บทที่ 2
ปริทัศน์วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

งานวิจัย เรื่อง การพัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ ได้ศึกษาแนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ดังนี้

2.1 ธุรกิจอาหารและเครื่องดื่ม

2.2 เครือข่ายสังคมออนไลน์ (Online Social Media)

2.3 บทวิจารณ์ออนไลน์ (Online Reviews)

2.4 การประมวลผลทางภาษา (Word segmentation)

2.5 การประมวลผลภาพดิจิทัล (Digital Image Processing)

2.6 โยโล่ (You Only Look Once: YOLO)

2.6 การเรียนรู้ของเครื่อง (Machine Learning)

2.7 การเรียนรู้เชิงลึก (Artificial Neural Network)

2.8 งานวิจัยที่เกี่ยวข้อง

2.9 กรอบแนวคิดการวิจัย

โดยจากการศึกษาแนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง มีรายละเอียดดังนี้

2.1 ธุรกิจอาหารและเครื่องดื่ม

2.1.1 ความหมายของธุรกิจอาหารและเครื่องดื่ม

ธุรกิจอาหารและเครื่องดื่ม หมายถึง กิจการที่ให้บริการอาหารและเครื่องดื่มแก่ทุกคน โดยมีสถานที่ให้บริการสำหรับผู้บริโภค หรือมีการบรรจุอาหารที่ผู้บริโภคไม่จำเป็นต้องนั่งรับประทานอาหารในสถานที่นั้น ๆ (Tourism Industry DPU, 2553)

2.1.2 ความเป็นมาของธุรกิจร้านอาหารและเครื่องดื่ม

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

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

2) ยุคกลาง ในยุคนี้ มีการบริการอาหารที่มีวิธีการปรุงแบบฝรั่งเศส ที่ผู้บริโภคสามารถรับประทานอาหารตามที่ต้องการได้โดยไม่ขึ้นอยู่กับเวลา และมีการปรับเปลี่ยนวิธีการให้บริการแก่นักท่องเที่ยว โดยให้บริการอาหารและเครื่องดื่มที่อยู่ในรายการของสถานที่นั้น ๆ ทำให้ มีร้านอาหารฝรั่งเศสที่เป็นที่รู้จักแก่ผู้คนเป็นจำนวนมาก นำไปสู่การตั้งชื่อร้านอาหารเป็นภาษาฝรั่งเศสในประเทศอังกฤษ

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

2.1.3 ประเภทของธุรกิจอาหารและเครื่องดื่ม

ประเภทของธุรกิจร้านอาหารและเครื่องดื่มออกในปัจจุบันออกได้ถูกแบ่งออกเป็น 7 ประเภท โดย นิศา ชัชกุล (2550) ดังนี้

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

2) ธุรกิจอาหารสำเร็จรูป (Deli Shops) เป็นอาหารประเภทแช่แข็ง เนย แซนวิช สลัด เป็นต้น ซึ่งส่วนใหญ่ให้บริการในบริเวณที่มีผู้คนพลุกพล่าน มีระยะเวลาในการเปิดและโต๊ะนั่งที่จำกัด

3) ธุรกิจอาหารบุฟเฟต์ (Buffets) เป็นธุรกิจที่ได้รับความนิยมในสถานที่พักที่เป็นโรงแรม ธุรกิจประเภทนี้ให้ผู้บริโภคสามารถเลือก เละตักอาหารได้ตามปริมาณและความต้องการ โดยคิดค่าบริการเป็นจำนวนคน

4) ธุรกิจประเภทคอฟฟี่ช็อป (Coffee Shops) เป็นธุรกิจที่มักตั้งอยู่ภายในห้างสรรพสินค้า หรือชุมชนที่ผู้คนผ่านไปผ่านมาเป็นจำนวนมาก ราคาถูก ให้บริการแก่ผู้บริโภคที่ต้องการความรวดเร็วในการรับประทานอาหาร

2.1.3.5 ธุรกิจอาหารเฉพาะกลุ่มเชื้อชาติ (Ethic Restaurants) เป็นธุรกิจอาหารของเต่ละเชื่อชาติ ตั้งอยู่ในท้องถิ่นของเชื่อชาตินั้น ๆ มีจุดเด่นที่การให้บริการที่เป็นเอกลักษณ์ของชาตินั้น ๆ ทั้งอาหาร บรรยากาศภายในร้าน และพนักงานที่ให้บริการ

2.1.3.6 ธุรกิจประเภทคาเฟทีเรีย (Cafeterias Self Services) เป็นธุรกิจที่รองรับผู้บริโภคได้จำนวนมากในระเวลาอันสั้น โดยผู้บริโภคต้องบริการตนเอง ไม่มีบริกรสำหรับให้บริการแก่ผู้บริโภค

2.1.3.7 ธุรกิจอาหารกูร์เมต์ (Gourmet Restaurants) เป็นธุรกิจที่ดึงดูดผู้บริโภคด้วยสถานที่ที่มีความหรูหรา มีเป้าหมายเป็นกลุ่มผู้บริโภคที่มีฐานะดี ให้บริการที่ดีกว่าธุรกิจประเภทอื่น ๆ แต่มีราคาแพง

จากความเป็นมาและประเภทของธุรกิจร้านอาหารและเครื่องดื่ม แสดงให้เห็นถึงการพัฒนาด้านธุรกิจร้านอาหารและเครื่องดื่มที่ปรับเปลี่ยนให้สอดคล้องกับพฤติกรรมของผู้บริโภคในแต่ละยุคสมัย สร้างความหลากหลาย และเป็นที่นิยมแก่ผู้บริโภคเป็นจำนวนมาก งานวิจัยนี้จึงพิจารณาธุรกิจอาหารและเครื่องดื่มที่ให้บริการแก่ผู้บริโภคที่ไม่เฉพาะเจาะจง เนื่องจากผู้บริโภคทุกคนสามารถเข้าถึง ได้รับบริการ และสามารถแบ่งปันข้อมูลด้านอาหารหรือร้านอาหารผ่านเครือข่ายสังคมออนไลน์ได้

2.2 สื่อสังคมออนไลน์

2.2.1 ความหมายของสื่อสังคมออนไลน์ (Social Media)

ราชบัณฑิตยสถาน (2554) ได้บัญญัติคำว่า สื่อสังคม ว่า “สื่อชนิดหนึ่งที่ผู้ใช้สามารถมีส่วนร่วม สร้าง และแลกเปลี่ยนความคิดเห็นต่าง ๆ ผ่านระบบอินเทอร์เน็ตได้ ในทางเทคนิค สื่อสังคม จะหมายถึง โปรแกรมกลุ่มหนึ่งที่ทำงานโดยใช้พื้นฐานและเทคโนโลยีของเว็บตั้งแต่รุ่น 2.0 เช่น บีโบ มายสเปซ เฟซบุ๊ก ทวิตเตอร์ วิกิพีเดีย ไฮไฟฟ์ และบล็อกต่าง ๆ ในทางธุรกิจสื่อสังคมคือสื่อที่ผู้บริโภคสร้างขึ้น (Consumer-Generated Media หรือ CGM) สำหรับกลุ่มบุคคลที่ติดต่อสื่อสารกันผ่านสื่อสังคม เพื่อแลกเปลี่ยนข้อมูลข่าวสาร และทำกิจกรรมที่สนใจร่วมกัน”

พิชิต วิจิตรบุญยรักษ์ (2554) ได้ให้ความหมายของสื่อสังคมออนไลน์ว่า เป็นสื่อที่ผู้ส่งสารแบ่งปันสารในรูปแบบต่าง ๆ ไปยังผู้รับสารผ่านเครือข่ายสังคมออนไลน์ โดยสามารถโต้ตอบกันระหว่างผู้ส่งสารและผู้รับสาร หรือผู้รับสารด้วยกันเอง

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

จากความหมายของเครือข่ายสังคมออนไลน์ สามารถสรุปได้ว่า สื่อสังคมออนไลน์ หมายถึง สังคมที่มีการรวมตัวกันเพื่อสร้างความสัมพันธ์บนอินเทอร์เน็ตหรือที่เรียกว่า ชุมชนออนไลน์ (Community Online) มีลักษณะเป็นสังคมเสมือน (Virtual Community) ที่ผู้คนสามารถทำความรู้จัก แลกเปลี่ยนความคิด แบ่งปันประสบการณ์ และเชื่อมโยงกันในทิศทางใดทิศทางหนึ่งผ่านการติดต่อสื่อสารกันเป็นเครือข่าย โดยงานวิจัยนี้ได้พัฒนาระบบแนะนำร้านอาหารจากบทวิจารณ์และภาพอาหารบนเครือข่ายสังคมออนไลน์ และพิจารณาข้อมูลการใช้งานเครือข่ายสังคมออนไลน์ของผู้ใช้ เพื่อให้สามารถเชื่อมโยงข้อมูลของผู้ใช้บนเครือข่ายสังคมออนไลน์และสามารถแนะนำร้านอาหารที่ตรงกับความต้องการของผู้ใช้ได้

2.2.2 ประเภทของเครือข่ายสังคมออนไลน์

เขมณัฏฐ์ ศิริธรรม (2556) ได้แบ่งประเภทของเครือข่ายสังคมออนไลน์ที่ผู้ใช้สามารถแสดงความคิดเห็น หรือแลกเปลี่ยนข้อมูลข่าวสารได้ ออกเป็น 5 ประเภท ได้แก่ บล็อก (Blog) หรือเว็บบล็อก ทวิตเตอร์ (Twitter) เครือข่ายสังคม (Social Network) สื่อแบ่งปัน (Media Sharing) และออนไลน์ฟอรั่ม (Online Forum)โดยประเภทของเครือข่ายสังคมออนไลน์ตามแนวคิดของ เขมณัฏฐ์ ศิริธรรม (2556) มีดังนี้

1) บล็อก (Blog) หรือเว็บบล็อก (Web Blog) เป็นระบบจัดการเนื้อหาที่ให้ผู้ใช้สามารถเขียนเนื้อหาเอง สามารถเผยแพร่แง่มุมของการตลาด ถูกนำมาใช้ในการสื่อสารการตลาด 2 รูปแบบ คือ สร้างตราสินค้าเพื่อสื่อสารกับผู้บริโภคและบล็อกที่เขียนจากผู้เขียนบล็อก (Blogger) อิสระที่มีผู้ติดตามจำนวนมากจนกลายเป็นผู้มีอิทธิพลทางการตลาด (Marketing Influencer)

2) ทวิตเตอร์ (Twitter) เป็นไมโครบล็อก (Micro Blog) ที่จำกัดตัวอักษรในการโพสต์ครั้งละ 140 ตัวอักษร ถูกนำไปใช้ทางธุรกิจ ทั้งเพื่อสร้างการบอกต่อ การเพิ่มยอดขาย การสร้างแบรนด์ หรือเป็นเครื่องมือในการบริหารความสัมพันธ์กับลูกค้า รวมถึงใช้เป็นเครื่องมือในการประชาสัมพันธ์

3) เครือข่ายสังคม (Social Network) หมายถึง เครือข่ายสังคมที่ถูกเชื่อมโยงในโลกออนไลน์ จนกลายเป็นสังคมเสมือน ซึ่งผู้ใช้สามารถสร้างตัวตนขึ้นเอง โดยมีข้อมูลส่วนตัว รูป การจดบันทึก หรือการใส่วิดีโอ เป็นต้น สามารถเพิ่มจำนวนเพื่อนได้ ทำให้นักการตลาดนำเครือข่ายสังคมมาใช้ในการสร้างปฏิสัมพันธ์กับลูกค้า รวมถึงจัดกลุ่มมความสนใจของคนในวงสังคมได้

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

5) ข่าวสังคมและบุ๊คมาร์ก (Social News and Bookmarking) เป็นเว็บไซต์ที่ประชาชนทั่วไปมีโอกาสให้คะแนนและโหวตเว็บไซต์ที่ชื่นชอบ ซึ่งสามารถเก็บไว้เป็นบุ๊คมาร์กได้ โดยนักสื่อสารการตลาดสามารถใช้เครื่องมือนี้ในการบอกต่อ และสร้างจำนวนคนที่เข้ามายังเว็บไซต์

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

จากประเภทของสื่อสังคมดังที่กล่าวมาข้องต้น งานวิจัยนี้ได้ใช้สื่อสังคมประเภทเครือข่ายสังคมออนไลน์ โดยพิจารณาข้อมูลของผู้ใช้จากการใช้งานสื่อสังคมออนไลน์ทวิตเทอร์ ประกอบด้วยบทวิจารณ์ และภาพเกี่ยวกับอาหาร และร้านอาหาร

2.3 บทวิจารณ์ออนไลน์

2.3.1 ความหมายของบทวิจารณ์ออนไลน์

บทวิจารณ์ออนไลน์ หมายถึง ความคิดเห็นส่วนตัวของผู้เขียนในเรื่องที่ตนมีประสบการณ์พบเจอมาซึ่งอาจแตกต่างจากผู้อื่นได้ตามความชื่นชอบส่วนบุคคล (บุณยาพร วุฒิธรรมคุณ, 2559)

2.3.2 บทวิจารณ์ออนไลน์บนเครือข่ายสังคมออนไลน์

ในการเขียนบทวิจารณ์ออนไลน์สามารถแบ่งได้ตามประเภทของเครือข่ายสังคมออนไลน์ ดังนี้ (บุณยาพร วุฒิธรรมคุณ, 2559)

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

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

3) เครือข่ายสังคม (Social Network) เป็นอีกหนึ่งช่องทางในการค้นหาบทวิจารณ์ร้านอาหาร และได้รับความนิยมเป็นอย่างมากในปัจจุบัน ผู้ใช้สามารถติดตามอ่านรีวิวร้านอาหารได้โดยการเป็นเพื่อนหรือกดถูกใจ

4) สื่อแบ่งปัน (Media Sharing) ได้รับความนิยมจากผู้ใช้ทั่วโลก ให้บริการทั้งในรูปแบบเว็บไซต์และแอพพลิเคชั่น โดยผู้ใช้สามารถอัปโหลดภาพวิดีโอ เปิดดูภาพวิดีโอ และแบ่งภาพวิดีโอให้ผู้อื่นได้โดยมีเสียค่าใช้จ่ายใด ๆ สำหรับผู้ที่ค้นหาการบทวิจารณ์ร้านอาหารนั้นสามารถทำได้โดยการพิมพ์คำสำคัญในช่องค้นหา

5) ข่าวสังคมและบุ๊คมาร์ก (Social News and Bookmarking) เป็นเว็บไซต์ที่รวบรวมและเชื่อมโยงบทความหรือเนื้อหาในอินเทอร์เน็ต และจัดกลุ่มเป็นหมวดหมู่ โดยผู้ใช้เป็นผู้ส่งและมีการให้คะแนนจากผู้ใช้คนอื่น ๆ เพื่อร่วมกันตัดสินใจว่าเนื้อหาใดเป็นที่น่าสนใจมากที่สุด

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

งานวิจัยนี้รวบรวมบทวิจารณ์จากทวิตเตอร์ เนื่องจากเป็นบทวิจารณ์ที่มีความยาวไม่มาก เพราะมีการจำกัดขนาดของข้อความ ทำให้ได้ข้อความที่กระชับ ซึ่งช่วยลดเวลาในการประมวลผลคำ และมีองค์ประกอบทั้งข้อความและภาพ ซึ่งสามารถรวบรวมบทวิจารณ์ทั้งข้อความและภาพมาใช้ในการวิเคราะห์ข้อมูลได้ รวมถึงในข้อความที่ผู้ใช้เผยแพร่ผ่านทวิตเตอร์ และการแบ่งปันต่อ ๆ กัน ทำให้สามาถนำมาวิเคราะห์ข้อมูลที่เกี่ยวข้อง ได้แก่ การแบ่งประเภทบทวิจารณ์ว่าเป็นด้านบวก ลบ หรือเป็นกลาง

2.4 การประมวลผลทางภาษา (Word Segmentation)

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

2.4.1 อักขรวิธีของภาษาไทย

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

ตารางที่ 2.1 ตัวอย่างการสะกดคำตามอักขรวิธี

องค์ประกอบ	ตัวอย่างของคำ
พยัญชนะ + สระ + วรรณยุกต์	หมู กา หมี
พยัญชนะ + สระ + วรรณยุกต์ + ตัวสะกด	ลิง ยุง แมว
พยัญชนะ + สระ + วรรณยุกต์ + ตัวสะกด + การันต์	สิงห์ หงส์ สังข์

2.4.2 วิธีการตัดคำ

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

1) การตัดคำโดยใช้กฎ (Rule-based) เป็นการตัดคำโดยการตรวจสอบกฎเกณฑ์ทางอักขระวิธีที่มีการประสมอักษร การเว้นวรรค และย่อหน้า เพื่อใช้กำหนดขอบเขตของคำ

2) การตัดคำโดยใช้พจนานุกรม (Dictionary) วิธีนี้เป็นการอ้างอิงด้วยพจนานุกรมโดยใช้สายอักขระมาเปรียบเทียบกับคำที่อยู่ในพจนานุกรม ซึ่งจะต้องมีการจัดเก็บคำไว้ในพจนานุกรม เป็นวิธีที่ใช้เวลานานในการประมวลผล

3) การตัดคำโดยใช้คลังคำ (Corpus) เป็นการตัดคำโดยใช้วิธีทางสถิติในการประมวลผคำ โดยใช้คลังคำที่เก็บไว้เป็นฐานความรู้ ใช้การนับความถี่ซึ่งประกอบด้วยการนับความถี่โดยใช้ความน่าจะเป็น (Probabilistic Word Segmentation) เป็นวิธีที่ใช้แบบจำลองเอ็นแกรม (N-gram) ในการหารูปแบบในการตัดคำ และหาลำดับของการตัดคำที่เป็นไปได้มากที่สุด โดยวิธีนี้เป็นการใช้คลังข้อมูลที่มีการตัดคำ และกำหนดหมวดคำไว้ก่อนแล้ว และการตัดคำโดยพิจารณาถึงคุณลักษณะของคำ (Feature-based Word Segmentation) เป็นวิธีในการลดข้อผิดพลาดในการตัดคำซึ่งใช้ความน่าจะของการจำกัดหมวดคำเป็นแบบจำลองสำหรับตัดคำ

2.4.3 เทคนิคการตัดคำ

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

1) เทคนิคการเทียบคำที่ยาวที่สุด (Longest word pattern matching) เป็นเทคนิคตรวจสอบสายอักขระจากซ้ายไปขวา แล้วนำไปเปรียบเทียบกับคำในพจนานุกรม หากพบว่ามีมากกว่า 1 พยางค์ จะเลือกพยางค์ที่ยาวที่สุด แล้วตัดคำต่อไปจนจบสายอักขระ

2) เทคนิคการเทียบคำที่สั้นที่สุด (Shortest word pattern matching) เป็นเทคนิคในการเลือกคำที่สั้นที่สุดที่พบก่อน แต่หากมีจำนวนที่มากที่สุดความถูกต้องในการตัดคำจะน้อยกว่าเมื่อเทียบกับการใช้เทคนิคการเทียบคำที่ยาวที่สุด

3) เทคนิคการใช้ความถี่ของคำ (Word usage frequency) เทคนิคนี้เป็นเทคนิคที่ช่วยแก้ปัญหาเรื่องคำกำกวมในประโยค โดยใช้การวิเคราะห์ความถี่ของคำที่ใช้ในชีวิตประจำวัน โดยเรียงคำในพจนานุกรมตามความถี่ของคำที่พบในการตัดคำ

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

2.4.4 ขั้นตอนวิธีในการตัดคำ

ในงานวิจัยนี้จะกล่าวถึงขั้นสตอนวิธีที่ใช้ในการตัดคำสพหรับการจำแนกประเภทบทวิจาณ์ ซึ่งใช้ขั้นตอนวิธีของ Python ที่ได้พัฒนาขั้นตอนวิธีในการตัดคำเรียกว่า PyThaiNLP

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

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

1) icu เป็นระบบการตัดคำระบบเริ่มแรกของ PyThaiNLP แต่มีความแม่นยำในการตัดอยู่ในระดับต่ำ และเป็นค่าเริ่มต้น

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

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

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

5) pylexto เป็นระบบที่ใช้ LexTo ในการตัดคำ

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

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

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

2.4.5.1 ความถี่ของคำ (Term Frequency: TF)

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

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

2.4.5.2 ความถี่ของเอกสารผกผัน (Inverse document Frequency: IDF)

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

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

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

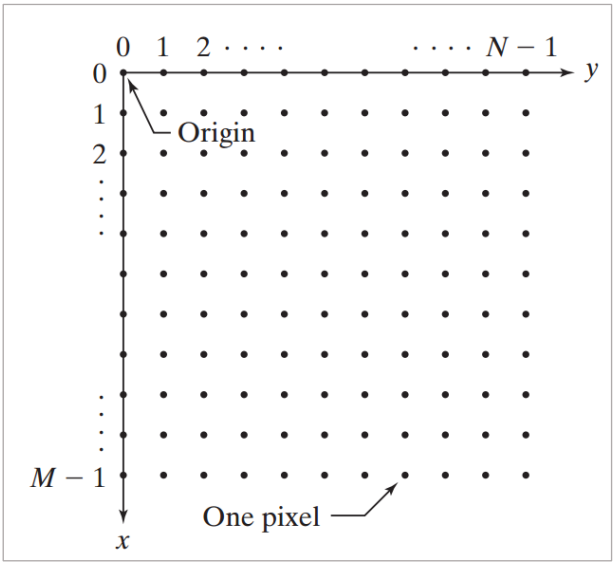
$$TF\text{-}IDF = TF \times IDF \tag{2.3}$$

2.5 การประมวลผลภาพดิจิทัล (Digital Image Processing)

2.5.1 ความหมายของการประมวลผลภาพ

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

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



ที่มา: Gonzalez, et al (2009)

รูปที่ 2.1 จุดพิกเซลบนภาพดิจิทัล

โดยภาพดิจิทัลขนาด M x N ในรูปแบบเมทริกซ์แสดงดังรูปที่ 2.2

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

ที่มา: Gonzalez, et al (2009)

รูปที่ 2.2 ภาพดิจิทัลขนาด M x N ในรูปแบบเมทริกซ์

2.5.2 ขั้นตอนการประมวลผลภาพ (Step of Image Processing)

แพรท (Pratt, 2001) ได้จำแนกขั้นตอนในการประมวลผลภาพดิจิทัล โดยมีรายละเอียด ดังนี้

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

- การจัดการกับวัตถุ หากไม่จัดการกับการวางตัวของวัตถุ อุปกรณ์ประมวลผลต้องหาทิศทางของวัตถุเอง

- ระยะระหว่างกล้องหรือเลนส์ถึงวัตถุ ซึ่งจะเป็นตัวกำหนดขนาดของวัตถุที่ระบบมองเห็น

- การจัดการเรื่องแสง โดยแสงเป็นองค์ประกอบที่สำคัญ เนื่องจากการมองเห็นภาพของระบบอัตโนมัติเกิดจากการที่แสงตกกระทบวัตถุ แล้วสะท้อนผ่านเลนส์เข้าตัวเซนเซอร์รับภาพของกล้อง ซึ่งการจัดการเกี่ยวกับแสงจำเป็นต้องพิจารณาเรื่องการเลือกใช้แหล่งกำเนิดแสง

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

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

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

- **การแยกบริเวณโดยใช้ค่าเทรสโชลด์ (Threshold)** โดยการแปลงภาพระดับเทาเป็นภาพไบนารี (Binary image) โดยค่าเทรสโชลด์มีค่าระหว่าง 0-255 ถ้าค่าความเข้มแสงที่พิกเซลตำแหน่งใดมีค่าต่ำกว่าหรือเท่ากับค่าเทรสโชลด์ จะกำหนดให้ค่าพิกเซลในตำแหน่งนั้นมีค่าเป็น 0 หรือเป็นด้านมืด และถ้าพิกเซลใดมีค่าสูงกว่าค่าเทรสโชลด์ ให้พิกเซลนั้นมีค่าเป็น 255 หรือเปลี่ยนเป็นด้านสว่าง

- **การแยกบริเวณโดยใช้ขอบของวัตถุ (Edge Based Segmentation)** การใช้วิธีนี้ต้องคำนวณหาขอบภาพของวัตถุ ซึ่งขอบในความหมายของการประมวลผลภาพคือพิกเซลที่มีการเปลี่ยนแปลงความเข้มแสงเกินค่าที่กำหนด

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

6) การจำแนกวัตถุและการแปลความหมาย (Classification and Interpretation) เป็นขั้นตอนการจัดกลุ่มให้วัตถุว่าเป็นวัตถุที่อยู่ในกลุ่มใด โดยอาศัยข้อมูลที่ได้จากการการคำนวณซึ่งเป็นคุณสมบัติของวัตถุนั้น ๆ เปรียบเทียบกับวัตถุตัวอย่างที่อยู่ในแต่ละกลุ่ม ก่อนที่ระบบจะสามารถตัดสินใจได้ ระบบจะต้องมีตัวอย่างของวัตถุในแต่ละกลุ่ม เพื่อใช้ในการเปรียบเทียบและสามารถแปลความหมายที่ถูกต้อง

2.5.3 องค์ประกอบในการประมวลผลภาพ (Components of Image Processing)

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

1) ฮาร์ดแวร์การประมวลผลภาพเฉพาะ (Specialized Image Processing Hardware) ประกอบด้วยฮาร์ดแวร์ที่ใช้งานการดำเนินงานแบบดั้งเดิมทำงานร่วมกับเครื่องมือในการแปลงค่าตัวเลข (Digitizer) มีหน้าในการรับส่งข้อมูลที่ใช้ความเร็วเป็นลักษณะเฉพาะ

2) คอมพิวเตอร์ (Computer) ในระบบประมวลผลภาพหมายถึง เครื่องคอมพิวเตอร์ทั่วไปที่สามารถใช้งานได้ไปจนถึงซูเปอร์คอมพิวเตอร์ ในการใช้งานเฉพาะบางครั้งคอมพิวเตอร์ที่กำหนดเองจะใช้เพื่อให้บรรลุระดับของประสิทธิภาพที่ต้องการ ซึ่งระบบแทบทุกระบบมีอุปกรณ์ครบถ้วนเหมาะสำหรับการประมวลผลภาพแบบออฟไลน์

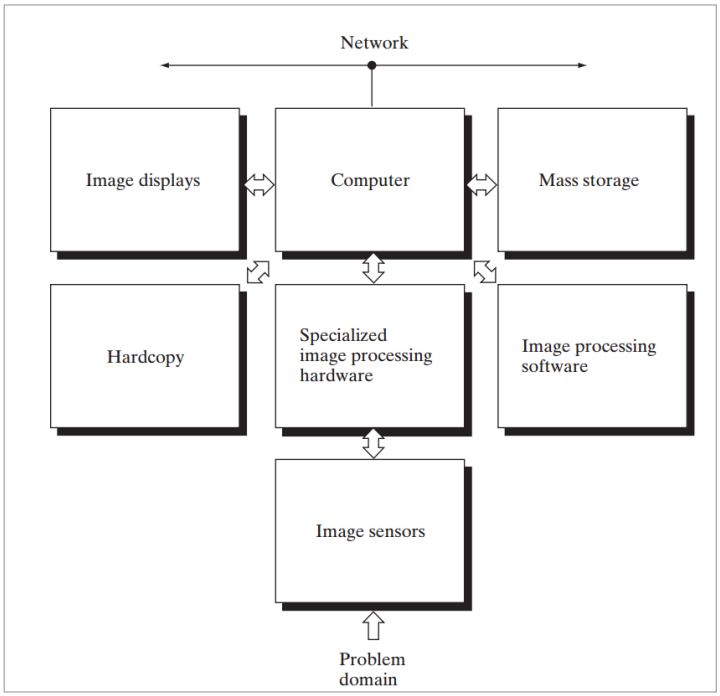
3) โปรแกรมสำหรับการประมวลผลภาพ (Image Processing Software) ประกอบด้วยโมดูลเฉพาะที่ดำเนินการเฉพาะ นอกจากนี้ยังมีความสามารถสำหรับผู้ใช้ในการเขียนโปรแกรมที่น้อยที่สุดโดยใช้โมดูลเฉพาะ โปรแกรมที่ซับซ้อนมากขึ้นจะช่วยให้สามารถรวมโมดูลต่าง ๆ เข้าด้วยกัน และใช้คำสั่งในการเขียนโปรแกรมทั่วไปจากภาษาคอมพิวเตอร์อย่างน้อยหนึ่งภาษา

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

5) จอภาพ (Image Displays) ภาพที่ใช้ในปัจจุบันส่วนใหญ่เป็นสี จอภาพถูกขับเคลื่อนด้วยผลลัพธ์ของการแสดงภาพและกราฟิก การ์ดที่เป็นส่วนสำคัญของระบบคอมพิวเตอร์มักไม่มีข้อกำหนดสำหรับแอปพลิเคชันการแสดงภาพที่การ์ดแสดงผล

6) สิ่งพิมพ์ออก (Hardcopy) ใช้สำหรับการบันทึกภาพ รวมถึงเครื่องพิมพ์เลเซอร์ กล้องฟิล์ม อุปกรณ์ที่ไวต่ออุณหภูมิหน่วยอิงค์เจ็ตและหน่วยดิจิทัล เช่น ดิสก์ และ CDROM ฟิล์มควรมีความละเอียดสูง อุปกรณ์ที่ใช้ควรเป็นอุปกรณ์ที่ได้รับการยอมรับและเป็นมาตรฐานสำหรับการนำเสนอภาพ

7) เครือข่าย (Networking) เป็นฟังก์ชันเริ่มต้นในระบบคอมพิวเตอร์ในปัจจุบัน เนื่องจากข้อมูลจำนวนมากที่มีอยู่ในแอปพลิเคชันการประมวลผลภาพ การพิจารณาที่สำคัญในการรับส่งข้อมูลคือแบนด์วิธ (Bandwidth)



ที่มา: Gonzalez and Woods (2008)

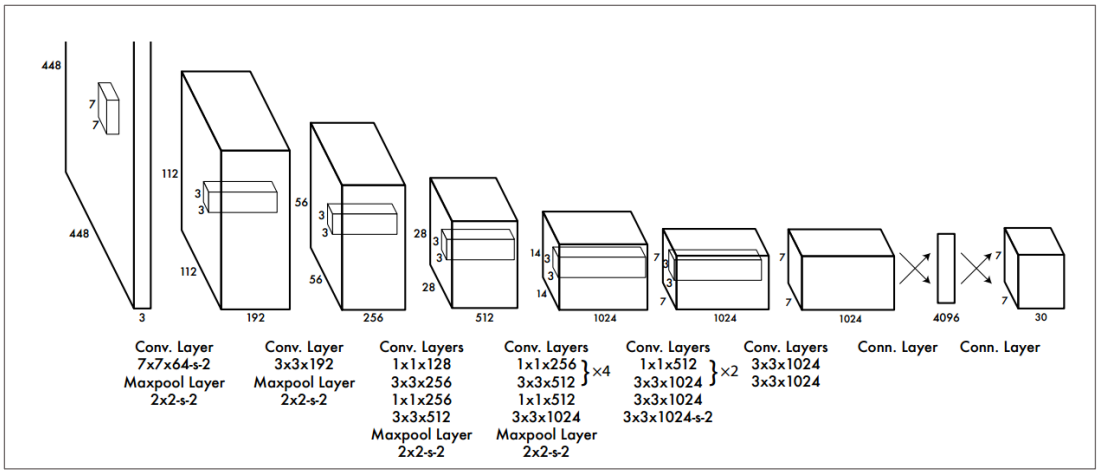
รูปที่ 2.3 องค์ประกอบในการประมวลผลภาพดิจิทัล

จากการทบทวนวรรณกรรมเกี่ยสกับการประมวลผลภาพดิจิทัล พบว่า ขั้นตอนในการประมวลผลภาพมีหลายขั้นตอน เละมีความละเอียดซับซ้อน โดยงานวิจัยนี้ได้นำขั้นตอนการประมวลผลภาพมาใช้ในการสกัดองค์ประกอบภายในภาพจากการรวบวรมภาพที่มาจากสื่อสังคมออนไลน์ เพื่อให้สามารถจำแนกคุณลักษณะที่แตกต่างกันของภาพได้ ซึ่งต้องเรียนรู้ขั้นตอนการประมวลผลภาพ ร่วมกับเทคนิคการเรียนรู้ของเครื่อง และการเรียนรู้เชิงลึก เพื่อลดข้อผิดพลาดในการประมวลผลภาพ

2.6 โยโล่ (You Only Look Once: YOLO)

2.6.1 คุณลักษณะของ YOLO

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

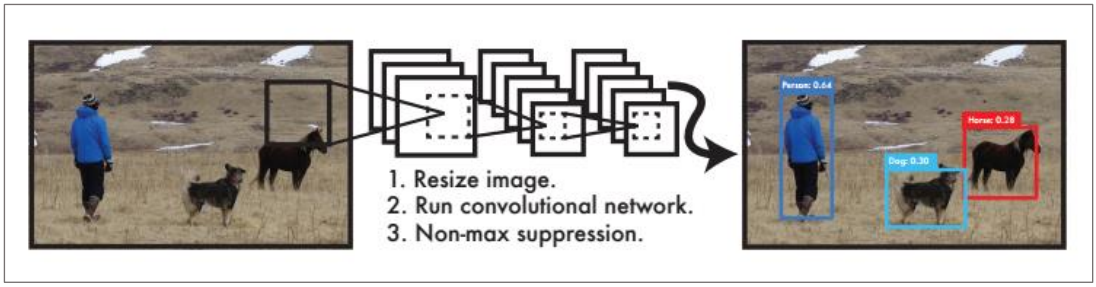


ที่มา: (Redmon, et al, 2016)

รูปที่ 2.4 สถาปัตยกรรมของ YOLO

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

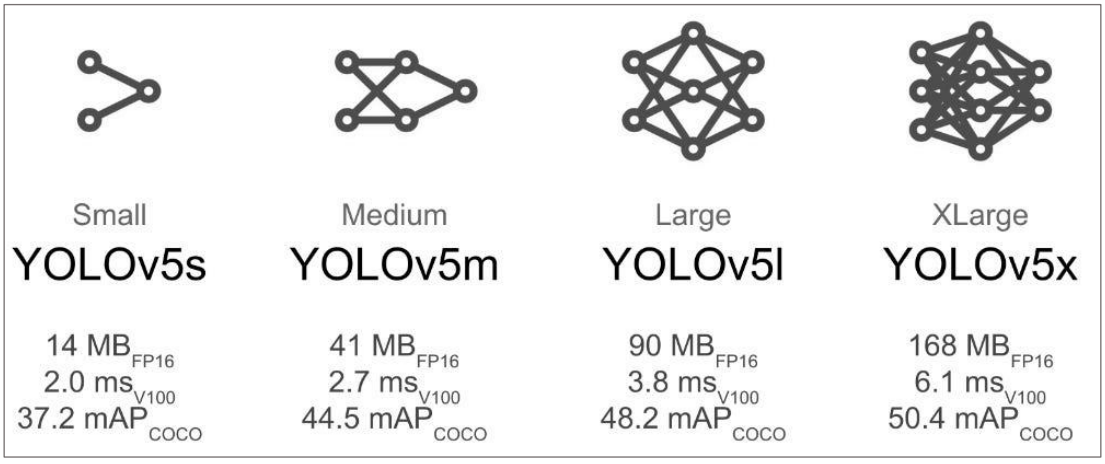
ในเริ่มแรกระบบการตรวจจับภาของ YOLO มีความเรียบง่ายและรวดเร็ว ประกอบด้วย (1) ข้อมูลนำเข้า ซึ่งเป็นภาพขนาด 448x448 (2) กระบวนการตรวจจับภาพโดยโครงข่ายแบบสังวัตนาการ และ (3) การตรวจสอบความเชื่อมั่นของแบบจำลอง ซึ่งองค์ประกอบในการทำงานของ YOLO แสดงดังรูปที่ 2.5



ที่มา: (Redmon, et al, 2016)

รูปที่ 2.5 องค์ประกอบในการทำงานของ YOLO

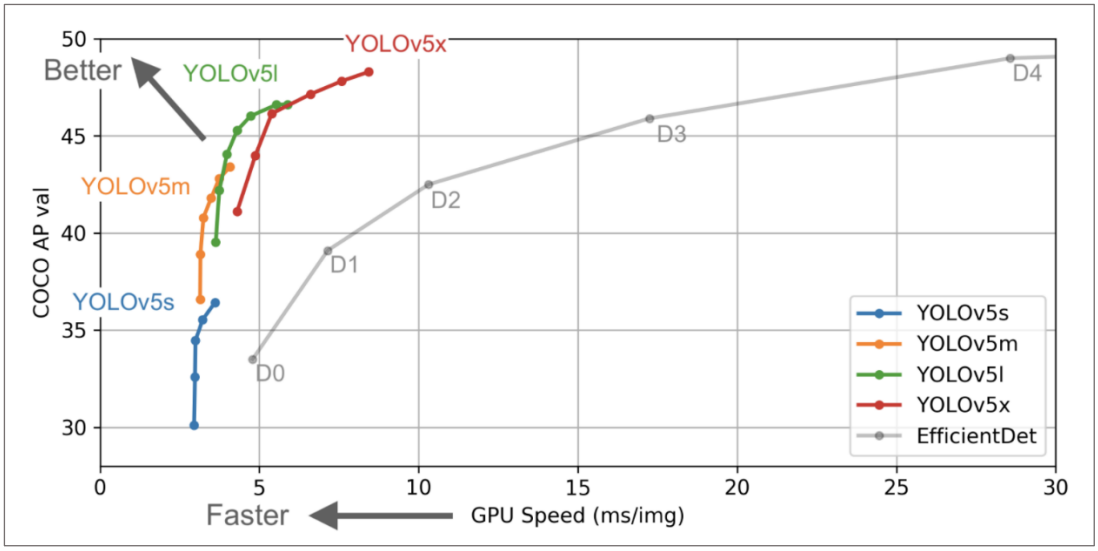
ในปัจจุบันได้มีการพัฒนา YOLO มาหลายเวอร์ชัน โดยเวอร์ชันปัจจุบันคือ YOLOv5 ซึ่งเป็นเวอร์ชันที่งานวิจัยนี้นำมาใช้ในการจำแนกประเภทของอาหารบนภาพ ซึ่ง YOLOv5 ถูกพัฒนามาจากYOLOv4 ที่มีการปรับปรุงประสิทธิภาพในการทำงานให้ดีขึ้น แต่ยังคงมีความรวดเร็วในการทำงานเช่นเดียวกับเวอร์ชันที่มีมาก่อนหน้านี้ โดย YOLOv5 แบ่งขั้นตอนวิธีในการทำนายผลและสร้างแบบจำลองออกเป็นแบบจำลองขนาดเล็ก ซึ่งใช้รหัสเป็นตัวอักษร s และแบบจำลองขนาดใหญ่ที่ใช้รหัสเป็นตัวอักษร x โดยขั้นตอนวิธีของ YOLOv5 แสดงดังรูปที่ 2.6



ที่มา: (Redmon, et al, 2016)

รูปที่ 2.6 ขั้นตอนวิธีของ YOLOv5

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

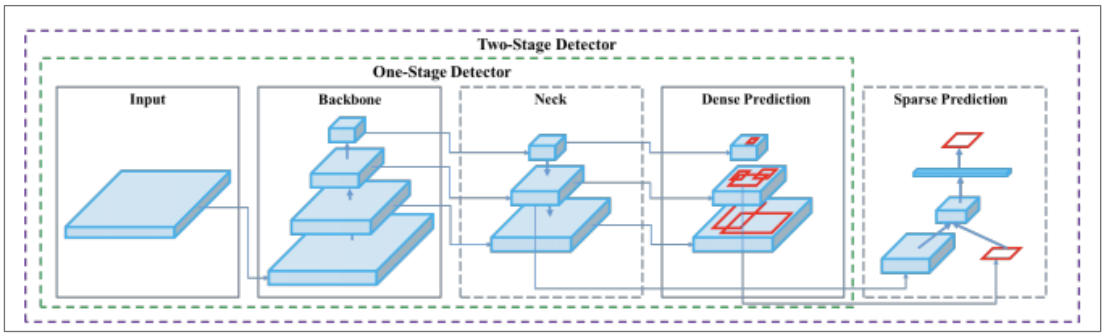


ที่มา: Redmon, et al, 2016

รูปที่ 2.7 ประสิทธิภาพการทำงานของ YOLOv5

2.6.2 โครงสร้างการทำงานของ YOLOv5

การทำงานของ YOLOv5 เริ่มต้นด้วยการนำเข้าข้อมูลนำเข้าเข้าสู่แบบจำลองของ YOLOv5 ซึ่งจะมีการแบ่งเซลล์บนภาพเป็นแบบตาราง (Grid cell) ออกเป็นช่องเล็ก ๆ ขนาด n x n (Bochkovskiy, et al, 2020) โดยโครงสร้างการทำงานของ YOLOv5 แสดงดังรูปที่ 2.8



ที่มา: Bochkovskiy and et al, 2020

รูปที่ 2.8 โครงสร้างการทำงานของ YOLOv5

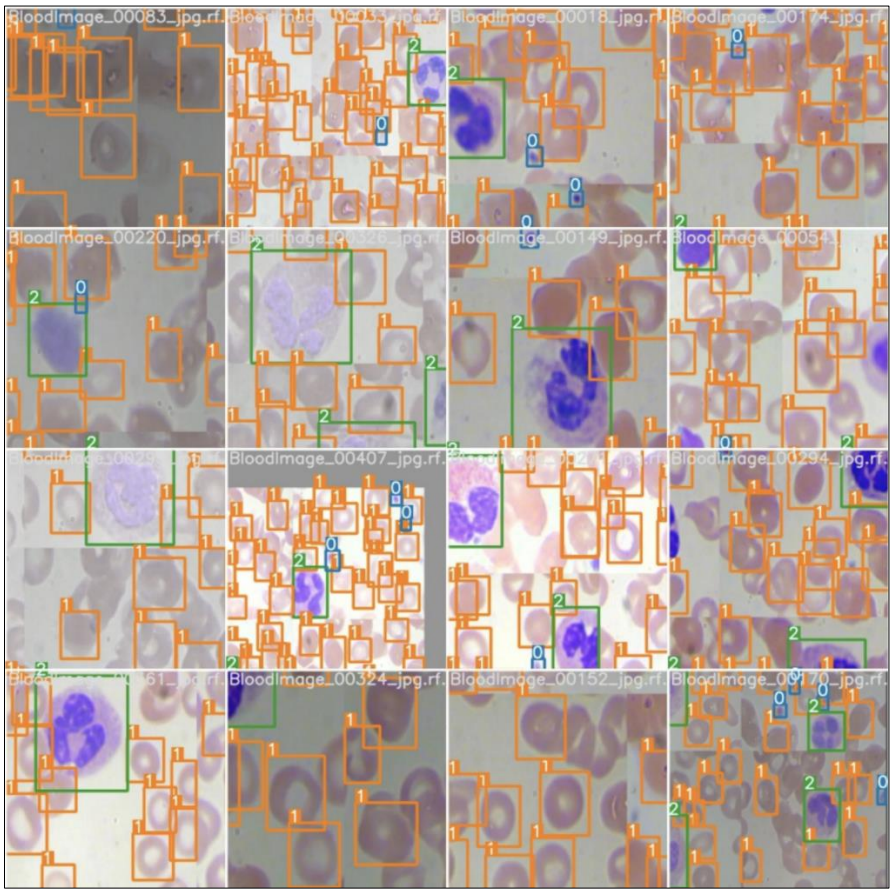
จากโครงสร้างการทำงานของ YOLOv5 ประกอบด้วยส่วนประกอบหลัก 3 ส่วน ดังนี้

1. Backbone คือส่วนที่ทำหน้าที่รับข้อมูลนำเข้าเข้าสู่แบบจำลอง แล้วสกัดคุณลักษณะที่โดดเด่นของภาพออกมา โดยจะส่งต่อไปยัง Neck ซึ่งเป็นส่วนที่รับคุณลักษณะของภาพ

2. Neck คือส่วนที่หน้าที่รับคุณลักษณะของภาพที่ได้จากการสกัดจากส่วน Backbone มาสร้างแผนที่คุณลักษณะ (Feature map) ของภาพ

3. Head คือส่วนที่ทำหน้าที่ตีกรอบตำแหน่งของวัตถุบนภาพที่สนใจ แล้วทำนายคลาส (Class) ของวัตถุนั้น ๆ และส่งต่อไปยังชั้นผลลัพธ์ (Output Layer) เพื่อแสดงผลลัพธ์จากการทำนาย

โดยตัวอย่างภาพที่มีการสร้างแบบจำลองสำหรับการตรวจจับวัตถุบนภาพด้วย YOLOv5 แสดงดังรูปที่ 2.9



รูปที่ 2.9 ตัวอย่างการตรวจจับวัตถุบนภาพด้วย YOLOv5

2.7 การเรียนรู้ของเครื่อง (Machine Learning)

2.7.1 ความหมายของการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง เป็นสาขาหนึ่งของปัญญาประดิษฐ์ที่พัฒนามาจากการศึกษาการรู้จำที่เกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยแบบจำลองที่สร้างมาจากชุดข้อมูลนำเข้าเพื่อการทำนายหรือตัดสินใจในภายหลัง ซึ่งจะไม่ทำงานตามลำดับของคำสั่งโปรแกรมคอมพิวเตอร์ (Han and Kamber, 2006)

2.7.2 เทคนิคการเรียนรู้ของเครื่อง

ในการจำแนกการเทคนิคการเรียนรู้ของเครื่องนั้น สามารถจำแนกออกเป็นหมวดหมู่ ดังนี้ (Ian and Eibe, 2005)

1) การจำแนกประเภท (Classification) เป็นการทำนายประเภทของข้อมูล จากการหารูปแบบของชุดข้อมูลที่มีความใกล้เคียงกันหรือเหมือนกันมาก โดยชุดข้อมูลที่ถูกทำนายเกิดจากการเรียนรู้จากชุดข้อมูลที่มีอยู่แล้ว แบบจำลองที่เกิดจากการทำนาย มีดังนี้

1.1) ต้นไม้ตัดสินใจ (Decision Tree)

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

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

โดยที่ S คือ ตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัวแปรตาม

E คือ เอนโทรปีของตัวอย่าง

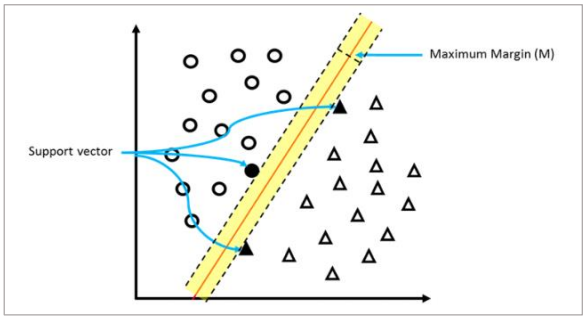
A คือ ตัวแปรต้นที่พิจารณา

value (A) คือ เซตของค่าของ A ที่เป็นไปได้

S_v คือ ตัวอย่างที่ A มีค่า v ทั้งหมด

1.2) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)

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



ที่มา: ศาสตรา วงศ์ธนวสุ (2560)

รูปที่ 2.10 ความกว้างระหว่างเส้นตรงกับซัพพอร์ตเวกเตอร์

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

$$h_0(x) = w_1x_1 + w_2x_2 + \cdots + w_nx_n + = w^Tx + b \qquad (2.5)$$

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

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

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

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

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

$$minimise_{w,b} = \frac{1}{2} w^T w$$
$$subject\ to\ t^{(i)}\left(w^T x^i + b\right) \geq 1 \quad (2.8)$$

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

$$minimise_{w,b} = \frac{1}{2} w^T w + C \sum_{i=1}^{m} \zeta^{(i)}$$
$$subject\ to\ t^{(i)}\left(w^T x^i + b\right) \geq 1 - \zeta^{(i)} \quad (2.9)$$

อย่างไรก็ตาม SVM มีข้อจำกัดคือสามารถสร้างเส้นแบ่งขอบเขตการตัดสินใจแบบเส้นตรงเท่านั้น ซึ่งอาจทำงานได้ไม่ดีหากความสัมพันธ์ของข้อมูลมีความซับซ้อน และทำให้ไม่สามารถแบ่งด้วยเส้นตรงไม่ได้ จึงใช้วิธีแก้ปัญหาที่เรียกว่าฟังก์ชันเคอร์เนลเคอร์เนล (Kernel Function)

Kernel คือวิธีทางคณิตศาสตร์ที่ทำให้ขั้นตอนวิธี SVM สามารถหาค่าตัวแปรแบบพหุนาม (Polynomial) ได้ โดยไม่ต้องไปเปลี่ยนรูปแบบและความสัมพันธ์ของคุณลักษณะตั้งต้น ซึ่งต้องดัดแปลงเป้าหมายการกำหนดค่าให้อยู่ในรูปแบบที่ Kernel ทำงานได้ โดยปรับเปลี่ยนจากรูปแบบในสมการที่ (2.9) ให้เป็นดังสมการที่ (2.10)

$$minimise_{w,b} = \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha^{(i)}\,\alpha^{(j)}t^{(i)}t^{(j)}x^{(i)T}x^{(j)} - \sum_{i=1}^{m}\alpha^{(i)}$$
$$subject\ to\ \alpha^{(i)} \geq 0 \qquad (2.10)$$

1.3) นาอีฟเบย์ (Naive Bayes)

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

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

โดยที่ $P(h)$ คือ ค่าความน่าจะเป็นของสมมติฐานที่คลาสเป็น h

$P(D)$ คือ ค่าความน่าจะเป็นของสมมติฐานที่ข้อมูลเป็น D

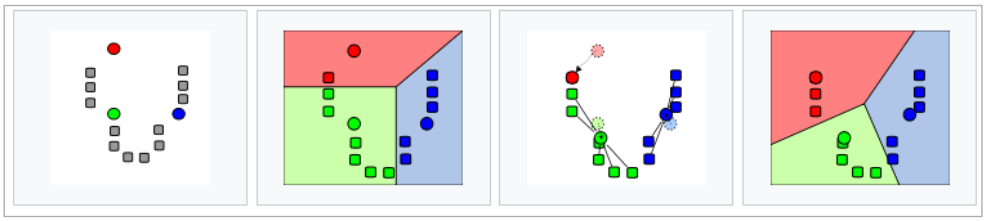
$P(h|D)$ คือ ค่าความน่าจะเป็นของสมมติฐานที่ข้อมูลเป็น D จะมีคลาสเป็น h

$P(D/h)$ คือ ค่าความน่าจะเป็นของสมมติฐานที่คลาส h จะมีข้อมูลเป็น D

ในกรณีที่มีตัวแปรต้นหรือข้อมูลที่ต้องพิจารณามากกว่า 1 ค่า ความน่าจะเป็นสามารถคำนวณได้จากผลคูณของความน่าจะเป็นของแต่ละข้อมูลที่มีคลาส h ดังสมการที่ (2.12)

$$P(h|D) = P(D_1|h) \times P(D_2|h) \times \ldots \times P(D_n|h) \times P(h) \qquad (2.12)$$

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



ที่มา: MacQueen (1967)

รูปที่ 2.11 ขั้นตอนวิธีค่าเฉลี่ย K

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

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

2.8 การเรียนรู้เชิงลึก (Deep Learning)

2.8.1 ความหมายของการเรียนรู้เชิงลึก (Deep Learning)

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

2.8.2 ประเภทของการเรียนรู้เชิงลึก (Way to Deep Learning)

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

2.8.2.1 การเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน (Unsupervised Deep Networks Learning)

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

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

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

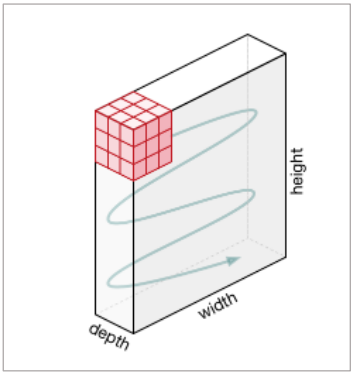
2.8.2.2 การเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน (Supervised Deep Networks Learning)

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

1) โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) เป็นโครงข่ายประสาทเทียมชนิดที่มีการคำนวณซ้ำ นิยมใช้ในการวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลา

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

ช่วยดึงคุณลักษณะที่ใช้ในการรู้จำวัตถุออก โดยปกติตัวกรองอันหนึ่งจะดึงคุณลักษณะที่สนใจออกมาได้หนึ่งอย่าง จึงจำเป็นต้องมีตัวกรองหลายตัว เพื่อหาคุณลักษณะทางพื้นที่ต่าง ๆ ประกอบกัน ซึ่งในการหาคุณลักษณะตัวกรองจะเลื่อนมีลักษณะการเลื่อนของตัวกรองแสดงดังรูปที่ 2.12



ที่มา: Saha (2018)

รูปที่ 2.12 ลักษณะการเลื่อนของตัวกรอง

2.8.2.3 การเรียนรู้โครงข่ายเชิงลึกแบบผสม (Hybrid Deep Networks Learning)

การเรียนรู้โครงข่ายเชิงลึกแบบผสม มีวัตถุประสงค์เพื่อเพิ่มประสิทธิภาพของการเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน หรือการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอน โดยสามารถใช้ทั้งการเรียนรู้โครงข่ายเชิงลึกแบบไม่มีผู้สอน และการเรียนรู้โครงข่ายเชิงลึกแบบมีผู้สอนร่วมกันได้

2.9 งานวิจัยที่เกี่ยวข้อง

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

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

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

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

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

นอกจากนี้ งานวัยที่มีการประมวลผลเกี่ยวกับภาพ ได้มีการพัฒนาระบบตรวจสอบองค์ประกอบของฮาร์ดดิสก์ไดร์ฟ (อนุภาค ปัญญทาสมบัติ, 2564) โดยใช้เทคนิคการเรียนรู้เชิงลึกด้วยโครงข่ายประสาทเทียมแบบสังวัตนาการ และโยโล่ เวอร์ชัน 5 ผลการศึกษาที่ได้พบว่า ขั้นตอนวิธีดังกล่าวเป็นขั้นตอนวิธีที่มีประสิทธิภาพและสามารถตรวจสอบได้อย่างรวดเร็ว โดยความถูกต้องในการตรวจสอบองค์ประกอบของฮาร์ดดิสก์ไดร์ฟเป็นร้อยละ 99.49 ส่วนงานวิจัยที่ใช้โยโล่สำหรับการประมวลผลภาพนั้น เป็นการพัฒนาแบบจำลองในการตรวจจับหนังสือแบบเรียลไทม์ (Kalinina, M., 2020) โดยทดสอบขั้นตอนวิธีของโยโล่ เวอร์ชัน 3 ทั้งหมด 3 ขั้นตอนวิธี คือแบบ 3 ผลลัพธ์ แบบ 2 ผลลัพธ์ และแบบขนาดเล็ก (Tiny) ผลการศึกษาพบว่า ค่าความถูกต้องในการตรวจจับหนังสือมีค่าเป็นร้อยละ 94.43 92.35 และ 88.75 ตามลำดับ อีกทั้งมีงานวิจัยที่ตรวจจับศัตรูพืชที่ส่งผลต่อผลผลิตทางการเกษตรโดยใช้ขั้นตอนวิธีโยโล่ (Luo, Y., et al, 2023) ผลการวิจัยที่ได้พบว่าสามารถระบุชนิดของศัตรูพืชได้ด้วยค่าความแม่นยำร้อยละ 96.8

โดยตารางการเปรียบเทียบงานวิจัยที่กล่าวมาข้างต้น แสดงดังตารางที่ 2.2

ตารางที่ 2.2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

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

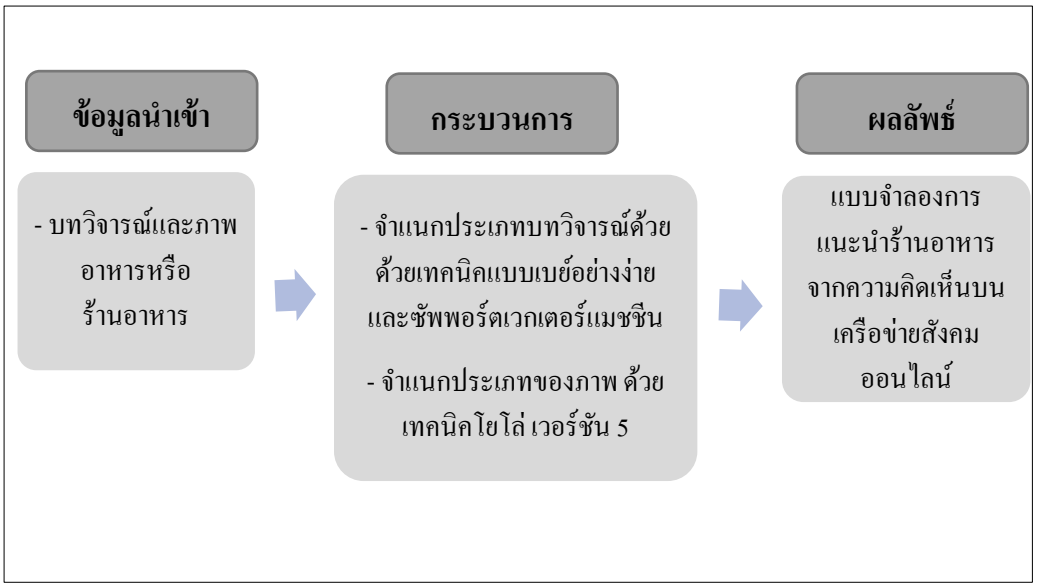
ตารางที่ 2.2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง (ต่อ)

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

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

2.10 กรอบแนวคิดการวิจัย

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



รูปที่ 2.13 กรอบแนวคิดการวิจัย