่น google translate

- spam detection

- named entity recognition การสกัดนิพจน์เฉพาะหรือชื่อเฉพาะในประโยค

- question answering การถามตอบ

- autocomplete เวลาพิมพ์ตัวอักษรไปในช่องรับข้อมูลแล้วจะแสดงข้อมูลที่เคยพิมพ์ขึ้นมา

- predictive typing การคาดเดาคำต่อไปเมื่อพิมพ์

การทำงานขั้นพื้นฐานของ NLP เช่น การทำ tokenization, parsing, lemmatization stemming, part-of-speech tagging, language detection และ identification of semantic relationships เป็นต้น ซึ่งกระบวนการทำงานย่อย ๆ เหล่านี้ใช้งานในกระบวนการทำงานด้าน NLP ที่ซับซ้อนยิ่งขึ้นไปอีกด้วย เช่น

- Content categorization เป็นการสรุปใจความหลักที่สำคัญของข้อความหรือชุดการสื่อสารต่าง ๆ ซึ่งรวมถึงการสืบค้น การทำความเข้าใจ การแจ้งเตือนเนื้อหาที่สำคัญ และการตรวจตราคำซ้ำ

- Topic discovery and modeling เป็นการทำงานเพื่อจับใจความสำคัญและภาพรวมที่อยู่ในชุดข้อความขนาดใหญ่ และนำเทคนิคการวิเคราะห์ระดับสูงหรือ advanced analytics มาใช้กับข้อความเหล่านั้น เช่น เทคนิค optimization and forecasting

- Contextual extraction เป็นกระบวนการสกัดข้อมูลที่มีโครงสร้าง ออกมาจากข้อมูลรูปแบบข้อความทั่วไป
- Sentiment analysis คือการตรวจหาความรู้สึก อารมณ์ หรือมุมมองความคิดเห็นของผู้คนหรือสาธารณชน ซึ่งอยู่ในข้อมูลข้อความหรือคำพูด ซึ่งรวมไปถึงความรู้สึกโดยเฉลี่ย และการทำเหมืองข้อมูลเพื่อสำรวจความคิดเห็นจากคนจำนวนมาก
- Speech-to-text and text-to-speech conversion คือการแปลงข้อความหรือคำสั่งรูปแบบเสียง ให้เป็นอักขระและตัวอักษร และ/หรือ แปลงอักขระให้เป็นข้อความเสียงพูด
- Document summarization คือการสร้างการสรุปสาระสำคัญของข้อความปริมาณมากโดยอัตโนมัติ
- Machine translation คือการแปลข้อความจากภาษาของมนุษย์ภาษาหนึ่ง ไปยังอีกภาษาหนึ่งโดยอัตโนมัติ

งานวิจัยการประมวลผลภาษาธรรมชาติ (Natural Language Processing)

นางเยาว สอนจะโปะ (2561) แหล่งข้อมูลขนาดใหญ่ที่อยู่บนอินเทอร์เน็ตล้วนเป็นข้อมูลที่มีประโยชน์ต่อการนำไปใช้งานทางด้านธุรกิจ การได้มาซึ่งข้อมูลสามารถนำมาใช้ประโยชน์ได้จริงต้องผ่านกระบวนการวิเคราะห์ข้อมูลทางสถิติและนำอัลกอริทึมเข้ามาช่วย ผ่านการประมวลผลด้วยเครื่องคอมพิวเตอร์ เพื่อเป็นเครื่องมือในการค้นพบองค์ความรู้ที่ซ่อนอยู่ในแหล่งข้อมูล การประมวลผลข้อความภาษาไทยมีความยุ่งยากกว่าภาษาอังกฤษ เพราะโครงสร้างประโยคและไวยากรณ์มีความซับซ้อนมากกว่า บทความนี้นำเสนอรูปแบบการจำแนกกลุ่มข้อความภาษาไทยแบบอัตโนมัติ โดยใช้การเรียนรู้ของเครื่อง (machine learning) ด้วยเทคนิค Unsupervised Learning ร่วมกับการประมวลผลภาษาธรรมชาติ (natural language processing) โดยรูปแบบดังกล่าวผ่านการสังเคราะห์และพัฒนากระบวนการทำงานแบ่ง ออกเป็น 3 มอดูล (modules) คือ

- 1) มอดูลการประมวลผลภาษาธรรมชาติ เป็นการวิเคราะห์โครงสร้างข้อความภาษาไทยให้อยู่ในรูปแบบที่เครื่องคอมพิวเตอร์สามารถนำไปประมวลได้ถูกต้อง
- 2) มอดูลการเรียนรู้ของเครื่องแบบไม่มีผู้สอนเป็นโมเดลการเรียนรู้สำหรับการจำแนกกลุ่มข้อความภาษาไทยให้เป็นไปอย่างอัตโนมัติ และ
- 3) มอดูลเหมืองความรู้ คือแหล่งจัดเก็บข้อมูลที่ผ่านการประมวลผลการเรียนรู้ของเครื่องแบบไม่มีผู้สอนในมอดูลที่ 2 ข้อมูลที่จัดเก็บไว้ในเหมืองความรู้จัดว่าเป็นแหล่งขุมทรัพย์ที่มีประโยชน์ต่อการนำไปใช้งานทางด้านธุรกิจ และด้านอื่น ๆ เพราะเป็นแหล่งข้อมูลที่ถูกรวบรวมมาจากทุกช่องทางบนเครือข่ายอินเทอร์เน็ต

เฉลิมชัย พิเศษ (2563) ตั๋วงาน คือ การเขียนคำร้องในระบบและส่งให้กับผู้สนับสนุนด้านไอที ช่วยแก้ไขปัญหา ข้อความในตั๋วงาน มีทั้งประโยคที่เป็นภาษาไทย ภาษาอังกฤษ และภาษาไทยผสม ภาษาอังกฤษ โดยบางครั้งผู้ใช้งานเขียนข้อความที่สั้นมาก ทำให้ไม่สามารถเข้าใจความหมายของ ข้อความและไม่สามารถทำการระบุประเภทของตั๋วงานได้อย่างถูกต้อง ในงานวิจัยนี้ได้ประยุกต์ใช้การ ประมวลผลภาษาธรรมชาติวิทยาการข้อมูล หลักการเรียนรู้ของเครื่องเพื่อช่วยในการแก้ปัญหาการ แยกประเภทตั๋วงาน โดยการวิเคราะห์คำที่ไม่สำคัญและใช้เครื่องมือการตัดคำภาษาไทยในขั้นตอนการ ประมวลผลภาษาธรรมชาติที่สามารถให้ผลการทดลองมีความถูกต้องสูงถึง 91 เปอร์เซ็นต์

2.8 การวิเคราะห์ความรู้สึก (Sentiment Analysis)

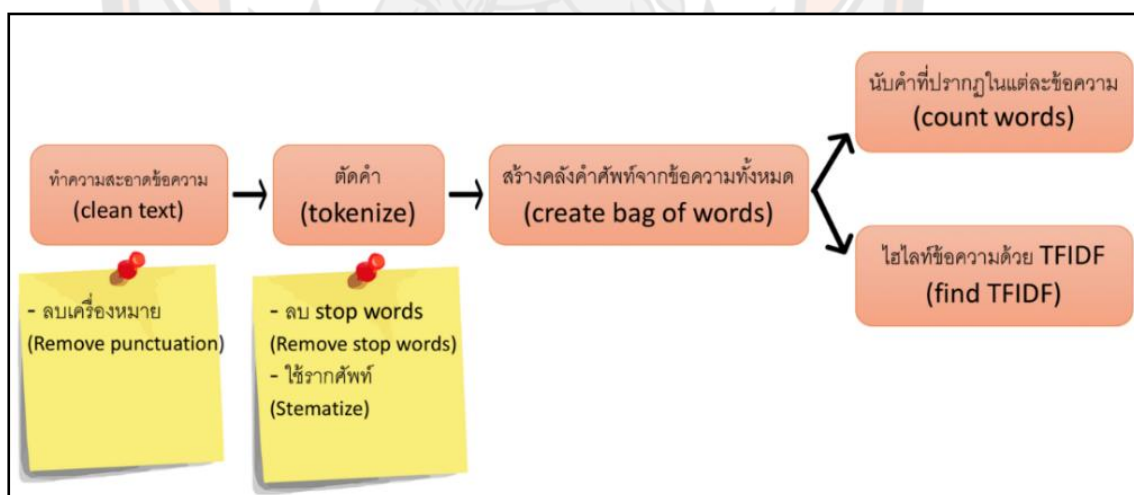
การวิเคราะห์ความรู้สึก (Sentiment Analysis) คือการวิเคราะห์ข้อความ เพื่อค้นหา ความรู้ หรือข้อมูลเชิงลึกที่ซ่อนอยู่ เช่น อารมณ์ของลูกค้า เป็นต้น การวิเคราะห์ความรู้สึกเป็นการ ประยุกต์ใช้งานอย่างหนึ่งของ Natural Language Processing โดยเป็นเทคนิคที่ใช้สำหรับการ จำแนกข้อความออกเป็นกลุ่ม ซึ่งส่วนใหญ่เป็นการวิเคราะห์เพื่อหาว่าข้อความกำลังแสดงอารมณ์ รูปแบบใด ซึ่งความรู้สึกส่วนใหญ่ที่ใช้ในการวิเคราะห์ ได้แก่ ความรู้สึกในทางที่ดี ความรู้สึกในทางที่ ไม่ดี ความรู้สึกปกติ เป็นต้น การวิเคราะห์ความรู้สึกจะใช้เทคนิคทางด้านการทำเหมืองข้อความ (Text Mining) ซึ่งเป็นหนึ่งในเทคนิคของ Natural Language Processing ในการวิเคราะห์ โดย เน้นไปทางด้านของการจำแนกประเภทของข้อความ (Text Classification) การวิเคราะห์ความรู้สึก นั้นได้นำไปประยุกต์ใช้ในงานด้านต่าง ๆ โดยในปัจจุบันสามารถสังเกตได้ คือ การใช้งาน Platform การซื้อขายบนอินเทอร์เน็ตที่มีการใช้การวิเคราะห์ความรู้สึกในการตัดสินใจเห็นต่อสินค้าว่าไป ในทิศทางใด เป็นต้น ซึ่ง Sentiment Analysis เป็นกระบวนการวิเคราะห์อารมณ์หรือความรู้สึก นั้นเอง อย่างเช่น มีลูกค้าโพสต์ผ่าน Social Media เกี่ยวกับสินค้าหรือผลิตภัณฑ์ โดยนำ Data ที่ได้ จากข้อความนั้นมาผ่านกระบวนการวิเคราะห์เพื่อทำให้รู้ว่าข้อความที่ลูกค้าโพสต์นั้นรู้สึกอย่างไรกับ สินค้า โดย Sentiment Analysis ที่ใช้วิเคราะห์ผลตอบรับของลูกค้า เช่น การสังเกตการณ์ Feedback ของลูกค้าผ่านทาง Social Media Post เช่น Twitter ซึ่งเป็นข้อมูลประเภท ข้อความ (text as data) การกำหนด Observation number(จำนวนของข้อมูล) และ Keyword เป็นสิ่งสำคัญที่สุด เพราะต้องการได้ความคิดเห็นส่วนใหญ่ของผู้ใช้บริการและสินค้า แต่ถ้า กล่าวถึงระยะเวลาการเก็บ Data ขึ้นอยู่กับระยะเวลาที่ต้องการทราบ Feedback เช่น ถ้าอยาก ทราบ Feedback ของลูกค้าหลังจากออกสินค้ารุ่นใหม่มาได้ 3 เดือน Data ที่ควรมีคือ โพสต์ หรือข้อความ ทั้งหมดที่กล่าวถึงสินค้าภายใน 3 เดือนย้อนหลังนี้ ซึ่งข้อมูลใน Social Media นั้น เกิดขึ้นมากมายมหาศาลและมีการเพิ่มขึ้นในทุกๆวินาที ไม่สามารถใช้กำลังคนมานั่งวิเคราะห์ sentiment ได้ ดังนั้นจึงมีการใช้เครื่องมือเข้ามาช่วยซึ่งปัจจุบันนี้มี Sentiment Analysis Tool

ให้เลือกใช้หลากหลาย โดยต้องเลือกใช้ให้เหมาะสมกับจุดประสงค์ของธุรกิจ ถ้าในกรณีที่ไม่มี Sentiment Analysis Tool และต้องการสร้างเอง จำเป็นต้องมี Observation number จำนวนมหาศาลเพื่อที่ใช้ศึกษาทุกลักษณะการพูดที่แตกต่างกัน เพราะว่าคำ หรือ วลี หนึ่งจะมีความหมายได้หลายอย่าง ซึ่งจะส่งผลต่อความแม่นยำในการวิเคราะห์ Sentiment (Digital Consult, 2020)

การวิเคราะห์ความรู้สึกด้วย Sentiment Analysis

Sentiment Analysis คือ การวิเคราะห์ความรู้สึก วิเคราะห์อารมณ์จากข้อความ ซึ่งเป็นสาขาหนึ่งของงานด้าน NLP (Natural Language Processing) สามารถใช้ sentiment analysis มาช่วยในการวิเคราะห์ความรู้สึกของคน เพื่อช่วยให้เข้าใจพฤติกรรม สามารถนำผลที่ได้มาปรับให้เข้ากับธุรกิจ เพื่อตอบสนองความต้องการของคนเหล่านั้นได้ตรงจุด มีประสิทธิภาพ ในการที่จะทำ sentiment analysis นั้น มีขั้นตอน ดังนี้ (Chengz, 2019)

1. Data Collection รวบรวมข้อมูลความเห็นจากแหล่งต่างๆที่สนใจ เช่น จาก Social
2. Data Pre-Processing เตรียมข้อมูลให้พร้อมก่อนนำไปวิเคราะห์ เช่น Tokenize
3. Sentiment Analysis นำข้อมูลที่ผ่านการ pre-process แล้วมาทำการวิเคราะห์ เพื่อจำแนก negative / positive sentiment ดังภาพที่ 25



ภาพ 25 ขั้นตอนการเตรียมข้อมูลประเภท text

ที่มา : <https://medium.com>

งานวิจัยการวิเคราะห์ความรู้สึก (Sentiment Analysis)

รวีสุดา เทศเมือง และ นิเวศ จิระวิจิตรชัย (2560) งานวิจัยนี้ได้นำเสนอวิธีการจำแนกความคิดเห็นโดยการวิเคราะห์ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์ ด้านการบริการห้องพัก โรงแรม รีสอร์ท จาก Agoda Thailand และ Twitter Thailand ที่จัดทะเบียนหลักทรัพย์ โดยเก็บข้อมูล จำนวน 2,890 ข้อมูล โดยใช้เทคนิคเหมืองข้อความ วิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิว สินค้าออนไลน์ และสร้างแบบจำลองด้วยอัลกอริทึม 4 วิธี ได้แก่ ซัพพอร์ตเวกเตอร์แมทซ์ ต้นไม้ตัดสินใจ นาอีฟเบย์ และเคเนียร์สเนเบอร์ เพื่อเปรียบเทียบประสิทธิภาพการวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์ จากการทดลองพบว่าคุณลักษณะที่ดีที่สุดคือ ซัพพอร์ตเวกเตอร์แมทซ์ ระดับรองลงมาเป็นนาอีฟเบย์ ต้นไม้ตัดสินใจ และเคเนียร์สเนเบอร์

พิชญะ พรหมลา (2563) การเปรียบเทียบประสิทธิภาพการวิเคราะห์ความพึงพอใจเกี่ยวกับการจัดการเรียนการสอนด้วยกระบวนการวิเคราะห์ความรู้สึกโดยใช้เทคนิคการรวมกลุ่มเพื่อจำแนกข้อมูล ปัจจุบันสถาบันการศึกษาให้ความสำคัญกับการประเมินผลการจัดการเรียนรู้ของครูผู้สอนเป็นอย่างมากนอกจากการประเมินทางตรงยังมีการประเมินทางอ้อมด้วยการสำรวจความพึงพอใจของผู้เรียน โดยใช้แบบสอบถามซึ่งการวิเคราะห์ข้อมูลคำถามปลายปิดในแบบสอบถามนั้นสามารถทำได้ง่ายแต่คำถามปลายเปิดทำได้ยาก และซับซ้อนรวมทั้งอาจเกิดความไม่แม่นยำเนื่องจากอคติจากผู้วิเคราะห์ ข้อมูลได้ในงานวิจัยนี้ได้ใช้กระบวนการวิเคราะห์ความรู้สึก เพื่อวิเคราะห์ความคิดเห็นที่ได้จากแบบสอบถามจำนวน 1,577 ข้อความให้เป็นข้อความพึงพอใจและเปรียบเทียบประสิทธิภาพการจำแนกโดยใช้เทคนิคการรวมกลุ่มเพื่อจำแนกข้อมูล ได้แก่ Vote, Bagging และ Random Forest กับเทคนิควิธีมาตรฐานได้แก่ Decision Tree, Naïve Bayes และ K-NN พบว่าเทคนิควิธีการรวมกลุ่มเพื่อจำแนกข้อมูลแบบ Vote มีประสิทธิภาพมากที่สุด

จากที่ผู้วิจัยได้ศึกษาหนังสือ ตำรา บทความวิจัย บทความวิชาการและเอกสารต่าง ๆ ที่เกี่ยวข้อง โดยผู้วิจัยได้นำมาใช้เป็นแนวคิดในการออกแบบงานวิจัย โดยศึกษาข้อมูลผลิตภัณฑ์ดูแลผิวหน้า เพื่อให้ทราบถึงคุณลักษณะของผลิตภัณฑ์ดูแลผิวหน้าว่ามีลักษณะอย่างไรที่ลูกค้าต้องการนำมาใช้เพื่อบำรุงผิวหน้า โดยลูกค้าส่วนใหญ่จะใช้ช่องทางตลาดซื้อขายสินค้าออนไลน์ในการเลือกซื้อผลิตภัณฑ์ ซึ่งมีความสะดวก รวดเร็วและเป็นที่ยอมรับมากในปัจจุบัน นอกจากนี้ลูกค้าจะสั่งซื้อผลิตภัณฑ์ดูแลผิวหน้าผ่านตลาดซื้อขายสินค้าออนไลน์แล้ว ยังสามารถแสดงความคิดเห็นหรือการรีวิวเพื่อตอบสนองต่อการซื้อว่ามีความรู้สึกอย่างไรกับการสั่งซื้อสินค้า คุณภาพสินค้า การขนส่ง การให้บริการหรืออื่น ๆ ที่เกี่ยวข้อง แต่การรีวิวนั้นอาจเป็นรีวิวจริงมาจากลูกค้าที่สั่งซื้อผลิตภัณฑ์นั้นโดยตรง หรืออาจจะเป็นรีวิวที่น่าสงสัย คือมาจากลูกค้าแต่ยังไม่เคยสั่งซื้อสินค้า อาจเป็นรีวิวที่ลูกค้าจ้างเขียน เขียน

รีวิวโดยได้รับผลตอบแทน เป็นต้น ซึ่งปัญหาของการวิจัยคือรีวิวที่แสดงอยู่ในตลาดซื้อขายสินค้าออนไลน์มีความน่าเชื่อถือหรือไม่ โดยผู้วิจัยใช้ทฤษฎีหลักการเรียนเรียนรู้ของเครื่อง (Machine Learning) มาช่วยในการวิเคราะห์ข้อมูลรีวิว การวิเคราะห์ความรู้สึก โดยใช้หลักการการประมวลผลภาษาธรรมชาติโดยการทำงานของรีวิวคือการแปลงข้อมูลดิบที่เป็นภาษาปกติของมนุษย์ ให้เป็นข้อมูลที่พร้อมต่อการทำงานของอุปกรณ์ ผ่านอัลกอริทึมและกระบวนการทางภาษาต่าง ๆ เพื่อให้สามารถนำข้อมูลนั้นไปใช้งานการวิเคราะห์และสามารถระบุรีวิวจริงหรือรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้านตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่องได้



บทที่ 3

วิธีดำเนินงานวิจัย

งานวิจัยนี้นำเสนอวิธีดำเนินงานวิจัยการระบุรีวิวนที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง ดังนี้

3.1 ข้อมูลเบื้องต้น

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ใช้ในการทำวิจัย ได้แก่ รีวิวผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ จะใช้ข้อมูลรีวิวที่เป็นภาษาไทยในการเก็บมาวิเคราะห์ข้อมูล ซึ่งจะแยกข้อมูลรีวิวผลิตภัณฑ์ดูแลผิวหน้า ที่ได้รับความนิยมในท้องตลาด จำนวน 10 แปรนต์ โดยข้อมูลจะไม่ปรากฏให้เห็นในงานวิจัยฉบับนี้ ซึ่งข้อมูลที่ได้ทั้งหมดอยู่ในรูปแบบชุดข้อความเท่านั้น

ตัวแปรที่ใช้ในการวิจัย

ข้อมูลที่ได้จากการรีวิวผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ จะทำการประมวลผลจากข้อมูลดิบของรีวิวที่กำหนดให้ที่อยู่ในรูปแบบไม่เป็นโครงสร้าง (Unstructured data) ให้เป็นข้อมูลที่อยู่ในรูปแบบเป็นโครงสร้าง (Structure data) และวิเคราะห์ข้อมูลออกมาตามกระบวนการเรียนรู้ด้วยเครื่อง

เครื่องมือและการพัฒนาเครื่องมือ

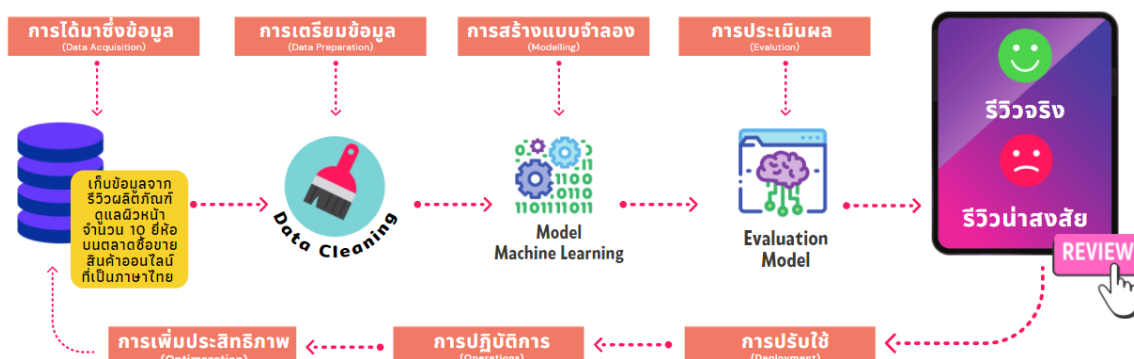
ในงานวิจัยนี้ใช้ซอฟต์แวร์และฮาร์ดแวร์ เพื่อดึงรีวิวผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์และใช้เทคนิควิธีการเรียนรู้ของเครื่อง (Machine Learning) ในการค้นหาคุณลักษณะเฉพาะและรูปแบบจำลองที่ตรวจสอบรีวิวจากผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ ประกอบด้วย

1. เครื่องมือด้านซอฟต์แวร์
 - Google Drive
 - Google Colab (Colaboratory)
 - Scikit-Learn (Sklearn)
 - PyThaiNLP
 - Machine Learning

2. เครื่องมือด้านฮาร์ดแวร์
 - Notebook DELL Inter® Core™ i7
3. ภาษาที่ใช้ในการพัฒนา
 - Python

3.2 วิธีดำเนินการ (Methods)

ผู้วิจัยมีแนวคิดในการนำหลักการของวงจรชีวิตของการเรียนรู้ของเครื่อง (Machine Learning Life Cycle) มาใช้ในวิธีดำเนินงานวิจัย โดยภาพรวมของวิธีดำเนินการทั้งหมด แสดงสรุปได้ดังภาพที่ 26



ภาพ 26 ภาพรวมของวิธีดำเนินการ

จากภาพรวมของวิธีดำเนินการ ประกอบไปด้วยการดำเนินงานโดยมีรายละเอียดทั้งหมด 7 ขั้นตอน ดังนี้

การได้มาซึ่งข้อมูล (Data Acquisition)

ผู้วิจัยได้เก็บข้อมูลจากรีวิวผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ที่เป็นภาษาไทย ระหว่างเดือนมกราคม - พฤศจิกายน 2564 จำนวน 60,081 รีวิว จากผลิตภัณฑ์ดูแลผิวหน้า 10 ยี่ห้อที่มีจำนวนขายสูงสุดในตลาดซื้อขายสินค้าออนไลน์ตามลำดับ โดยเลือกเก็บข้อมูล 8 คุณลักษณะ ได้แก่ ยี่ห้อสินค้า ราคา ชื่อร้านค้า ความเร็วในการตอบกลับ ชื่อลูกค้า ข้อมูลรีวิว วันเวลาที่รีวิว จำนวนดาว เพื่อนำข้อมูลทั้งหมดมาใช้ในการพิจารณาเพื่อประเมินความน่าเชื่อถือของข้อมูลรีวิว ว่ามีความผิดปกติของข้อมูลที่มีความเกี่ยวข้องกับข้อมูลรีวิวหรือไม่ โดยข้อมูลไม่ปรากฏให้เห็นในงานวิจัยฉบับนี้ ข้อมูลที่ได้ทั้งหมดอยู่ในรูปแบบชุดข้อความเท่านั้น เพื่อรวบรวมคำคุณลักษณะที่แสดงความรู้สึกรีวิวในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้า

การเตรียมข้อมูล (Data Preparation)

เมื่อได้จัดเก็บข้อมูลรีวิวลักษณะที่ดูแล้วหน้าของลูกค้านตลาดซื้อขายสินค้าออนไลน์ โดยใช้เงื่อนไขของการดึงรีวิว (Scrape Reviews) แล้ว ในขั้นตอนถัดไปต้องเตรียมข้อมูลให้พร้อม เพราะเป็นกระบวนการที่สำคัญมากอย่างหนึ่ง หากการเตรียมข้อมูลทำได้ไม่ดี มีโอกาสสูงที่ก่อให้เกิดความเสียหายในขั้นตอนอื่น ๆ ที่ตามมาได้หรืออาจส่งผลให้ผลการวิเคราะห์หรือการตีความจากการนำข้อมูลไปใช้ผิดเพี้ยนไปจากที่ควรจะเป็น จากนั้นนำข้อมูลรีวิวทั้งหมดมาเข้าสู่กระบวนการเตรียมข้อมูลโดยใช้ Google Colab (Colaboratory) ที่นำไปรันอยู่บนคลาวด์ โดยใช้ภาษา Python เป็นภาษาหลักในการเขียนโปรแกรม ดังขั้นตอนต่อไปนี้

1. Download Library ซึ่งในแต่ละตัวไว้ใช้ในการประมวลผลภาษาไทย ดังนี้
 - PythaiNLP คือ Library แพคเกจของภาษา Python ใช้สำหรับประมวลผลข้อความและการวิเคราะห์ทางภาษา ใช้กับภาษาไทยโดยเฉพาะ ซึ่งมีฟังก์ชันอักษรไทย คำไทย เรียงคำภาษาไทย
 - Stop_Words คือ ภาษาไทย ตัดคำภาษาไทย วิเคราะห์ชนิดของคำทางไวยากรณ์ ตรวจตัวสะกด แก่คำผิด การแบ่งส่วนคำภาษาไทย
 - Pipeline คือ การบอกขั้นตอนต่าง ๆ ที่จำเป็นในการสร้างแบบจำลองโดย เรียงเป็นขั้นตอน จากจุด 1 จุด 2 จุด 3 ให้รู้ว่าขั้นตอนอะไรเริ่มก่อนโดย Pipeline ทำการ “ต่อท่อ” ให้กับขั้นตอนเหล่านั้นจนได้เป็นโมเดลสุดท้ายออกมา มีความสะดวกสบายในการเขียน Code การ Training model หลาย ๆ แบบ
 - Python-crfsuite คือ ต้อง install พร้อมกัน เพื่อเชื่อมกับ google drive
 - Openpyxl คือ เป็น Library ตัวหนึ่งในภาษา Python ซึ่งมีไว้สำหรับการอ่านหรือเขียนไฟล์ Excel 2010 เช่น xlsx, xlsxm, xlsx, xlsxm ต้องติดตั้งตัว OpenPyXL เพิ่มเพราะไม่ได้เป็น Library ที่มีมาตั้งแต่แรกใน Python ดังภาพที่ 29

```
!pip -q install pythainlp
!pip -q install stop_words
!pip -q install pipelinehelper
!pip -q install python-crfsuite
!pip -q install openpyxl==3.0.9
!pip install --upgrade openpyxl
```

ภาพ 29 Download Library

2. เชื่อมต่อจาก Google Colab ไปยัง Drive เพื่อเรียกข้อมูลรีวิวทั้งหมดที่ถูกจัดเก็บไว้ออกมาใช้งาน และติดตั้ง Pandas ซึ่งเป็นเครื่องมือในการวิเคราะห์และจัดการข้อมูลที่เร็ว มีประสิทธิภาพ มีความยืดหยุ่น และใช้งานง่าย สามารถนำเข้าข้อมูลเช่นไฟล์ CSV, TSV หรือจาก SQL database นำไปสร้างเป็น Python Object ที่มีลักษณะเป็น row และ column ทำให้สามารถใช้งานข้อมูลเหล่านั้นได้ง่าย โดยแบบจำลองนี้เป็นแบบจำลองประเภท Supervised Learning คือการเรียนรู้แบบมีผู้สอนในข้อมูลที่ใช้ฝึก (Training Data) ใช้คนมาแยกประเภทบอกผลลัพธ์ (Label) ที่ควรเป็นไปไว้ด้วย ซึ่งใช้หลักการทางทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการแบ่งประเภทของรีวิวออนไลน์ที่ผู้วิจัยได้ศึกษามา โดยนำมาใช้ในการสอนให้แบบจำลอง Machine Learning เรียนรู้จากข้อมูลรีวิว Labeled Data โดยบอกแบบจำลองว่าข้อมูลรีวิวนั้นคือรีวิวอะไร เช่น จะสอนให้แบบจำลองจำรีวิวว่ารีวิวไหนจริง หรือรีวิวไหนที่น่าสงสัย โดยนำรีวิวทั้งหมดที่มีความหลากหลายมาดูพร้อมใส่ Label ว่ารีวิวนี้เป็นรีวิวจริง หรือรีวิวที่น่าสงสัย แบบจำลองนี้จะสร้าง Label เป็นตัวกำกับกับข้อมูลในแต่ละ record ซึ่งเทคนิคการเรียนรู้ของเครื่องจะเรียนรู้จากสิ่งนี้โดยการเรียนรู้ทีละ record เพื่ออัปเดตตัวเองไปเรื่อย ๆ จนได้ผลทำนายที่แม่นยำมากขึ้น เห็นว่าการจำแนกข้อมูลรีวิว (Classification) ทำการสร้างแบบจำลองเพื่อทำนายหมวดหมู่ของรีวิว (Categorical) โดยใช้ข้อมูลรีวิวทั้งหมด 60,081 รีวิว, 13 คอลัมน์ ได้แก่ ลำดับ ยี่ห้อสินค้า ราคา ชื่อร้านค้า ความเร็วในการตอบกลับ ชื่อลูกค้า ข้อมูลรีวิว วันเวลาที่รีวิว จำนวนดาว รีวิวจริง (Real) รีวิวที่น่าสงสัย (Suspicious) รีวิวที่ไม่มีข้อความ (Missing value) ความรู้สึก (Sentiment) ซึ่งข้อมูลที่สนใจคือ ข้อมูลรีวิวผลิตภัณฑ์ดูแลผิวหน้า (Content) รีวิวจริง (Real) รีวิวที่น่าสงสัย (Suspicious) รีวิวที่ไม่มีข้อความ (Missing value) และความรู้สึก (Sentiment) และนำข้อมูลรีวิวทั้งหมดมาจำแนก จากจำนวนรีวิวทั้งหมด 60,081 รีวิว แยกเป็น รีวิวจริง (Real) รีวิวที่น่าสงสัย (Suspicious) รีวิวที่ไม่มีข้อความ (Missing value) และค่าที่ไม่เกี่ยวข้อง (Error) ดังภาพที่ 30


```
# function for clean text
def cleanText(raw_html):
    cleanweb = re.compile("<.*?>")
    cleanhashtag = re.compile("#")
    cleantext = re.sub(r'http\S+', 'http', raw_html) ## remove HTML
    for (k,v) in special.items(): ## remove special Char
        cleantext = cleantext.replace(k,v)
    cleantext = re.sub(cleanhashtag,"",cleantext) ## remove hashtag

    for c in string.punctuation:
        cleantext = re.sub(r'\{',format(c),"",cleantext) # clean punctuation char
    cleantext = ''.join(cleantext.split())

    # Remove EMOJI
    RE_EMOJI = re.compile("[")
        u"\U0001F600-\U0001F64F" # emoticons
        u"\U0001F300-\U0001F5FF" # symbols & pictographs
        u"\U0001F680-\U0001F6FF" # transport & map symbols
        u"\U0001F1E0-\U0001F1FF" # flags (iOS)
        u"\U00002500-\U00002BEF" # chinese char
        u"\U00002702-\U000027B0"
        u"\U00002702-\U000027B0"
        u"\U000024C2-\U0001F251"
        u"\U0001F926-\U0001F937"
        u"\U00010000-\U0010ffff"
        u"\u2640-\u2642"
        u"\u2600-\u2B55"
        u"\u200d"
        u"\u200b"
        u"\u23cf"
        u"\u23e9"
        u"\u231a"
        u"\ufe0f" # dingbats
        u"\u3030"
        u"8"
        u"000"
        u"000"
        u"RT"
        "]+", flags=re.UNICODE) # ตัวแทน Code ใหม่
    #cleantext = RE_EMOJI.sub(r'', cleantext)
    return cleantext
```

ภาพ 32 การลบข้อมูลที่มีสิ่งรบกวน (noise)

```
print(df.content[24500:24510])

40114 ข้อครั้งแรก ยังไม่เคยใช้เจ้าตัวนี้ ถ้าใช้ดีแล...
40115 ข้อครั้งที่สองแล้วจ๋าาาา ไม่รู้ใช้แล้วดีมีข...
40116 ข้อครั้งที่สองแล้วจ๋า รับรองต้องมีครั้งที่สาม...
40117 ข้อครั้งที่สองแล้วคะ มีของแถมตลอด
40118 ข้อครั้งที่สองแล้ว ใช้ดีมากๆ ใช้แล้วเสียหาย ใ...
40119 ข้อครั้งที่สองแล้ว ครั้งแรกข้อมาลองใช้ดู ดีว...
40120 ข้อครั้งที่สองแล้ว คนที่ใช้พอใจสินค้า ได้รับส...
40121 ข้อครั้งที่3แล้วคะ คุณภาพดี สินค้าถึงไว มีกา...
40122 ข้อครั้งที่2แล้ว เลือกร้านนี้เพราะมันใจในสิน...
40123 ข้อครั้งที่2 ประทับใจเหมือนเดิม ร้านค้าบริการ...
Name: content, dtype: object

print(df.content[24500:24510].apply(cleanText))

40114 ข้อครั้งแรก ยังไม่เคยใช้เจ้าตัวนี้ ถ้าใช้ดีแล...
40115 ข้อครั้งที่สองแล้วจ๋าาาา ไม่รู้ใช้แล้วดีมีข...
40116 ข้อครั้งที่สองแล้วจ๋า รับรองต้องมีครั้งที่สาม ...
40117 ข้อครั้งที่สองแล้วคะ มีของแถมตลอด
40118 ข้อครั้งที่สองแล้ว ใช้ดีมากๆ ใช้แล้วเสียหาย ใ...
40119 ข้อครั้งที่สองแล้ว ครั้งแรกข้อมาลองใช้ดู ดีว...
40120 ข้อครั้งที่สองแล้ว คนที่ใช้พอใจสินค้า ได้รับส...
40121 ข้อครั้งที่3แล้วคะ คุณภาพดี สินค้าถึงไว มีกา...
40122 ข้อครั้งที่2แล้ว เลือกร้านนี้เพราะมันใจในสิน...
40123 ข้อครั้งที่2 ประทับใจเหมือนเดิม ร้านค้าบริการ...
Name: content, dtype: object
```

ภาพ 33 ตัวอย่างวิธีที่ถูกลบหรือลดข้อมูลที่มีสิ่งรบกวน (noise)

4. การทำ Tokenization เนื่องจากภาษาไทย เป็นภาษาที่เขียนติดกันหมด ไม่มีการเว้นคำด้วย Space เหมือนภาษาอังกฤษ ทำให้การตัดคำภาษาไทย หรือ Tokenization มีความซับซ้อน และการตัดคำที่ถูกต้องมีความสำคัญต่อการนำข้อมูลคำศัพท์ไปประมวลผลต่อ โดยใช้ Word_tokenize ตัดคำ Word, Sentence โดยอัลกอริทึมใหม่ล่าสุด และเป็น Default ของ PyThaiNLP tokenize ณ ขณะนี้ คือ newmm ที่ใช้ อัลกอริทึม หา maximum matching จาก Dictionary ที่ทำงานได้อย่างรวดเร็ว และถูกต้องพอสมควร รวมทั้งลบตัวเลข ลบช่องว่าง และการจับกลุ่มคำที่ปรากฏด้วยกัน การหาความสัมพันธ์ของคำ รวมไปถึงการเลือกกำจัดคำบางคำที่ไม่มีความสำคัญกับการวิเคราะห์ เปรียบเทียบการตัดคำภาษาไทยด้วยอัลกอริทึมต่าง ๆ โดยนำไปเชื่อมโยงกับ WordNet เป็นฐานข้อมูลคำศัพท์ของความสัมพันธ์ทางความหมายระหว่างคำต่างๆ ที่เชื่อมโยงคำเข้าไปในความสัมพันธ์ทางความหมายรวมทั้งคำพ้องความหมาย และเข้าสู่กระบวนการในการแปลง Word ด้วยรายการคำศัพท์ใน Dictionary การวิเคราะห์หลักไวยากรณ์ของภาษาอย่างเหมาะสม ในการแปรคำ ผันคำ เพื่อกำจัดความผันผวนของคำ (Inflection) โดย Lemmatization ตัดส่วนท้ายของคำ ให้เหลือแต่รูปฟอร์มพื้นฐานเป็นคำใน Dictionary เรียกว่า Lemma ซึ่งมีลักษณะเป็น list ของคำต่างๆ โดยเหลือข้อมูลรีวิวที่ได้ทำการวิเคราะห์แบบพยากรณ์เพียง 34,123 รีวิว จากนั้นจะนำข้อมูลที่ใช้ฝึกที่มี Label ไปผ่านอัลกอริทึมสำหรับสร้างแบบจำลองที่ไว้ทำนายผลลัพธ์ โดยมีการแบ่งชุดข้อมูล Split ออกเป็น 2 ส่วน คือ Training Set และ Test Set โดยแบ่งเป็น 90% มาให้โมเดลเรียนรู้ Training Set และอีก10% เอาไว้ทดสอบประสิทธิภาพ Test Set ดังภาพที่ 34 โดยแสดงตัวอย่างการทำ Tokenization ดังภาพที่ 35

```

# Text token process
def tokenization(text):
    tokens = word_tokenize(text)
    #print(tokens)
    tokens = [i for i in tokens if not i in thai_stopwords() and not i in en_stop]
    #print(tokens)
    # Thai
    tokens_temp=[]
    for i in tokens:
        w_syn = wordnet.synsets(i)
        if (len(w_syn)>0) and (len(w_syn[0].lemma_names('tha'))>0):
            tokens_temp.append(w_syn[0].lemma_names('tha')[0])
        else:
            tokens_temp.append(i)

    tokens = tokens_temp
    # ลบตัวเลข
    tokens = [i for i in tokens if not i.isnumeric()]
    # ลบช่องว่าง
    tokens = [i for i in tokens if not ' ' in i]
    # print(tokens)
    return tokens

df['tokenized_text'] = df.content.apply(str).apply(cleanText).apply(tokenization)
df.shape
(34123, 3)

```

ภาพ 34 การทำ Tokenization

'ลอง',	'coins',
'ใช่แล้ว',	'คัม',
'ดีมาก',	'แกม',
'กก',	'จัดส่ง',
'กก',	'ไว',
'กก',	'แพ',
'เนื้อ',	'ด',
'ครีม',	'ดี',
'หน้า',	'สั่ง',
'รอย',	'พริงนี่',
'แดง',	'เย็น',
'รอ',	'👍👍'],
'สั๊กพัก',	['🤔',
'เซ',	'สินค้า',
'รัม',	'เรียบรอย',
'แกม',	'แพ',
'เนื้อ',	'ด',
'นุ่ม',	'ดี',
'ไกล',	'จัดส่ง',
'ขมขื่น',	'กันกระแทก',
	'ซื้อ',
	'ดอน',
	'ลดราคา',
	'Coins',
	'คืน',
	'อีกด้วย',
	'คัม',

ภาพ 35 ตัวอย่างการทำ Tokenization

```

from sklearn.model_selection import train_test_split
import numpy as np

X = df.tokenized_text.copy()
y = df.spam.copy()
y = y.astype(np.int)

X_train_, X_test_, y_train_, y_test_ = train_test_split(
    X, y, test_size=0.10, random_state=42)

X_test_
20606  [เขียน, โปร, นะคะ, โควิด, แบบนี้, ใตราคา, ย...
6925   [สินค้า, ดี, คุณภาพ, ราน, จัดส่ง, ดีมาก, คา, ...
46462  [ขนส่ง, แท, ค, สินค้า, ดี, สินค้า, ช่างโน, ส...
34935  [การจัดส่ง, คุณภาพ, สินค้า, ดี]
28274  [สินค้า, การจัดส่ง, ดี, คุณภาพ, สินค้า, หมดอาช...
...
17676  [พี, packaging, ดี, โฟม, ล้าง, หน้าที่, ลอง, ...
11329  [สินค้า, คุณภาพดี, ขนส่ง, รวดเร็วทันใจ, จจจจจ...
48010  [โปรโมชัน, คุ่มคำ, คุ่ม, คำ, แบบนี้, นะคะ]
56192  [ใช่, ดี, ประจำ, เซลล์]
18555  [ليب, เนื้อ, งาม, ดีมาก, ไข่แล้ว, ปาก, แห้ง, ...
Name: tokenized_text, Length: 3413, dtype: object

```

ภาพ 36 การนำ Scikit-learn มาใช้งาน

นำ Scikit-learn (sklearn) มาใช้งาน ดังภาพที่ 36 ซึ่งเป็นคำสั่งที่ช่วยในการทำนายข้อมูลที่เกิดขึ้นโดยนำหลักการจาก Supervised Machine Learning เป็นไลบรารีฟรีในภาษาไพธอน สำหรับการพัฒนาโปรแกรมโดยใช้การเรียนรู้ของเครื่อง มีจุดเด่นคือฟังก์ชันในการแบ่งประเภทข้อมูล การแบ่งกลุ่มข้อมูล การวิเคราะห์การถดถอยหลายอย่างไม่ว่า เช่น ซัพพอร์ตเวกเตอร์แมชชีน การเรียนรู้ต้นไม้ตัดสินใจ และการแบ่งกลุ่มข้อมูลแบบเคมีน เป็นต้น มีการแบ่งประโยคที่เอาไว้ train (X_{train}) ประโยคที่เอาไว้ test (X_{test}) ที่บอกไว้ว่าเป็นรีวิวจริงหรือรีวิวที่น่าสงสัย โดยการประยุกต์ใช้ Generalization เป็นสิ่งที่แบบจำลองเรียนมากับข้อมูลชุดใหม่ และหลังจากที่เรียนรู้ว่าจริงหรือไม่ว่ารีวิวไหนเป็นรีวิวจริงหรือรีวิวไหนเป็นรีวิวที่น่าสงสัย เป็นการจับกลุ่มรีวิวที่แบ่งแล้ว จับมาเข้าด้วยกัน โดยใช้ spacebar คั่นไว้ จากนั้นนำข้อมูลที่ใช้ฝึกที่มี Label ไปผ่านอัลกอริทึมสำหรับสร้างแบบจำลองที่ไว้ทำนายผลลัพธ์ ในขั้นตอนถัดไปไม่สามารถนำข้อมูลด้านบนเข้าแบบจำลองได้โดยตรง แต่ต้องแปลงที่แบ่งคำไว้แล้วให้เข้ากลับมาเป็นประโยคก่อน โดยต้องจับกลุ่มคำที่แบ่งแล้วเข้ามาไว้ด้วยกัน โดยใช้ Spacebar เป็นตัวแยกกันระหว่างประโยค จะเห็นว่าคำที่แยกกันจะมาอยู่รวมกัน แต่จะมี spacebar คั่นอยู่ ดังภาพที่ 37


```

from sklearn.feature_extraction.text import TfidfVectorizer

vectorizer = TfidfVectorizer(tokenizer = lambda x: x.split())
X_train_vec = vectorizer.fit_transform(space_X_train)
X_test_vec = vectorizer.transform(space_X_test)

```

ภาพ 38 การกำหนด TfidfVectorizer จาก sklearn library

หลังจากที่ได้ข้อความต้อง Test ข้อความนี้ให้เป็นตัวเลขก่อน คือ การทำ Vectorizer เป็นการเปลี่ยนข้อความเป็นตัวเลข ก่อนที่คอมพิวเตอร์จะเข้าใจต้องเปลี่ยนจากข้อความให้เป็นตัวเลข เช่น คำว่า “มีปัญหา ดีเยี่ยม หรือ สุดยอด” โดยจำว่า “มีปัญหา” เป็นตัวเลข 5315, “ดีเยี่ยม” เป็นตัวเลข 3159, “สุดยอด” เป็นตัวเลข 6700 และหลังจากที่รู้แล้วว่าตัวเลขคืออะไร มีการเพิ่มเติมว่ารีวิวนี้มีตัวเลขเท่าไร ตัวเลขนี้มีกี่ตัว จำความถี่ของคำนั้น คำทุกตัวที่แปลงเป็นตัวเลขและถูกเก็บอยู่ในชุดข้อมูลที่เป็น Dictionary ดังภาพที่ 39

'พลาสติก': 4858,	'ยอด': 5502,
'บัง': 4292,	'ดีเลิศ': 3161,
'ฝา': 4770,	'เด็ดขาด': 7972,
'กล้วย': 1608,	'อันดับ': 7194,
'หุง': 4946,	'ศักดิ์สิทธิ์': 6383,
'กาล': 1753,	'มหัศจรรย์': 5186,
'ตอนกลางคืน': 3317,	'สุดยอด': 6700,
'สำหรับ': 6630,	'ยอดเยี่ยม': 5508,
'ทางการ': 3715,	'ชนะเลิศ': 2737,
'มีปัญหา': 5315,	'ดีเด่น': 3158,
'สรุป': 6507,	'สุดเหวี่ยง': 6701,
'ความไว': 2264,	'วิเศษ': 6321,
'ดีเยี่ยม': 3159,	'ฮีโร่': 7340,
'บอมเบอร์': 5605,	'ชั้นเลิศ': 2792,
'สะดวกสบาย': 6544,	'ล้ำ': 6126,
'พัสด': 4908,	'เด็ด': 7971,
	'หนึ่งเดียว': 6810.

ภาพ 39 การทำ Vectorizer

ขั้นตอนถัดไปทำการนำชุดข้อมูลเข้าแบบจำลอง ที่ Import จาก Dictionary โดยแบ่งชุดข้อมูลออกเป็น 2 ส่วนคือ Training set สำหรับการ Train และ Test set สำหรับทดสอบ โดยแบ่งมาจากชุดข้อมูลเพื่อนำมาทดสอบความแม่นยำ ความถูกต้องของแบบจำลองที่ Train เรียบร้อยแล้ว ซึ่งการ Train กับตัวข้อมูลที่เป็นรีวิวกับตัวข้อมูลที่เป็น

เป้าหมาย (Target) ต้องการให้ระบบ Machine Learning มีความสามารถทำนาย จำแนก แยกแยะรีวิวจริงและรีวิวที่น่าสงสัยได้ ทำแบบนี้เหมือนกันทุกตัวคือ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซ์พอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เทคนิค ต้นไม้ตัดสินใจ (Decision Tree) หลังจากที่ทำเสร็จแล้วจะมีการวัดประสิทธิภาพโดยทำนาย (Predict) รีวิวที่ใช้ในการไว้ Test ดังภาพที่ 40

```
[ ] dt = DecisionTreeClassifier()
dt.fit(X_train_vec,y_train_)

DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='gini',
max_depth=None, max_features=None, max_leaf_nodes=None,
min_impurity_decrease=0.0, min_impurity_split=None,
min_samples_leaf=1, min_samples_split=2,
min_weight_fraction_leaf=0.0, presort='deprecated',
random_state=None, splitter='best')

[ ] svc = SVC(kernel='linear')
svc.fit(X_train_vec, y_train_)

SVC(C=1.0, break_ties=False, cache_size=200, class_weight=None, coef0=0.0,
decision_function_shape='ovr', degree=3, gamma='scale', kernel='linear',
max_iter=-1, probability=False, random_state=None, shrinking=True,
tol=0.001, verbose=False)

[ ] nbv = BernoulliNB()
nbv.fit(X_train_vec, y_train_)

BernoulliNB(alpha=1.0, binarize=0.0, class_prior=None, fit_prior=True)

[ ] nbv_pred = nbv.predict(X_test_vec)
svc_pred = svc.predict(X_test_vec)
dt_pred = dt.predict(X_test_vec)
print("----- Naives Bayes -----")
print(classification_report(y_test_, nbv_pred, target_names = ['real','spam']))
print("\n----- Support Vector Machine -----")
print(classification_report(y_test_, svc_pred, target_names = ['real','spam']))
print("\n          Decision Tree          ")
print(classification_report(y_test_, dt_pred, target_names = ['real','spam']))
```

ภาพ 40 การทำ Machine Learning ด้วยเทคนิคทั้ง 3 วิธี

การประเมินผล (Evaluation)

หลังจากที่ได้แบบจำลองแล้ว จะทำการประเมินผลว่าแบบจำลองนั้นมีความถูกต้องแม่นยำมากน้อยเพียงใด โดยประเมินผลข้อมูลรีวิวเพื่อทดสอบแบบจำลองโดยชุดข้อมูลทดสอบ (Testing Data) โดยวัดประสิทธิภาพด้วยค่า ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) และ ค่าประสิทธิภาพโดยรวม (F1-Score) คือค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall เพื่อวัดความสามารถของโมเดล จากนั้นทำการเปรียบเทียบประสิทธิภาพค่าเฉลี่ยในแต่ละเทคนิค เพื่อให้ได้แบบจำลองจากเทคนิควิธีการเรียนรู้ของเครื่องที่ดีที่สุด

เมื่อได้แบบจำลองแล้วจะทำการค้นหาแบบกริด (Grid Search) ซึ่งเป็นเครื่องมือในการพัฒนาแบบจำลองให้เก่งขึ้น เป็นเทคนิคที่ใช้ในการหาค่า Hyperparameter ที่เป็นการสุ่มที่เข้าใจง่ายและตรงไปตรงมา ด้วยการใช้พารามิเตอร์ที่กำหนดไว้ล่วงหน้าทุกชุด และประเมินประสิทธิภาพหรือความแม่นยำของโมเดลแต่ละชุด โดยค่าทั้งหมดอยู่ในรูปของเมทริกซ์ (Matrix) พารามิเตอร์แต่ละชุดถูกนำมาพิจารณาและสังเกตความถูกต้อง โมเดลที่มีชุดพารามิเตอร์ที่ให้ความแม่นยำสูงสุดถือว่าดีที่สุด การเปลี่ยนแปลง Version ของแบบจำลอง Machine Learning ที่ได้มาเพื่อให้เก่งขึ้น library ที่ Import มีส่วนช่วยในการทำ pipeline คือ ทำการ train หลายครั้ง หลาย version จะใช้ pipeline เป็นเหมือนตัว Automation ที่จะ Run แบบจำลองโดยไม่ต้องสร้าง pipeline ขึ้นมาก่อน ประกอบไปด้วย 2 ส่วน คือ การทำ Vectorizer คือ แปลงข้อความให้เป็นตัวเลข และการทำ Classifier คือ แบบจำลอง Machine Learning ทั้ง 3 ก่อนหน้านี้ ดังภาพที่ 41 หลังจากนั้นสามารถ Apply ใช้เครื่องมือ Gridsearch เข้าไป ซึ่งเป็นชุดข้อมูลที่ไม่เคยเห็นมาก่อน โดยแบ่งชุดข้อมูลออกเป็นชุดการฝึก (Train) และชุดตรวจสอบ (Test) ได้เลย ใช้วิธี K-fold cross validation คือการที่เราแบ่งข้อมูลเป็นจำนวน K ส่วนโดยการในแต่ละส่วนจะต้องมาจากสุ่มเพื่อที่จะให้ข้อมูลของเรากระจายเท่าๆ กัน นอกจากการหา Training Data ที่ดีที่สุดแล้ว Cross Validation ยังสามารถใช้เปรียบเทียบได้อีกว่าควรใช้วิธีไหนที่เหมาะสมที่สุดในการสร้างโมเดลของเรา ซึ่งในปัจจุบันเป็นวิธีที่นิยมใช้มากที่สุด โดยมีขั้นตอนในการแบ่งข้อมูล Training Data ออกเป็น N ส่วน จากนั้นจะสร้างโมเดลด้วย N-1 ส่วน และทดสอบโมเดลกับส่วนที่เหลือ ในการวนซ้ำแต่ละครั้งจะมีการเปลี่ยนส่วนของที่ใช้ในการทดสอบโมเดล เมื่อสร้างโมเดลครบ N ครั้ง อัลกอริทึมจะทำการหาค่าเฉลี่ยของผลของการสร้างโมเดล เพื่อให้ได้ผลลัพธ์ที่มีประสิทธิภาพที่สุด เพื่อวัดประสิทธิภาพของแบบจำลอง มีทั้งหมด 608 candidates หรือ 600 version ทำเป็น kcode 10 ชุด คือ 600 version ทำทั้งหมด 10 ชุด data ที่แบ่งออกมา กลายเป็นว่ามีทั้งหมด 6000 ต้องทำการ Train 6000 ครั้ง ดังภาพที่ 42

```

pipe = Pipeline([
  ('vect', PipelineHelper([
    ('tfidf', TfidfVectorizer(tokenizer = lambda x: x.split()))
  ])),
  ('classifier', PipelineHelper([
    ('svm', LinearSVC()),
    ('nb', BernoulliNB()),
    ('dt', DecisionTreeClassifier())
  ]))
])

params = {
  'vect__selected_model': pipe.named_steps['vect'].generate({
    'tfidf__ngram_range': [(1,2),(2,3)],
    'tfidf__max_df': [0.9,0.8],
    'tfidf__min_df': [5,6]
  }),
  'classifier__selected_model': pipe.named_steps['classifier'].generate({
    'svm__C': [0.1, 1.0],
    'svm__tol': [0.0001, 0.001],
    'svm__class_weight': [None, 'balanced'],
    'dt__max_depth': [30, 40],
    'dt__criterion': ['gini','entropy'],
    'dt__min_samples_split': [4,5],
    'dt__min_samples_leaf': [3,4],
    'dt__max_features': ['auto','sqrt'],
    'dt__max_leaf_nodes': [20,30],
    'nb__alpha': [1, 0.1, 0.01, 0.001],
  })
}

grid = GridSearchCV(pipe, params, verbose=2, scoring='f1_micro', cv =10, n_jobs =4)
grid.fit(space_X, y)

```

ภาพ 41 การทำ Pipeline

```

Fitting 10 folds for each of 608 candidates, totalling 6080 fits
[Parallel(n_jobs=4)]: Using backend LokyBackend with 4 concurrent workers.
[Parallel(n_jobs=4)]: Done 33 tasks      | elapsed: 23.3s
[Parallel(n_jobs=4)]: Done 154 tasks   | elapsed: 2.0min
[Parallel(n_jobs=4)]: Done 357 tasks   | elapsed: 4.9min
[Parallel(n_jobs=4)]: Done 640 tasks   | elapsed: 9.1min
[Parallel(n_jobs=4)]: Done 1005 tasks  | elapsed: 14.0min
[Parallel(n_jobs=4)]: Done 1450 tasks  | elapsed: 19.9min
[Parallel(n_jobs=4)]: Done 1977 tasks  | elapsed: 26.9min
[Parallel(n_jobs=4)]: Done 2584 tasks  | elapsed: 34.9min
[Parallel(n_jobs=4)]: Done 3273 tasks  | elapsed: 44.1min
[Parallel(n_jobs=4)]: Done 4042 tasks  | elapsed: 54.3min
[Parallel(n_jobs=4)]: Done 4893 tasks  | elapsed: 84.6min
[Parallel(n_jobs=4)]: Done 5824 tasks  | elapsed: 101.4min
[Parallel(n_jobs=4)]: Done 6080 out of 6080 | elapsed: 105.9min finished

```

ภาพ 42 การค้นหาแบบกริด GridSearch

การปรับใช้ (Deployment)

การสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง หลังจากที่ได้แบบจำลองที่มีคุณภาพและความถูกต้องแม่นยำตามที่ต้องการแล้ว ซึ่งผลลัพธ์ที่ได้แสดงให้เห็นถึงคุณลักษณะของรีวิวบนตลาดซื้อขายสินค้าออนไลน์ในรูปแบบรีวิวจริงหรือรีวิวที่น่าสงสัย เพื่อเป็นข้อมูลสนับสนุนในการตัดสินใจซื้อของลูกค้า สามารถนำแบบจำลองไปใช้ในการพัฒนาแพลตฟอร์มเพื่อระบุรีวิวจริงหรือรีวิวที่น่าสงสัยได้

ปฏิบัติการ (Operations)

ต้องมีการตรวจสอบประสิทธิภาพของแบบจำลองตามข้อมูลที่เปลี่ยนไปในสภาวะการณ์การแข่งขันทางธุรกิจที่มีความรุนแรง ยังสามารถนำไปปรับใช้งานกับผลิตภัณฑ์สินค้า หรือบริการอื่น ๆ เพื่อให้ได้ข้อมูลในการสร้างแบบจำลองที่เป็นปัจจุบันและมีประสิทธิภาพ

การเพิ่มประสิทธิภาพ (Optimization)

เป็นขั้นตอนสุดท้าย โดยต้องมีการพัฒนาแบบจำลองแมชชีนเลิร์นนิง (Model Machine Learning) ในการผลิตเมื่อมีแหล่งข้อมูลใหม่เข้ามาหรือทำตามขั้นตอนที่จำเป็นสำหรับการอัปเดตประสิทธิภาพของโมเดลแมชชีนเลิร์นนิง



บทที่ 4 ผลการวิจัย

งานวิจัยนี้ได้สรุปผลการวิจัยการระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง โดยผลการวิจัยสอดคล้องตามวัตถุประสงค์ของงานวิจัย ดังนี้

4.1 ผลการวิเคราะห์รีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง

ผลการวิจัยตามวัตถุประสงค์ข้อที่ 1 เพื่อวิเคราะห์รีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง โดยการวิเคราะห์แบบพยากรณ์ (Predictive Analytics) สกัดข้อมูลที่เหมาะสมมาใช้ในการประมวลผล โดยการได้มาซึ่งข้อมูล (Data Acquisition) เก็บข้อมูลจากรีวิวผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ที่เป็นภาษาไทย ระหว่างเดือนมกราคม - พฤศจิกายน 2564 จำนวน 60,081 รีวิว จาก 10 ยี่ห้อ ซึ่งข้อมูลที่ได้ทั้งหมดอยู่ในรูปแบบชุดข้อความ หลังจากนั้นนำชุดข้อความมาเตรียมข้อมูล (Data Preparation) การนำข้อมูลรีวิวมาล้างกรองข้อมูล ทำความสะอาดข้อมูล (Data Cleaning) การรวมข้อมูล (Data Integration) ที่มีความซ้ำซ้อนกันไว้ในรูปแบบเดียวกัน และแปลงข้อมูล (Data Transformation) เพื่อเตรียมข้อมูลให้อยู่ในรูปแบบที่พร้อมนำไปใช้โดยการวิเคราะห์แบบพยากรณ์ (Predictive Analytics) สกัดข้อมูลที่เหมาะสมมาใช้ในการประมวลผล ดังนี้

ตาราง 1 ตารางแสดงการวิเคราะห์แบบพยากรณ์

ค่าคำตอบ	ค่าที่ทำนาย
รีวิวจริง	31,304
รีวิวที่น่าสงสัย	6,581
รีวิวที่ไม่มีข้อความ	22,195
ค่าที่ไม่เกี่ยวข้อง	1
รวม	60,081

จากตาราง 1 การวิเคราะห์แบบพยากรณ์ (Predictive Analytics) โดยนำข้อมูลรีวิวทั้งหมด 60,081 รีวิว มาวิเคราะห์พบว่ามีความค่าตอบ/ค่าที่ทำนาย เป็นรีวิวจริง (Real) 31,304 รีวิว รีวิวที่น่าสงสัย (Suspicious) 6,581 รีวิว รีวิวที่ไม่มีข้อความ (Missing value) 22,195 รีวิว และค่าที่ไม่เกี่ยวข้อง (Error) 1 รีวิว และหลังจากนั้นจะทำการลบ Missing value, Error ทำการลบรีวิวที่มีความซ้ำกัน เขียนเหมือนกัน โดยเก็บไว้แค่หนึ่งรีวิวเท่านั้น ซึ่งเหลือข้อมูลรีวิวที่ได้ทำการวิเคราะห์แบบพยากรณ์โดยเทคนิคการเรียนรู้ของเครื่องเพียง 34,123 รีวิว โดยค่าคำตอบ/ค่าที่ทำนาย จากข้อมูลรีวิวทั้งหมด 60,081 รีวิว ผู้วิจัยได้นำข้อมูลรีวิวมาวิเคราะห์ความรู้สึก (Sentiment Analysis) โดยแบ่งเป็น ความรู้สึกเชิงบวก (Positive) ความรู้สึกเชิงลบ (Negative) และความรู้สึกเป็นกลาง (Neutral) ดังนี้

ตาราง 2 ตารางแสดงการวิเคราะห์ความรู้สึก

ค่าคำตอบ	ความรู้สึกเชิงบวก (Positive)	ความรู้สึกเชิงลบ (Negative)	ความรู้สึกเป็นกลาง (Neutral)
รีวิวจริง	29,894	1,333	77
รีวิวที่น่าสงสัย	6,290	61	230
รวม	36,184	1,394	307

จากตาราง 2 การวิเคราะห์ความรู้สึก (Sentiment Analysis) โดยนำข้อมูลรีวิวทั้งหมด 60,081 รีวิว มาวิเคราะห์ความรู้สึก พบว่า **รีวิวจริง (Real)** จำนวน 31,304 รีวิว คิดเป็น 52.10% โดยแยกเป็นการวิเคราะห์ความรู้สึก ดังนี้ **รีวิวจริงแสดงความรู้สึกเชิงบวก (Positive)** 29,894 รีวิว คิดเป็น 95.50% โดยข้อมูลรีวิวมีลักษณะอธิบายถึง ความพึงพอใจคุณภาพของสินค้า ความคุ้มค่า ราคา ส่วนลด บรรจุกุภัณฑ์ การหีบห่อ ความรวดเร็วในการจัดส่ง การให้บริการจากร้านค้า ความถูกต้องของสินค้า มีของแถมให้ ตัวอย่างเช่น

“สินค้าคุณภาพดี ห่อแน่นหนามาก”

“สินค้า บริการ คุณภาพ ดี มากมาย”

“คุณภาพสินค้าดี ระยะการจัดส่งปานกลาง”

“สั่งครั้งแรกประทับใจเลยคะ ของแถมคือตะลึงมาก ให้เยอะมากๆคะ”

“สินค้าเป็นของแท้เหมือนที่ใช้อยู่เป็นประจำคะ ร้านแพ็คบับเบิ้ลและมีของแถม ใส่ใจ”

“จัดส่งสินค้าอย่างรวดเร็ว ได้ราคาตอนโปร299 คุ้มค่ามากๆ ช่วยกันแดดได้ดี ไม่อุดตัน จะมาซื้ออีก”

“หลอดที่สองแล้ว ที่ใช้มา ใช้ดีมาก ไม่อุดตันเหมาะสำหรับผิวมันเป็นสิ่ว ใช้แล้วหน้าไม่มันเลยยยยย ไปรททั้งวันตบแป้งรอบเดียวจ้า จัดส่งเร็วดีนะคะ“ เป็นต้น

รีวิวจริงแสดงความรู้สึกเชิงลบ (Negative) 1,333 รีวิว คิดเป็น 4.26% โดยข้อมูลรีวิวมี่ลักษณะอธิบายถึง คุณภาพของสินค้าไม่ดี ไม่สะอาด ค้างสต็อก บรรจุน้ำไม่ดี ไม่ได้รับของแถม ของแถมไม่มีคุณภาพ การเตรียมสินค้าในการจัดส่งใช้เวลานาน สินค้าที่ได้รับไม่ตรงกับสินค้าที่สั่งซื้อ ส่งของไม่ครบ การขนส่งล่าช้า สินค้าเสียหายจากการขนส่ง ไม่ประทับใจในการให้บริการของร้านค้าและขนส่ง ตัวอย่างเช่น

“ไม่มีของแถมเหมือนที่โฆษณาไว้”

“สินค้าดี แต่ขนส่งล่าช้า เอาของไปเก็บไว้ข้ามวัน”

“ร้านจัดส่งสินค้าไม่ครบ กดสั่งได้ ถึงเวลาไม่มีของส่ง”

“กล่องใส่สินค้าสภาพบอบมาก ดีที่ของข้างในไม่เป็นไร”

“คุณภาพของสินค้าดี ขอหัก1ดาวจัดส่งช้าไปหน่อยค่ะ”

“สูตรนี้ดีมากกกก ช่วยลดปัญหาสิว แต่ร้านส่งช้าไปหน่อย”

“โปรโมชัน ตัวอักษรเล็กมาก อ่านไม่เห็น ทำให้เข้าใจผิดว่าจะได้ของแถม”

“ส่งเกินเวลาตลอดต้องคอยกดขยายเวลา ร้านไม่แจ้งอะไรทั้งนั้น ต้องถามต้องตามเอง เคยสั่งแล้วไม่ส่งซึกที่จนถูกยกเลิก เย่มากๆ”

“สินค้าที่ได้รับสมบรูณ์ดีค่ะ ร้านค้าจัดส่งให้รวดเร็ว แพคเกจมาดีค่ะ การให้บริการร้านค้าดี แต่ผลิตภัณฑ์ตั้งแต่ปี2018ลื้ดเกี่ยวกับของหลายเดือนที่แล้วเลยคะเหมือนสินค้าค้างสต็อก” เป็นต้น

รีวิวจริงแสดงความรู้สึกเป็นกลาง (Neutral) ไม่ได้แสดงถึงความเชื่อมั่นที่เฉพาะเจาะจง 77 รีวิว คิดเป็น 0.24% โดยข้อมูลรีวิวมี่ลักษณะอธิบายถึงสินค้า บริการ การขนส่ง ตัวอย่างเช่น

“ยังไม่ได้ใช้”

“ยังไม่เห็นผลเท่าไร”

“อยากลองใช้ยังไม่เคยใช้”

“ลองสั่งมาใช้เห็นรีวิวเยอะ”

“วันหมดอายุตรงทั้งกล่องทั้งสินค้า” เป็นต้น

รีวิวที่น่าสงสัย (Suspicious) จำนวน 6,581 รีวิว คิดเป็น 10.95% **รีวิวที่น่าสงสัยแสดงความรู้สึกเชิงบวก (Positive) 6,290** รีวิว คิดเป็น 95.58% โดยข้อมูลตัวอย่างในรีวิวมี่ลักษณะมีการพิมพ์คำซ้ำเติม ๆ หลายครั้งในรีวิวเดียว ซึ่งมีประโยคเติม ๆ มากกว่า 3-4 ประโยคขึ้นไป หรืออาจใช้สัญลักษณ์แสดงความรู้สึกประกอบไปในรีวิว ตัวอย่างเช่น

“ดีมากๆๆดีมากๆๆดีมากๆๆดีมากๆๆ”

“ดีมากเลิศลดบ่อยๆๆดีมากเลิศลดบ่อยๆๆ”

“การจัดส่งเร็ว.....การจัดส่งเร็ว.....”

4.2 ผลการสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง

ผลการศึกษาตามวัตถุประสงค์ข้อที่ 2 เพื่อสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง ซึ่งหลังจากที่ได้ทำการวิเคราะห์แบบพยากรณ์ เหลือข้อมูลรีวิวที่ใช้สำหรับการ Train Model ที่นำเข้ามาจำนวน 34,123 รีวิว วัตจากค่าผลลัพธ์ของแบบจำลองทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิค ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เทคนิคต้นไม้ตัดสินใจ (Decision Tree) โดยจำแนกประเภทของข้อมูล (Classifier) คำนวณหาค่าประสิทธิภาพของการจำแนกประเภทข้อมูล ได้แก่ การวัดประสิทธิภาพและค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) ค่าประสิทธิภาพโดยรวม (F1-Score) คือ ค่าเฉลี่ยระหว่าง precision และ recall และ ค่าเฉลี่ยผลรวมค่าความแม่นยำ (Macro avg) เป็นการแสดงผลว่าแบบจำลองทั้ง 3 เทคนิค มีการทายถูกแบบสมดุลระหว่างทั้ง 2 กลุ่ม คือ รีวิวจริง (Real) กับ รีวิวที่น่าสงสัย (Suspicious) หรือไม่ หากทายถูกเพียงหนึ่งกลุ่ม ค่า Macro avg (F1-score) ไม่สูงมาก แต่หากทายถูกทั้ง 2 กลุ่ม ค่า Macro avg (F1-score) จะสูงซึ่งหากดูว่าแบบจำลองทั้ง 3 เทคนิค ทายกลุ่มใดเก่งจะดูตรงส่วนของรีวิวจริง (Real) (F1-score) และรีวิวที่น่าสงสัย (Suspicious) (F1-score) โดยค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (F1-score) ต้องสูงขึ้นด้วย ซึ่งผลลัพธ์และการเปรียบเทียบจากการทดลองการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่อง การระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้านตลาดซื้อขายสินค้าออนไลน์ ดังตารางที่ 2-5

ตาราง 3 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิคนาอิวเบย์ (Naives Bayes)

	ค่าความแม่นยำ Precision	ค่าความครบถ้วน Recall	ค่าประสิทธิภาพ โดยรวม F1
รีวิวจริง (Real)	90%	87%	88%
รีวิวที่น่าสงสัย (Suspicious)	46%	54%	49%
ค่าความถูกต้อง(Accuracy)	-	-	81%
ค่าเฉลี่ยของผลรวม (Macro avg)	68%	70%	69%

จากตาราง 3 ตารางแสดงผลลัพธ์การทดลองพบว่าการทำงานแบบจำลองด้วยเทคนิคนาอิวเบย์ (Naives Bayes) ค่าความถูกต้อง (Accuracy) มีค่าเท่ากับ 81% โดยรีวิวจริง (Real) มีค่าความแม่นยำ (Precision) 90% ค่าความครบถ้วน (Recall) 87% ค่าประสิทธิภาพโดยรวม (F1-Score) 88% และ

รีวิวน่าสงสัย (Suspicious) มีค่าความแม่นยำ (Precision) 46% ค่าความครบถ้วน (Recall) 54% ค่าประสิทธิภาพโดยรวม (F1-Score) 49% และค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (F1-score) 69%

ตาราง 4 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิค ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

	ค่าความแม่นยำ Precision	ค่าความครบถ้วน Recall	ค่าประสิทธิภาพ โดยรวม F1
รีวิวจริง (Real)	83%	99%	91%
รีวิวน่าสงสัย (Suspicious)	66%	6%	10%
ค่าความถูกต้อง(Accuracy)	-	-	83%
ค่าเฉลี่ยของผลรวม (Macro avg)	75%	53%	51%

จากตาราง 4 ตารางแสดงผลลัพธ์การทดลองพบว่าการทำนายของแบบจำลองด้วยเทคนิค ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ค่าความถูกต้อง (Accuracy) มีค่าเท่ากับ 83% โดยรีวิวจริง (Real) มีค่าความแม่นยำ (Precision) 83% ค่าความครบถ้วน (Recall) 99% ค่าประสิทธิภาพโดยรวม (F1-Score) 91% และรีวิวน่าสงสัย (Suspicious) มีค่าความแม่นยำ (Precision) 66% ค่าความครบถ้วน (Recall) 6% ค่าประสิทธิภาพโดยรวม (F1-Score) 10% และค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (F1-score) 51%

ตาราง 5 ตารางแสดงผลลัพธ์การจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree)

	ค่าความแม่นยำ Precision	ค่าความครบถ้วน Recall	ค่าประสิทธิภาพ โดยรวม F1
รีวิวจริง (Real)	86%	92%	89%
รีวิวน่าสงสัย (Suspicious)	42%	28%	34%
ค่าความถูกต้อง(Accuracy)	-	-	81%
ค่าเฉลี่ยของผลรวม (Macro avg)	64%	60%	61%

จากตาราง 5 ตารางแสดงผลลัพธ์การทดลองพบว่าการทำนายของแบบจำลองด้วยเทคนิคต้นไม้ตัดสินใจ (Decision Tree) ค่าความถูกต้อง (Accuracy) มีค่าเท่ากับ 83% โดยรีวิวจริง (Real) มีค่าความแม่นยำ (Precision) 86% ค่าความครบถ้วน (Recall) 92% ค่าค่าประสิทธิภาพโดยรวม (F1-Score) 89% และรีวิวที่น่าสงสัย (Suspicious) มีค่าความแม่นยำ (Precision) 42% ค่าความครบถ้วน (Recall) 28% ค่าประสิทธิภาพโดยรวม (F1-Score) 34% และค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (F1-score) 61%

ตาราง 6 ตารางแสดงผลลัพธ์การเปรียบเทียบการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคต้นไม้ตัดสินใจ (Decision Tree)

เทคนิค วิธีการเรียนรู้ ของเครื่อง	ร้อยละค่าการเปรียบเทียบการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่อง					
	ค่าความ ถูกต้อง (Accuracy)	ค่าความ แม่นยำ (Precision)	ค่าความ ครบถ้วน (Recall)	ค่า ประสิทธิภาพ โดยรวม (F1)	รีวิว จริง (Real)	รีวิว น่าสงสัย (Suspicious)
Naives Bayes	81%	68%	70%	69%	88%	49%
Support Vector Machine	83%	75%	53%	51%	91%	10%
Decision Tree	81%	64%	60%	61%	89%	34%

จากตาราง 6 ตารางแสดงผลลัพธ์การเปรียบเทียบจากการทดลองการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่องทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคต้นไม้ตัดสินใจ (Decision Tree) การทดลองพบว่าเทคนิค Naives Bayes กับ Decision Tree มีค่าความถูกต้อง (Accuracy) เท่ากัน คือ 81% แต่ค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (f1-score) ของเทคนิค Naives Bayes มีค่าอยู่ที่ 69% และเทคนิค Decision Tree มีค่า 61% สรุปได้ว่า เทคนิค Naives Bayes ดีกว่าเทคนิค Decision Tree และเมื่อนำเทคนิค Naives Bayes มาเปรียบเทียบกับ เทคนิค Support Vector Machine โดยมีค่าความถูกต้อง (Accuracy) เท่ากับ 83% ซึ่งมีมากกว่า เทคนิค Naives Bayes แต่ค่าประสิทธิภาพโดยรวมของผลรวมค่าความแม่นยำ (Macro avg) (f1-score) ของเทคนิค Support Vector Machine เท่ากับ 51% ซึ่งมีน้อยกว่า เทคนิค Naives Bayes สรุปได้ว่า เทคนิค Naives Bayes เป็นแบบจำลองที่ดีที่สุด เพราะวัดค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความครบถ้วน (Recall) ในการทำนายรีวิวจริง (Real) และรีวิวที่น่าสงสัย (Suspicious) ได้ถูกต้องเมื่อเปรียบเทียบกับจำนวนรีวิวทั้งหมดที่ปรากฏอยู่ในชุดข้อมูล

4.3 ผลการทดสอบประสิทธิภาพแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง

ผลการศึกษาตามวัตถุประสงค์ข้อที่ 3 เพื่อทดสอบประสิทธิภาพแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง โดยการค้นหาแบบกริด (Grid Search) เป็นการค้นหาแบบสุ่มด้วยการหาค่าพารามิเตอร์ที่กำหนดไว้ล่วงหน้าทุกชุด และประเมินประสิทธิภาพความแม่นยำของแบบจำลองแต่ละชุด โดยค่าทั้งหมดอยู่ในรูปของเมทริกซ์ (Matrix) พารามิเตอร์แต่ละชุดถูกนำมาพิจารณาและสังเกตความถูกต้อง โมเดลที่มีชุดพารามิเตอร์ที่ให้ความแม่นยำสูงที่สุดถือว่าดีที่สุด โดยแบ่งข้อมูลออกเป็น 2 ส่วน ได้แก่ ข้อมูลสำหรับสร้างแบบจำลอง (Training Data) และข้อมูลสำหรับทดสอบแบบจำลอง (Test Data) เพื่อทดสอบประสิทธิภาพของแบบจำลองกับชุดข้อมูล ซึ่งได้ Best parameters ออกมาและเลือกค่าความถูกต้อง (Accuracy) ที่เหมาะสม

ตาราง 7 ตารางแสดงผลลัพธ์การทดสอบประสิทธิภาพด้วยวิธีการเรียนรู้ของเครื่องทั้ง 3 เทคนิค ด้วยวิธี K-fold cross validation

เทคนิคการเรียนรู้ของเครื่อง	ค่าความถูกต้อง (Accuracy)
Naives Bayes	76%
Support Vector Machine	85%
Decision Tree	83%

จากตาราง 6 ตารางแสดงผลลัพธ์การทดสอบประสิทธิภาพด้วยวิธีการเรียนรู้ของเครื่องทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) มีค่าความถูกต้อง (Accuracy) เท่ากับ 76% เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) มีค่าความถูกต้อง (Accuracy) เท่ากับ 85% และเทคนิคต้นไม้ตัดสินใจ (Decision Tree) มีค่าความถูกต้อง (Accuracy) เท่ากับ 83% จากการทดลองพบว่า เทคนิค Support Vector Machine สามารถช่วยเพิ่มค่าความถูกต้อง (Accuracy) ที่ดีที่สุด

บทที่ 5

บทสรุป

งานวิจัยนี้ได้วิเคราะห์ สร้างแบบจำลองและทดสอบประสิทธิภาพแบบจำลองการระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง สามารถสรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะ ดังนี้

5.1 สรุปผลการวิจัย

การระบุรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง ผู้วิจัยสรุปผลและอภิปรายผลการวิจัยตามประเด็นที่ได้ศึกษา โดยเก็บข้อมูลจากผู้ที่ใช้ที่รีวิวผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์แห่งหนึ่ง โดยเป็นภาษาไทย ระหว่างเดือนมกราคม - พฤศจิกายน 2564 จำนวน 60,081 รีวิว จาก 10 ยี่ห้อ ซึ่งข้อมูลที่ได้ทั้งหมดอยู่ในรูปแบบชุดข้อความ โดยงานวิจัยนี้เลือกเก็บข้อมูล 8 คุณลักษณะ ได้แก่ ยี่ห้อสินค้า ราคา ชื่อร้านค้า ความเร็วในการตอบกลับ ชื่อลูกค้า ข้อมูลรีวิว วันเวลาที่รีวิว จำนวนดาว ซึ่งจาก 8 คุณลักษณะที่เก็บข้อมูลมาเพื่อใช้ในการประเมินความน่าเชื่อถือของรีวิว พบว่าข้อมูลที่เกี่ยวข้องทั้งหมดไม่มีความผิดปกติ ซึ่งข้อมูลรีวิวในผลิตภัณฑ์ดูแลผิวหน้าของลูกค้าบนตลาดซื้อขายสินค้าออนไลน์ทั้งหมดมีรีวิวที่เป็นข้อความแสดงความรู้สึกเชิงบวก เชิงลบ รีวิวที่เป็นสัญลักษณ์แสดงความรู้สึก รีวิวที่เป็นข้อความผสมสัญลักษณ์แสดงความรู้สึก รีวิวที่มีคำว่าง โดยเนื้อหาของการรีวิวเกี่ยวข้องกับ สินค้า คุณภาพสินค้า ความคุ้มค่า ราคา ของแถม การรับออเดอร์ การบรรจุหีบห่อ การให้บริการตอบคำถาม การให้บริการหลังการขาย การจัดส่งสินค้า ความรวดเร็วในการจัดส่ง การให้บริการของขนส่ง การซื้อสินค้าในช่วงโปรโมชั่น การตอบแชท พบว่าการรีวิวของลูกค้าจากชุดข้อมูลนี้ไม่ได้แสดงความรู้สึกถึงความพึงพอใจของสินค้าเท่านั้น แต่มีปัจจัยอื่น ๆ เข้ามาเกี่ยวข้องในการรวิวนี้อีกด้วย หลังจากนั้นนำชุดข้อความมาเตรียมข้อมูล (Data Preparation) โดยการวิเคราะห์แบบพยากรณ์ (Predictive Analytics) สกัดข้อมูลที่เหมาะสมมาใช้ในการประมวลผล การสร้างแบบจำลอง (Modelling) โดยพัฒนาแบบจำลองโดยใช้เทคนิควิธีการเรียนรู้ของเครื่อง (Machine Learning) ประกอบด้วย 3 เทคนิค ได้แก่ เทคนิคนาอิวส์เบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) เทคนิคต้นไม้ตัดสินใจ (Decision Tree) จากนั้นทำการทดสอบและเปรียบเทียบประสิทธิภาพของแบบจำลองผลการวิจัยพบว่าเทคนิคที่เหมาะสมสำหรับการสร้างแบบจำลองรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ คือ เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ซึ่งให้ค่าความถูกต้อง (Accuracy) มากที่สุดเท่ากับ 85%

5.2 อภิปรายผล

การสร้างแบบจำลองและการทดสอบประสิทธิภาพแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง โดยนำชุดข้อมูลมาสร้างแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง ทั้ง 3 เทคนิค ได้แก่ เทคนิคนาอิวเบย์ (Naives Bayes) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) และเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และแสดงผลลัพธ์การเปรียบเทียบจากการทดลองการจำแนกข้อมูลด้วยวิธีการเรียนรู้ของเครื่อง ซึ่งจากผลการทดลองพบว่าเทคนิคนาอิวเบย์ (Naives Bayes) ซึ่งให้ค่าความถูกต้อง (Accuracy) มากที่สุดเท่ากับ 81% เป็นแบบจำลองที่ดีที่สุด สอดคล้องกับงานวิจัยของ (จารี ทองคำ และธีรยุทธ คุณสุข, 2564) การวิเคราะห์หัวข้อภาษาไทยที่มีผลต่อการจำแนกประเภทความคิดเห็น มีวัตถุประสงค์เพื่อการวิเคราะห์หัวข้อภาษาไทยที่มีผลต่อการจำแนกประเภทความคิดเห็น ความคิดเห็นโดยส่วนใหญ่จะเป็นวลี ความคิดเห็นของลูกค้าสามารถคัดกรองตามชนิดคำที่ส่งผลกระทบต่อประสิทธิภาพของการจำแนกความคิดเห็น ดังนั้นวิเคราะห์การใช้คำประเภทต่างๆ จากความคิดเห็นของลูกค้าเกี่ยวกับร้านอาหารเป็นงานวิจัยที่ทำทนายงานวิจัยนี้ข้อมูลถูกเก็บรวบรวมจากเว็บไซต์ www.wongnai.com จำนวนทั้งหมด 4,487 ข้อความ และได้นำเอาประเภทคำ 7 ประเภท เพื่อทำการเปรียบเทียบ ได้แก่ คำนาม คำกริยา คำวิเศษณ์ คำนามและคำกริยา คำนามและคำวิเศษณ์ คำกริยาและคำวิเศษณ์ และคำนาม คำกริยาและคำวิเศษณ์ เทคนิคนาอิวเบย์ เทคนิคซัพพอร์ตเวกเตอร์แมชชีน เทคนิคซี 4.5 เทคนิคขั้นตอนวิธีการเพื่อนบ้านใกล้ที่สุด เทคนิคป่าสุ่ม และ เทคนิคเพอร์เซปตรอนหลายชั้น ได้ถูกนำมาใช้ในการสร้างแบบจำลองเพื่อการจำแนก ในการวัดประสิทธิภาพของแบบจำลอง 10-fold cross validation ได้ถูกนำมาใช้ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ และใช้ค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึกลับ มาทำการแสดงประสิทธิภาพของเทคนิคในการสร้างแบบจำลอง จากการทดลองพบว่า เทคนิคนาอิวเบย์ สามารถสร้างแบบจำลองที่มีประสิทธิภาพในการจำแนกสูงสุด เมื่อใช้คำนาม คำกริยา และคำวิเศษณ์ร่วมกัน โดยมีค่าความถูกต้องสูงถึง 89.08% ค่าความแม่นยำสูงถึง 89.12% และค่าความระลึกลับสูงถึง 89.10% หลังจากนั้นได้มีการทดสอบประสิทธิภาพแบบจำลองรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง โดยการค้นหาแบบกริด (Grid Search) พบว่า เทคนิคที่เหมาะสมสำหรับการสร้างแบบจำลองการระบุรีวิวน่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ คือ เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ซึ่งให้ค่าความถูกต้อง (Accuracy) มากที่สุดเท่ากับ 85% ซึ่งสอดคล้องกับงานวิจัยของ (นุชนาฏ ปิ่นเมือง และจารี ทองคำ, 2561) ได้ใช้หลักการ 10-fold cross validation ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูล ทดสอบ และวัดประสิทธิภาพการจำแนกของแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ

(Precision) และค่าความระลึก (Recall) เมื่อทำการทดสอบและวัดประสิทธิภาพของโมเดลพบว่ามีเทคนิค Naïve Bayes ให้ผลดีที่สุดในการจำแนก ข้อความความคิดเห็น โดยให้ค่าความถูกต้อง 93.88% ค่าความแม่นยำ 94.02% และค่าความระลึก 93.54% โดยทั้งสองส่วนมีความสอดคล้องในประเด็นการใช้เครื่องมือ 10-fold cross validation ในการประเมินประสิทธิภาพ แต่เนื่องจากใช้ชุดข้อมูลที่ไม่เหมือนกัน ทำให้ค่าความถูกต้องของแบบจำลองที่มีประสิทธิภาพในการจำแนกต่างกัน และยังสอดคล้องกับงานวิจัยของ (สุปัญญา อภิวงค์โสภณ, 2561) การตรวจจับข่าวปลอมบนเครือข่ายสังคมออนไลน์ทวิตเตอร์ด้วยวิธีการเรียนรู้ด้วยเครื่อง โดยใช้การเรียนรู้ด้วยเครื่องสามวิธีได้แก่ Naïve Bayes, Neural Network และ Support Vector Machine โดยเก็บข้อมูลจากหัวข้อข่าวที่เป็นภาษาไทย ระหว่างเดือนตุลาคมถึงพฤศจิกายน พ.ศ. 2560 ผลการวิจัยพบว่าทั้งสามวิธีสามารถตรวจจับข่าวปลอมในชุดข้อมูลได้อย่างถูกต้อง ร้อยละความถูกต้องของวิธี Naïve Bayes คือ 96.08 เปอร์เซ็นต์ Neural Network 99.89 เปอร์เซ็นต์และ Support Vector Machine 99.89 เปอร์เซ็นต์ นอกจากนี้ได้ทำการวิเคราะห์ข้อมูลข่าวปลอมและชี้ให้เห็นลักษณะของข่าวปลอมที่พบในชุดข้อมูล โดยมีความสอดคล้องในประเด็นการเก็บข้อมูลย้อนหลังและค่าความถูกต้องของแบบจำลองของวิธี Support Vector Machine ที่มีประสิทธิภาพในการจำแนกที่เหมือนกัน

5.3 ข้อเสนอแนะ

ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1. ลูกค้าสามารถอ่านรีวิวก่อนตัดสินใจในการซื้อผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ สามารถรู้ได้ว่ารีวิวนั้นเป็นรีวิวจริงหรือเป็นรีวิวที่น่าสงสัยได้
2. ผู้ประกอบการสามารถนำเทคนิควิธีการเรียนรู้ของเครื่องไปใช้ในการพัฒนาแพลตฟอร์มระบบรีวิวที่น่าสงสัยในผลิตภัณฑ์ดูแลผิวหน้าบนตลาดซื้อขายสินค้าออนไลน์ได้
3. ผู้ประกอบการสามารถประเมินสถานการณ์ของธุรกิจในปัจจุบันได้อย่างแม่นยำ โดยนำข้อมูลเชิงลึก (Insight) ที่เกี่ยวข้องกับการรีวิวของลูกค้าไปใช้ในการกำหนดกลยุทธ์และวางแผนธุรกิจในอนาคตได้

ข้อจำกัดและข้อเสนอแนะแนวทางวิจัยในอนาคต

1. เทคนิควิธีการเรียนรู้ของเครื่องมีข้อจำกัดในการตัดคำ การลดรูปคำ การรวมคำที่เก็บข้อมูลเป็นภาษาไทย ต้องเชื่อมโยงกับแหล่งข้อมูลภายนอก
2. การเพิ่มปริมาณข้อมูลในชุดข้อมูลเพื่อช่วยยืนยันประสิทธิภาพของแบบจำลองที่สร้างขึ้นด้วยเทคนิควิธีการเรียนรู้ของเครื่อง
3. การทดลองใช้เทคนิควิธีการเรียนรู้ของเครื่องด้วยการจำแนกประเภทข้อมูลในรูปแบบอื่น เพื่อเพิ่มประสิทธิภาพของแบบจำลอง

บรรณานุกรม

- โกเมศ อัมพวัน. (2561). การจำแนกประเภทข้อมูล. <https://staff.informatics.buu.ac.th/>
- จारी ทองคำ และธีรยุทธ คุณสุข. (2564). การวิเคราะห์วลีภาษาไทยที่มีผลต่อการจำแนกประเภทความคิดเห็น. วารสารเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏอุบลราชธานี, 11(1).
<https://ph01-ohno.tci-thaijo.org/index.php/jitubru/about>
- เฉลิมชัย พิเดช. (2563). ระบบจำแนกตัวงานอัจฉริยะด้วยเทคนิควิเคราะห์ข้อความไทย-อังกฤษ และการเรียนรู้ของเครื่อง คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ]. <http://ir-thesis.swu.ac.th/dspace/handle/123456789/1164>.
- ณัฐพร เหล่าสีพงษ์. (2562). กลยุทธ์การจัดการทางธุรกิจของผลิตภัณฑ์อินทรีย์ (organic) ประเภทเครื่องสำอางและ
- ชนวัฒน์ มาลาบุปผา. (2564). ภาพรวมวงการ E-commerce ของปี 2021.
<https://www.thumbsup.in.th/ecommerce-2021>
- นกยูง. (2564). สัดส่วน E-commerce ในประเทศไทย 256-2564.
<https://www.marketingoops.com/reports/research/value-of-e-commerce-survey-in-thailand-2021/>
- นงเยาว์ สอนจะโปะ. (2561). รูปแบบการจำแนกกลุ่มข้อความภาษาไทยแบบอัตโนมัติโดยใช้การเรียนรู้ของเครื่อง (Machine Learning) ด้วยเทคนิค Unsupervised Learning ร่วมกับการประมวลผลภาษาธรรมชาติ (Natural Language Processing). วารสารวิชาการศรีปทุม ชลบุรี, 14 (4).
- นุชนาฏ ปิ่นเมือง และจारी ทองคำ. (2561). การจำแนกความคิดเห็นของคนไทยเกี่ยวกับสื่อออนไลน์โดยใช้การทำเหมืองข้อความ. *J Sci Technol MSU, Vol 37(3)*, 372-379.
- บศักดิ์ ลีวัลักษณ์. (2564). บทบาทของ E-Marketplace. Retrieved 16 March 2021 from <https://www.krungsri.com/th/plearn-plearn/marketplace-you-should-present-your-goods-now>.
- ปรีดี นกุลสมปรารถนา. (2563). *Customer Touchpoint* มีอะไรบ้าง.
<https://www.popticles.com/marketing/get-to-know-customer-touchpoints/>
- เปรมกมล หงส์ยนต์. (2562). ปัจจัยที่มีอิทธิพลต่อการตัดสินใจซื้อสินค้าผ่านทางแอปพลิเคชันออนไลน์ (ลาซาด้า). Retrieved 25 March 2021 from <https://e-research.siam.edu/wp-content/>
- พิชญะ พรหมลา. (2563). การเปรียบเทียบประสิทธิภาพการวิเคราะห์ความพึงพอใจเกี่ยวกับการจัดการเรียนการสอนด้วยกระบวนการวิเคราะห์ความรู้สึกโดยใช้เทคนิคการรวมกลุ่มเพื่อจำแนกข้อมูล. วารสารวิจัย มข. (ฉบับบัณฑิตศึกษา) 20(4), 140-149.

- แพรวไพลิน มณีชิตย์. (2563). อิทธิพลของรีวิวออนไลน์ที่ส่งผลต่อพฤติกรรมการวางแผนท่องเที่ยวของนักท่องเที่ยวชาวไทยกลุ่มเจนเนอเรชันวาย. วารสารวิทยาลัยดุสิตธานี, 14(2).
- ภาณุวัฒน์ เมฆะ. (2564). เทคนิคเหมืองข้อมูล. <https://erp.mju.ac.th/acticleDetail.aspx?qid=424>
- ภาวธ พงษ์วิทย์ภานุ. (2564). ธุรกิจ e-Commerce ไทยอยู่ในมือของ e-Marketplace ต่างชาติ. <https://www.prachachat.net/columns/news-589871>.
- รวีศ หาญอุตสาหะ. (2563). การตัดสินใจเลือกซื้อสินค้า. <https://www.krungsri.com/th/plearn-plearn/customers-decide-to-buy>
- รวีสดา เทศเมือง และ นิเวศ จิระวิจิตชัย. (2560). การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์โดยใช้ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมทซ์. *Engineering Journal of Siam University*, 18(34)(1), 1-12.
- รัชกฤต อารีราษฎร์. (2561). การจำแนกข้อความเข้าข่ายหมิ่นประมาทบนสื่อสังคมออนไลน์ จุฬาลงกรณ์มหาวิทยาลัย. <https://www.car.chula.ac.th/display7.php?bib=b2256998>
- ไลฟ์ สไตล์อิชชู. (2019). ประเภทของสกินแคร์ (Skincare) แบ่งตามเนื้อผลิตภัณฑ์. <https://www.lifestyleissue.com/beauty/skincare-types/>
- วารสารกรมพัฒนาธุรกิจการค้า. (2562). ขายของออนไลน์ทำไมต้องรีวิว. วารสารกรมพัฒนาธุรกิจการค้า. <http://magazine.dbd.go.th/magazine>
- ศูนย์วิจัยกสิกรไทย. (2563). เมื่อโควิด -19 เขย่าตลาดเครื่องสำอาง. https://www.kasikornresearch.com/th/analysis/k-social-media/Pages/cosmetic_market.aspx
- สกินแคร์ ในเขตกรุงเทพมหานคร : มุมมองจากผู้บริโภค <https://archive.cm.mahidol.ac.th/bitstream/123456789/3353/1/TP%20MS.054%202562.pdf>
- สมศักดิ์ ศรีสุวรรณ และสมัย ศรีสวย. (2563). การวิเคราะห์เหมืองความคิดเห็นโดยใช้เทคนิคการสกัดคำ. *Journal of Applied Information Technology*, ปีที่ 6(2), 95-104.
- สลิลาทิพย์ ทิพย์ไกรสร. (2561). เส้นทางของผู้บริโภคในยุคดิจิทัล โจทย์ใหม่ทางการตลาด. วารสารปัญญาภิวัตน์, 10(2), 294-302.
- สำนักข่าวอินโฟเควสท์. (2021). อังกฤษสั่งตรวจสอบแอมะซอน-กูเกิล ปมรีวิวปลอมระบาด. <https://www.infoquest.co.th/2021/100589>
- สินินาฏ เพิ่มสวัสดิ์. (2564). Marketplace คืออะไร. <https://www.tibdglobal.com/marketplac>
- สุปัญญา อภิวงษ์โสภณ. (2561). การตรวจสอบข่าวปลอมด้วยวิธีการเรียนรู้ด้วยเครื่อง จุฬาลงกรณ์มหาวิทยาลัย. <https://www.cp.eng.chula.ac.th/~prabhas/thesis/supanya-final-5771425821.pdf>
- อชิบัติ คำเจริญ และสุมาลย์ ปานคำ. (2563). รูปแบบความสัมพันธ์เชิงสาเหตุความตั้งใจซื้อครีมนิยมกซ์ สกินแคร์ ผ่านเพจเฟซบุ๊ก ของผู้บริโภคในเขตกรุงเทพมหานครและปริมณฑล. *Journal of Arts Management*, 4(2).

- Chengz. (2019). วิเคราะห์ความรู้สึกด้วย *sentiment analysis*.
<https://medium.com/@cheng3374>
- comets. (2020). แนวโน้มผลิตภัณฑ์สกินแคร์ปี 2021. <https://www.cometsintertrade.com/>
- Davide Proserpio et al. (2020). The Market for Fake Reviews.
https://www.researchgate.net/publication/343807868_The_Market_for_Fake_Reviews
- Digital Consult. (2020). การวิเคราะห์ความรู้สึก (*Sentiment Analysis*)
www.jibdigitalconsult.com
- FEED NEWS. (2021). การเขียนรีวิวปลอมบน AMAZON เพื่อแลกเงินและสินค้า.
<https://techfeedthai.com/2021/>
- G-Able. (2022). ขั้นตอนการทำเหมืองข้อมูล (*Data Mining*) <https://www.g-able.com/digital-review/digital-transformation/big-data-analytics/>
- google. (2021). วิธีลบรีวิวเชิงลบปลอมบน Google <https://support.google.com/business/>
- Kaewsiri Srisam-ang. (2020). ประเภทของผลิตภัณฑ์ดูแลผิวหน้า
<https://www.vogue.co.th/beauty/skincarerroutinehow>
- Lifestyleissue. (2021). ผลิตภัณฑ์ดูแลผิวหน้า.
<https://www.lifestyleissue.com/beauty/skincare-types/>
- M.Pussapol. (2020). Facebook ใช้ *Machine Learning* ตรวจสอบบัญชีปลอมอย่างไร.
<https://www.brandage.com/article/17474/facebook>
- mycloudfulfillment. (2020). จุดเด่น ข้อดี ที่ช่วยตอบโจทย์ให้ผู้ซื้อ-ผู้ขาย
<https://www.mycloudfulfillment.com/blog/>
- Natratanon Kanraweekultana. (2019). *Natural Language Processing (NLP)*.
<https://medium.com/@natratanonkanraweekultana/natural-language-processing-nlp-40af75482b62>
- Nessessence. (2018). กระบวนการของ *Machine Learning*.
<https://www.thaiprogrammer.org/>
- nsgcosmetics. (2021). อุตสาหกรรมผลิตภัณฑ์บำรุงผิว. <https://www.nsg.shopping/th/blog>
- Nutmos. (2018). Google ใช้ *machine learning* ร่วมกับมนุษย์ กำจัดรีวิวขยะบน *Play Store*
<https://www.blognone.com/node/>
- pigabyte. (2021). เปิดเทรนด์ *E-Commerce* 2021.
<https://www.marketingoops.com/news/ecommerce/ctc2021-ecommerce-trends/>
- quickservice. (2019). การรีวิว. <https://www.quickservice.co.th/review/>
- SAS Institute Inc. (2020). การวิเคราะห์เหมืองข้อมูล.
https://www.sas.com/th_th/insights/analytics/data-mining.html#dmtechnical
- SAS Institute Inc. (2021). การเรียนรู้ของเครื่อง. Retrieved 15 February 2021 from
<https://www.sas.com/>

- STEPS Academy. (2021). การเติบโตของตลาดที่ใช้งานเทคโนโลยี NLP ทั่วโลก.
<https://stepstraining.co/foundation/natural-language-processing>
- SafetyDetectives. (2021). *Amazon Fake Reviews Scam* เปิดเผยในการละเมิดข้อมูล.
<https://www.safetydetectives.com/blog/amazon-reviews-leak-report/>.
- Sarunjade. (2020). บทบาทของรีวิวจีนค้ากับการทำการตลาดที่เปลี่ยนไป.
<https://www.marketingoops.com/exclusive/online-review-and-the-new-marketing/>
- SMEONE. (2020). ขายของออนไลน์ทำไมต้องมีรีวิว. <https://www.smeone.info/posts/view/208>
- SME Update. (2021). ความคิดเห็นลูกค้ามีค่าอย่างไรต่อธุรกิจ.
<https://www.bangkokbanksme.com/en/comments-value-for-your-business>
- Techsauce. (2020). มูลค่าตลาด *E-commerce* ปี 2020. <https://techsauce.co/tech-and-biz/thailand-e-commerce-landscape-covid-19>
- Tripadvisor. (2020). รีวิวปลอม. <https://tripadvisor.mediaroom.com/TH-content-integrity-policy>
- Tirasak W.Warm. (2021). *TouchPoint* ที่ดีไม่ใช่แค่นี้ แต่ต้องตอบใจത്യ.
<https://www.everydaymarketing.co/>



ภาคผนวก

มหาวิทยาลัยนครพนม