

บทที่ 2

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

ในบทที่ 2 นี้จะกล่าวถึงทฤษฎีแนวคิดที่เกี่ยวข้องในการดำเนินโครงการอันได้แก่ ความหมายของ Recommendation Systems วิธีการของระบบแนะนำในปัจจุบัน ข้อดีข้อเสียของแต่ละวิธีในปัจจุบัน ซึ่งทั้งหมดมีรายละเอียดของเนื้อหาดังต่อไปนี้

2.1 ความหมายของระบบแนะนำ

Recommendation Systems หรือระบบแนะนำ คือ ระบบที่แนะนำข้อมูล ผลิตภัณฑ์ หรือ ผู้คนให้กับผู้ใช้ระบบ โดยอ้างอิงจากสมมติฐานการเรียนรู้ข้อมูลความชอบหรือความต้องการ ณ ขณะนั้นของผู้ใช้ โดยระบบแนะนำกลายเป็นหัวข้องานวิจัยที่สำคัญตั้งแต่มีการปรากฏตัวของ งานวิจัยเรื่อง Collaborative Filtering ในช่วงกลางทศวรรษที่ 1990 โดยความสำเร็จในช่วงแรก ๆ ของการทำระบบแนะนำเกิดมาจากการเจริญเติบโตของธุรกิจทางด้าน E-Commerce โดยตัวอย่าง ของ Applications ที่ใช้ระบบแนะนำในการเลือกซื้อสินค้า ซีดีเพลง หรือผลิตภัณฑ์อื่นๆ ได้แก่ เว็บ Amazon.com และเว็บ CDNow.com เป็นต้น ทั้งนี้เพื่อให้ลูกค้าเกิดความภักดีต่อองค์กร และเพิ่ม โอกาสทางธุรกิจ อย่างไรก็ตามไม่ว่าระบบแนะนำจะก้าวหน้าไปเช่นไร ระบบแนะนำก็ยังคง ต้องการการพัฒนาความสามารถมากกว่านี้ เพื่อที่จะทำให้ระบบสามารถแนะนำสิ่งต่างๆ ให้แก่ผู้ ใช้ได้อย่างมีประสิทธิภาพมากยิ่งขึ้นและใช้งานได้ง่ายขึ้นแม้แต่กิจกรรมต่างๆ ไปชีวิตประจำวัน เช่น การแนะนำแพ็คเกจช่วงเวลาการพักร้อน การซื้อสินค้าในร้านที่มีระบบ Smart Shopping Cart

ทั้งนี้ ระบบแนะนำสามารถจำแนกประเภทจากวิธีการแนะนำเป็น 3 ประเภท ได้แก่ 1. วิธีพิจารณาจากข้อมูลสิ่งของ (Content-based approach) 2. วิธีพิจารณาการคัดกรองร่วมกัน (Collaborative filtering) และ 3. แบบผสม (Hybrid) สำหรับระบบแนะนำด้วยวิธีพิจารณาจากข้อมูล สิ่งของ ระบบจะแนะนำสิ่งของที่มีความคล้ายคลึงกับสิ่งของอื่น ที่ผู้ใช้พึงใจในอดีต สำหรับระบบ แนะนำด้วยวิธีพิจารณาการคัดกรองร่วมกัน ระบบจะแนะนำสิ่งของจากสิ่งของที่ผู้ใช้คนอื่น ใน ระบบพึงใจ โดยกลุ่มผู้ใช้อย่างนี้นั้น มีความพึงใจคล้ายคลึงกับผู้ใช้ระบบในปัจจุบันและสำหรับ ระบบแนะนำแบบผสม ระบบจะนำวิธีพิจารณาจากข้อมูลสิ่งของ และวิธีพิจารณาการคัดกรอง ร่วมกัน มาใช้ร่วมกัน ซึ่งสามารถนำมาประยุกต์ได้หลากหลายรูปแบบ นอกจากนี้ระบบแนะนำยังสามารถจำแนกประเภทได้จากเทคนิคของขั้นตอนวิธี (Algorithm) ได้เป็น 2 ประเภท ได้แก่ 1. วิธี

พิจารณาจดจำ (Memory-based) และ 2. วิธีพิจารณาตัวแบบ (Model-based) สำหรับวิธีพิจารณาจดจำกล่าวได้ว่าเป็นวิทยาการศึกษาศาสนิก (Heuristics) อย่างหนึ่ง กล่าวคือระบบจะคำนวณการแนะนำอย่างอัตโนมัติโดยอาศัยข้อมูลกิจกรรมในอดีตของผู้ใช้ในทางตรงกันข้าม สำหรับวิธีพิจารณาตัวแบบ ในขั้นต้นระบบจะเรียนรู้กิจกรรมของผู้ใช้ และจัดทำตัวแบบเฉพาะผู้ใช้นั้นจากนั้นจึงนำตัวแบบดังกล่าวมาใช้เพื่อการแนะนำต่อไป

ระบบแนะนำจะทำการเปรียบเทียบคำบรรยายลักษณะผู้ใช้ (User profile) กับข้อมูลเพิ่มเติมอื่นๆ เพื่อคาดการณ์คะแนนสิ่งของที่ผู้ใช้นั้นดังกล่าว ยังไม่เคยให้คะแนนมาก่อน วิธีการให้ได้มาซึ่งคะแนนดังกล่าวนี้ อาจได้มาจากวิธีพิจารณาจากข้อมูลสิ่งของ (Content-based approach) หรือจากวิธีพิจารณาจากข้อมูลของผู้ใช้ในระบบ (Collaborative filtering approach)

เมื่อเราต้องการจัดทำคำบรรยายลักษณะผู้ใช้ เราอาจดำเนินการจัดเก็บข้อมูลได้ 2 ลักษณะคือการเก็บข้อมูลแบบชัดเจน (Explicit) หรือ การเก็บข้อมูลแบบแฝงนัยยะ (Implicit)

ตัวอย่างของการเก็บข้อมูลแบบชัดเจน ได้แก่

1. สร้างแบบฟอร์มให้ผู้ใช้ให้คะแนนสิ่งของ โดยแบ่งเป็นระดับความชอบ
2. สร้างแบบฟอร์มให้ผู้ใช้เรียงลำดับสิ่งของจากความพึงพอใจสูงสุดจนถึงต่ำสุด
3. นำเสนอสิ่งของ 2 ชิ้นให้กับผู้ใช้ จากนั้นจึงให้เลือกสิ่งของชิ้นที่ผู้ใช้พึงพอใจ
4. สร้างแบบฟอร์มให้ผู้ใช้ระบุสิ่งของที่ตนพึงพอใจ

ตัวอย่างของการเก็บข้อมูลแบบแฝงนัยยะ ได้แก่

1. สังเกตสิ่งของที่ผู้ใช้เลือกเข้าชม ในร้านค้าออนไลน์ (Online store)
2. วิเคราะห์จำนวนครั้งเข้าชมสิ่งของ ของผู้ใช้
3. จัดเก็บรายการการสั่งซื้อผ่านระบบอินเทอร์เน็ตของผู้ใช้
4. จัดเก็บรายการสิ่งของที่ผู้ใช้เคยเข้าชมหรือรับฟังจากคอมพิวเตอร์ของผู้ใช้
5. วิเคราะห์ข้อมูลความสนใจในเรื่องต่างๆ จากเครือข่ายสังคมออนไลน์

ปกติแล้วปัญหาของระบบแนะนำได้ถูกลดลงไปจนเหลือเพียงปัญหาของการคำนวณเรตติ้งสำหรับไอเท็มที่ผู้ใช้ยังไม่เคยเห็น โดยการคำนวณนี้จะอาศัยเรตติ้งที่ผู้ใช้นั้นๆ ได้ให้ไว้กับไอเท็มอื่นๆ และอาศัยข้อมูลอื่นๆ อีกด้วย ถ้าหากเราสามารถคำนวณหาเรตติ้งให้กับไอเท็มที่ยังไม่ได้รับการให้เรตติ้ง เราจะสามารถแนะนำไอเท็มที่มีค่าคำนวณสูงสุดให้กับผู้ใช้ได้

Adomavicius & Tuzhilin (2005) ได้ให้แนวความคิดเกี่ยวกับระบบแนะนำคือระบบที่แนะนำข้อมูลต่าง ๆ ให้กับผู้ใช้ระบบโดยอ้างอิงจากสมมุติฐานการเรียนรู้ข้อมูลความชอบหรือความต้องการของผู้ใช้ในขณะนั้น ซึ่งในปัจจุบันแบ่งออกเป็น 3 ประเภทคือ วิธีการที่ใช้พื้นฐานจากข้อมูล (Content-based Recommendation) วิธีที่ใช้ข้อมูลในอดีต (Collaborative Filtering) และวิธีการแบบ

ผสม (Hybrid Approaches) โดยงานวิจัยนี้ได้ใช้หลักการประเมินค่าน้ำหนักคะแนนแบบหลายปัจจัยและหลายมิติเพื่อหาผู้ใช้ที่มีลักษณะความชอบหรือความต้องการใกล้เคียงกัน โดยผู้ใช้จะต้องทำการให้คะแนนสถานที่ท่องเที่ยว ที่พัก ร้านอาหาร ร้านขายของฝาก และร้านขายสินค้าหนึ่งตำบลหนึ่งผลิตภัณฑ์ ซึ่งจะเป็นค่าระดับคะแนน 1 ถึง 5

2.2 งานวิจัยที่เกี่ยวข้อง

ศิริลักษณ์และสุขุมล [1] มีการทำนายความชอบของผู้ชมที่มีต่อภาพวาด โดยบูรณาการเทคนิคการทำนายซึ่งพิจารณาจากคุณลักษณะพื้นฐานของ (Basic Characteristics) และความหมาย (Semantics) ของภาพวาด เช่น อารมณ์ ความรู้สึกที่ผู้ชมรับรู้ได้เมื่อชมภาพวาด หรือ โทนีสี เป็นต้น ร่วมกับคำทำนายโดยพิจารณาจากคำแนะนำจากผู้ชมผู้อื่น (Hybrid Content-Based Classification and Collaborative Filtering) ซึ่งงานวิจัยนี้มีวัตถุประสงค์เพื่อการทดสอบคุณลักษณะที่เหมาะสมในการทำนายโดยเปรียบเทียบระหว่างคุณลักษณะสองแบบ แบบที่หนึ่งคือคุณลักษณะความหมายร่วมกับคุณลักษณะพื้นฐานและแบบที่สองคุณลักษณะพื้นฐานเพียงอย่างเดียว นอกจากนี้ยังเปรียบเทียบประสิทธิภาพการทำนายสามเทคนิค คือ Content-Based Classification , Collaborative Filtering และ Hybrid Content-Based and Collaborative Filtering

Hung-Chen and Are L.P. Chen [2] ได้มีการออกแบบระบบ Music Recommendation System (MRS) เพื่อทำการแนะนำ Music ให้กับ User แต่ละคน โดยจะใช้ข้อมูลของ MIDI ในการนำมาวิเคราะห์ ซึ่งได้ทำการแบ่งกลุ่มของ music ออกเป็น 6 กลุ่ม ตามคุณลักษณะของ music และได้นำเอาข้อมูลของ user ของการเข้ามาใช้งานในอดีตนำมาวิเคราะห์ด้วย โดยนำเอาเทคนิค content Based , collaborative และ statistic-based recommendation มาใช้ในการวิเคราะห์ ซึ่งการวิเคราะห์นั้นอยู่บนพื้นฐานของคิกริความนิยมของ User ต่อกลุ่มของ music

Choochart, Chatchawal and Sarawoot [3] บทความนี้ได้กล่าวถึงประโยชน์ของระบบการให้คำแนะนำสินค้าซึ่งสามารถเพิ่มยอดขายได้มากขึ้น และเพื่อให้สามารถตอบสนองความต้องการของลูกค้าแบบทันทีทันใด (Real Time) ซึ่งใช้เทคนิค Content-Based Filtering กับ Collaborative Filtering และมีการเพิ่มเทคนิคการ Clustering ทำให้ระบบให้คำแนะนำมีประสิทธิภาพมากขึ้นเนื่องจากช่วยลดเวลาในการแนะนำสินค้าได้

Bo Yang, Toa Mei, Xian-Sheng Hua, Linjun Yang, Shi-Qiang Yang and Mingjing Li (2007) ศึกษาพบว่างานวิจัยแบบเก่าเริ่มตั้งแต่ปี ค.ศ.1990 หลากๆระบบแนะนำถูกนำเสนอในพื้นที่ที่หลากหลาย เช่น ภาพยนตร์ ทีวี เว็บไซต์ ซึ่งใช้สันนิษฐานว่า Profile ของผู้ใช้จะสามารถเป็นประโยชน์ โดยทั่วไป Profile ของผู้ใช้มาจาก 2 แหล่ง คือ 1. เลือกเองว่าจะใช้อะไรด้วยตนเอง 2.

ถูกแนะนำให้ใช้โดยอ้อมจากโดยสรุปแล้วไม่ต้องคำนึงถึงแง่ที่ถูกแนะนำวัตถุประสงค์หลัก คือ การแนะนำรายการที่เกี่ยวข้องกับการแนะนำแบบเดิมซึ่งขึ้นอยู่กับความสนใจของผู้ใช้เอง อย่างไรก็ตาม ขณะที่ท่องเว็บไซต์อยู่นั้นคนส่วนใหญ่มักเข้าชมหน้าเว็บที่สนใจวิดีโอออนไลน์ หลายกรณีการเข้าชมนั้นเข้าไม่ได้เพราะขาดการติดต่อจาก User

เทคนิคที่ใช้ในการค้นหาวิดีโอ แบ่งออกเป็น 2 ประเภท คือ text-based (เน้นตามตำรา) และ content-based (เน้นตามเนื้อหา) การค้นหาวิดีโอ แบบ Text-based คือ การค้นหาวิดีโอตามตำราหรือเอกสารที่เกี่ยวข้อง ขณะที่การค้นหาวิดีโอแบบ content-based คือ การใช้เนื้อหาที่เห็นเด่นชัดเพื่อการค้นหา เมื่อไม่นานมานี้มีการใช้ความพยายามอย่างมากในการใช้การค้นหาวิดีโอแบบ content-based ซึ่งเนื้อหาถูกนำมาใช้โดยตรงกับการคำนวณความเหมือนกันระหว่าง video ต่างๆ หรือ นำมาใช้กับการประเมินผลแบบมีปฏิสัมพันธ์ของผู้ใช้ โดยใช้วิธีการค้นหาแบบ text based อย่างไรก็ตาม การแนะนำโดยใช้ video แตกต่างจากการค้นหา video อยู่มาก ซึ่งทำให้การค้นหาไม่เสถียรในการปรับเทคนิคที่ใช้ในการค้นหาวิดีโอไปสู่การเสนอแนะวิดีโอ พวกเราจึงสรุปความแตกต่าง ดังต่อไปนี้

วัตถุประสงค์แตกต่างกัน (Different objectives) กล่าวคือ การค้นหาวิดีโอ เป็นการค้นหาวิดีโอที่ส่วนมากเข้ากับ (match) การสืบค้นต่างๆ หรือ ภาพที่ค้นหาในระดับความคิด อย่างไรก็ตาม วัตถุประสงค์ของการแนะนำวิดีโอเป็นการจัดลำดับวิดีโอที่ส่วนมากตรงกับแหล่งที่มาของวิดีโอ

การป้อนข้อมูลแตกต่างกัน (Different inputs) การป้อนข้อมูลแบบ video search มาจากชุดของ Keywords (คำสำคัญ) หรือ รูปภาพต่างๆ ซึ่งการป้อนข้อมูลทุกชนิดไม่มีอุปกรณ์ใดๆเลย ขณะที่การป้อนข้อมูลของการแนะนำวิดีโอประกอบด้วย video และอุปกรณ์เสริมที่เกี่ยวข้องยิ่งกว่านั้น อุปกรณ์เสริมต่างปกติก็น่าจะมีคุณสมบัติเฉพาะ เช่น การค้นหา(query) ชื่อเรื่อง (title) การแท็ก (tags) การวิจารณ์ (comments) เป็นต้น

ดังนั้น การจัดการกับความแตกต่างเกี่ยวกับข้อมูลที่มีรูปแบบหลากหลายจึงเป็นสิ่งที่น่าสนใจ

ชาริน ถิ่นพังงา และไพโรสน ศิริพูล (2552) พัฒนาระบบแนะนำร้านอาหารบนเว็บไซต์และการเชื่อมโยงผ่าน Google Map โดยใช้หลักการพิจารณาจากความชอบของผู้ใช้ โดยอาศัยความรู้สึกที่มีต่อร้านค้านั้นๆ ยกตัวอย่าง เช่น การเลือกร้านอาหาร ผู้ใช้อาจเลือกจากรู้สึกต่างๆ เช่น ความอร่อย ความสะอาดของร้าน ความรวดเร็ว การบริการ ความคุ้มค่า บรรยากาศภายในร้าน นอกเหนือจากนั้นระบบแนะนำร้านอาหารนี้ยังเชื่อมต่อเข้ากับระบบให้บริการแผนที่ของ Google หรือที่รู้จักกันว่า Google Maps API ซึ่งแผนที่นี้มีความสามารถพิเศษคือ ผู้ใช้สามารถปักหมุด

ตำแหน่งของสถานที่ต่างๆได้ โดยอ้างอิงตำแหน่งจากละติจูดและลองจิจูดของสถานที่นั้นๆ จึงได้พัฒนาระบบนี้มาใช้ร่วมกับระบบแนะนำร้านอาหาร เพื่อให้ผู้ใช้สามารถปักหมุดร้านอาหาร และสามารถตรวจสอบตำแหน่งของร้านอาหารต่างๆ ได้ เพื่อให้ผู้ใช้ได้เห็นภาพว่าร้านอาหารที่กล่าวถึงอยู่ตำแหน่งใด ใกล้หรือไกลจากตัวผู้ใช้น้อยแค่ไหน เพื่อให้เป็นอีกทางเลือกหนึ่งในการช่วยตัดสินใจสำหรับผู้ใช้

ทัศนวรรณ แก้วใส และสุพจน์ นิตย์สุวัฒน์ (2552) ได้พัฒนาระบบแนะนำภาพยนตร์ด้วยเทคนิคตัวกรองเชิงร่วมมือร่วมกับวิธีเคมีน ซึ่งเป็นระบบการแนะนำภาพยนตร์โดยใช้เทคนิคตัวกรองเชิงร่วมมือมาประยุกต์เพื่อแก้ปัญหา Information Overload และนำเสนอการแก้ปัญหาของเทคนิคตัวกรองเชิงร่วมมือด้านขนาดของข้อมูลด้วยอัลกอริทึมแบบเคมีน แบ่งข้อมูลเป็นกลุ่มก่อนเข้าสู่กระบวนการกรองเชิงร่วมมือ ต่อมาเสนอการออกแบบระบบและพัฒนาระบบด้วยภาษา PHP และ MySQL ทำการประเมินความพึงพอใจของผู้ใช้ต่อระบบด้วยแบบประเมินตามวิธีของไลเคอร์ท พบว่าผู้ใช้งานมีความพึงพอใจต่อระบบอยู่ในระดับดี และระบบสามารถนำไปใช้งานได้จริง

อาคม รัตนศิริ และสุพจน์ นิตย์สุวัฒน์ (2552) ได้พัฒนาระบบแนะนำโฆษณาสำหรับโปรแกรมสร้างรูปสไลด์โชว์โดยใช้เทคนิค Text Classification เป็นงานวิจัยเพื่อศึกษาระบบช่วยในการแนะนำโฆษณาผ่านเว็บไซต์ (Advertisements Recommend) โดยระบบดังกล่าวสามารถที่จะช่วยให้เกิดการสื่อสารกับผู้บริโภคได้อย่างตรงกลุ่มเป้าหมาย โดยทางผู้จัดทำได้นำหลักการดังกล่าวเข้ามาใช้งานกับระบบสร้างรูปสไลด์โชว์โดยใช้เทคนิค Naïve Bayesian Classification อันเป็นการจัดกลุ่มข้อมูลส่วนบุคคล รวมทั้งข้อมูลของรูปใช้ข้อมูลจาก Tag Metadata ที่ผู้ใช้งานระบบได้นำไปจัดทำเป็นสไลด์ เพื่อนำมาใช้ในการพิจารณาในการจับคู่ผู้ใช้งานกับโฆษณาที่มีอยู่ในระบบ อันจะก่อให้เกิดประสิทธิภาพของการใช้งาน โฆษณาได้อย่างตรงกลุ่มเป้าหมายของผู้ใช้งานอย่างแท้จริง