รายงานการปฏิบัติงานสหกิจศึกษา

ระบบแนะนำสินค้า

PRODUCT RECOMMENDATION SYSTEM

โดย

นายปรินทร์ กิตติโสภณธรรม 6305000007

รายงานนี้เป็นส่วนหนึ่งของวิชา สหกิจศึกษา

ภาควิชาธุรกิจดิจิทัล

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

ภาคการศึกษาที่ 3 ปีการศึกษา 2565

หัวข้อโครงงาน ระบบแนะนำสินค้า

Product Recommendation System

รายชื่อผู้จัดทำ นายปรินทร์ กิตติโสภณธรรม 6305000007

ภาควิชา ธุรกิจดิจิทัล

อาจารย์ที่ปรึกษา อาจารย์ศรัญธร มั่งมี

อนุมัติให้รายงานฉบับนี้เป็นส่วนหนึ่งของการปฏิบัติงานสหกิจศึกษา ภาควิชาธุรกิจดิจิทัล ประจำภาคการศึกษาที่ 3 ปีการศึกษา 2565

คณะกรรมการสอบโครงงาน

.....อาจารย์ที่ปรึกษา

(อาจารย์ ศรัญธร มั่งมี)

.....พนักงานที่ปรึกษา

(คุณธิติวุฒ สีม่วง)

.....กรรมการกลาง

(ผู้ช่วยศาสตราจารย์ ดร.พิชญากร เลค)

.....ผู้ช่วยอธิการบดีและผู้อำนวยการสำนักสหกิจศึกษา

(ผู้ช่วยศาสตราจารย์ ดร.มารุจ ลิมปะวัฒนะ)

จดหมายนำส่งรายงาน

วันที่ 21 เดือน สิงหาคม พ.ศ. 2566

เรื่อง ขอส่งรายงานการปฏิบัติงานสหกิจศึกษา

เรียน อาจารย์ที่ปรึกษาสหกิจศึกษา ภาควิชาธุรกิจดิจิทัล

อาจารย์ศรัญธร มั่งมี

ตามที่ ผู้จัดทำ นายปรินทร์ กิตติโสภณธรรม นักศึกษาภาควิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศมหาวิทยาลัยสยาม ได้ไปปฏิบัติงานสหกิจศึกษาระหว่างวันที่ 22 พฤษภาคม 2566 ถึงวันที่ 1 กันยายน 2566 ในตำแหน่ง IT ณ บริษัทยูไนเต็ดฟูดส์และได้รับมอบหมายงานจากพนักงานที่ปรึกษาให้ศึกษาทำรายงาน เรื่องระบบแนะนำสินค้าบัดนี้ การปฏิบัติงานสหกิจศึกษาได้สิ้นสุดแล้ว นายปรินทร์ กิตติโสภณธรรม จึงขอส่งรายงานดังกล่าวมาพร้อมกันนี้ จำนวน 1 เล่ม เพื่อขอรับคำปรึกษาต่อไป

จึงเรียนมาเพื่อโปรดพิจารณา

ขอแสดงความนับถือ

นายปรินทร์ กิตติโสภณธรรม

นักศึกษาสหกิจศึกษา

ภาควิชาธุรกิจดิจิทัล

คณะเทคโนโลยีสารสนเทศ มหาลัยสยาม

กิตติกรรมประกาศ

(Acknowledgement)

การที่ผู้จัดทำได้มาปฏิบัติงานในโครงการสหกิจศึกษา ณ บริษัทยูไนเต็ดฟูดส์ ตั้งแต่วันที่ 22 พฤษภาคม 2566 ถึงวันที่ 1 กันยายน 2566 ส่งผลให้ผู้จัดทำรับความรู้และประสบการณ์ต่างๆ ที่มีค่ามากมาย สำหรับรายงานสหกิจศึกษาฉบับนี้ สำเร็จลงได้ด้วยดีจากความร่วมมือและสนับสนุนจากหลายฝ่าย ดังนี้

1. คุณปรีชา ตั้งเกียรติกำจาย ผู้จัดการฝ่าย IT
2. คุณธิติวุฒ สีม่วง พนักงานที่ปรึกษา
3. อาจารย์ศรัญธร มั่งมี อาจารย์ที่ปรึกษา

และบุคคลท่านอื่นๆ ที่ไม่ได้กล่าวนามทุกท่านที่ได้ให้คำแนะนำช่วยเหลือในการจัดทำรายงาน

ผู้จัดทำ ขอขอบพระคุณผู้ที่มีส่วนเกี่ยวข้องทุกท่านที่มีส่วนร่วมในการให้ข้อมูลและเป็นที่ปรึกษาในการทำรายงานฉบับนี้จนเสร็จสมบูรณ์ ตลอดจนให้การดูแลและให้ความเข้าใจกับชีวิตของการทำงานจริง ซึ่งผู้จัดทำขอขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้ด้วย

คณะผู้จัดทำ
นายปรินทร์ กิตติโสภณธรรม
21 สิงหาคม 2566

หัวข้อโครงงาน : ระบบแนะนำสินค้า

หน่วยกิต : 5

คณะผู้จัดทำ : นายปรินทร์ กิตติโสภณธรรม

อาจารย์ที่ปรึกษา : อาจารย์ศรัญธร มังมี

ชื่อปริญญา : บริหารธุรกิจบัณฑิต

ภาควิชา : ธุรกิจดิจิทัล

คณะ : เทคโนโลยีสารสนเทศ

ภาคการศึกษา / ปีการศึกษา : 3/2565

บทคัดย่อ

บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน) เป็นบริษัทชั้นนำในการผลิตขนม ของประเทศไทย โดยมีผลิตภัณฑ์ขนมชนิดต่างๆเช่น เวเฟอร์ เซี่ยงไฮ้, ช็อกโกแลต อัลมอนด์ , เยลลี่ โยโย่ , โตโร่ป๊อปคอร์นและอื่นๆการเพิ่มยอดขายและนำเสนอสินค้าให้ตรงตามความต้องการของผู้บริโภคเป็นช่องทางที่ทำให้เพิ่มรายได้ให้กับผู้ประกอบการและเพิ่มความพึงพอใจให้กับผู้บริโภคที่มีต่อสินค้ามากยิ่งขึ้นวัตถุประสงค์ของงานวิจัยนี้คือการพัฒนาแบบจำลองการแนะนำสินค้ามาช่วยในการแนะนำสินค้าโดยใช้เทคนิคการกรองแบบอิงเนื้อหา (Content-Based Filtering) ซึ่งใช้หลักการความคล้ายคลึงแบบโคไซน์และเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) โดยนำไลบรารีของ Non-negative matrix factorization มาสร้างแบบจำลองทำให้แนะนำสินค้าให้กับลูกค้าได้อย่างเหมาะสม

คำสำคัญ : ระบบแนะนำ, การกรองแบบอิงเนื้อหา, การกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

Project Title	: Product Recommendation System
Credits	: 5
By	: Mr. Parin Kittisophontham
Advisor	: Miss Sarathon Maungmee
Degree	: Bachelor of Business Administration
Major	: Digital Business
Faculty	: Information Technology
Semester / Academic year	: 3/2022

Abstract

United Foods Public Company Limited is a leading confectionery manufacturing company in Thailand. offering a wide range of confectionery products such as Wafer Shanghai, Almond Chocolate, Yo-Yo Jelly and Toro Popcorn. Selling and presenting products to meet the needs of consumers is a way to increase revenue for entrepreneurs and increases consumer satisfaction with products. The objective of this research was to develop a product recommendation model to assist with content-based filtering techniques. Content-based filtering, uses the cosine similarity principle and user-dependent filtering techniques. Collaborative filtering uses a library of non-negative matrix factorization. that creates a model to appropriately recommend products to customers.

Keywords: recommendation system, content-based filtering, collaborative filtering

Appproved by

.....

สารบัญ

หน้า

จดหมายนำส่งรายงาน ..... ก

กิตติกรรมประกาศ ..... ข

บทคัดย่อ ..... ค

Abstract ..... ง

บทที่ 1 บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตของโครงงาน 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 2

บทที่ 2 การทบทวนวรรณกรรมที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง 3

2.2 งานวิจัยที่เกี่ยวข้อง 17

2.3 เว็บไซต์ที่เกี่ยวข้อง 21

บทที่ 3 รายละเอียดการปฏิบัติงาน

3.1 ชื่อและที่ตั้งของสถานประกอบการ 23

3.2 ลักษณะการประกอบและการให้บริการหลัก 24

3.3 รูปแบบการจัดองค์การและการบริหารงานขององค์กร 24

3.4 ตำแหน่งและลักษณะงานที่ได้รับมอบหมาย 25

3.5 ชื่อและตำแหน่งงานของพนักงานที่ปรึกษา 25

3.6 ระยะเวลาปฏิบัติงาน 25

3.7 ขั้นตอนและวิธีการดำเนินงาน 25

3.8 อุปกรณ์และเครื่องมือที่ใช้ 26

บทที่ 4 รายละเอียดของโครงงาน

4.1 ขั้นตอนการใช้งาน 27

4.2 ขั้นตอนการวิเคราะห์ข้อมูล 27

บทที่ 5 สรุปผลและข้อเสนอแนะ

5.1 สรุปผลโครงงาน 37

5.2 สรุปผลการปฏิบัติงานสหกิจศึกษา 37

สารบัญ (ต่อ)

หน้า

บรรณานุกรม.....38

ภาคผนวก ก.....40

ประวัตินักศึกษาสหกิจศึกษา.....43

สารบัญตาราง

หน้า

ตารางที่ 3.1 ระยะเวลาในการดำเนินงาน 25

สารบัญรูปภาพ

หน้า

รูปที่ 2.1 ระบบคำแนะนำของ Netflix 3
รูปที่ 2.2 วิธีการสำหรับระบบแนะนำ 4
รูปที่ 2.3 ตัวอย่างของเทคนิคการกรองแบบอิงเนื้อหา 4
รูปที่ 2.4 ตัวอย่างของคำแนะนำภาพยนตร์สำหรับผู้ใช้งาน Netflix 5
รูปที่ 2.5 การสร้าง User Vector 5
รูปที่ 2.6 เวกเตอร์รายการสำหรับภาพยนตร์ 6
รูปที่ 2.7 การคูณเวกเตอร์ 2 มิติ 6
รูปที่ 2.8 ตัวอย่างการแนะนำแบบ User-based 7
รูปที่ 2.9 ตัวอย่างการแนะนำแบบ Item-based 8
รูปที่ 2.10 เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมแต่ละประเภท 8
รูปที่ 2.11 ความสัมพันธ์ระหว่างผู้ใช้งาน 9
รูปที่ 2.12 ความสัมพันธ์ระหว่างสินค้า 10
รูปที่ 2.13 เวกเตอร์แต่ละผู้ใช้งานที่ให้ความสำคัญกับคุณลักษณะของโรงแรม 10
รูปที่ 2.14 คุณลักษณะของโรงแรม 11
รูปที่ 2.15 ตัวอย่างการค้าดเดาผู้ใช้งาน B จะมีความชื่นชอบโรงแรมใดระบบ 12
รูปที่ 2.16 เมทริกซ์ของผู้ใช้งานและการให้คะแนน 12
รูปที่ 2.17 ตัวอย่างการแยกตัวประกอบของเมทริกซ์ 13
รูปที่ 2.18 การหาเมทริกซ์ของผู้ใช้งาน X และเมทริกซ์ของสินค้า Y 14
รูปที่ 2.19 ตัวอย่างของระบบแนะนำแบบผสมผสาน 14
รูปที่ 2.20 การหาความคล้ายคลึงแบบโคไซน์ 15
รูปที่ 2.21 การประยุกต์ใช้การหาความคล้ายคลึงแบบโคไซน์ 16
รูปที่ 2.22 ตัวอย่างการลดมิติของวิธี Non-negative Matrix Factorization 16
รูปที่ 2.23 แดชบอร์ดของระบบแนะนำ 18
รูปที่ 2.24 ขั้นตอนระบบการแนะนำหลักสูตรแบบผสมผสาน 19
รูปที่ 2.25 โครงสร้างของระบบแนะนำแบบผสมผสาน 20
รูปที่ 2.26 https://www.w3schools.com/python/default.asp 21
รูปที่ 2.27 https://colab.research.google.com/?hl=th 22
รูปที่ 2.28 https://www.w3schools.com/sql/default.asp 22
รูปที่ 3.1 แผนที่บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน) 23

สารบัญรูปภาพ (ต่อ)

หน้า

รูปที่ 3.2 โครงสร้างการจัดองค์การและการบริหารงานขององค์กร 24

รูปที่ 4.1 ข้อมูลลูกค้าและรายการสินค้าการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม 28

รูปที่ 4.2 ข้อมูลลูกค้าการกรองแบบอิงเนื้อหา.....28

รูปที่ 4.3 การสร้างเมทริกซ์และจัดรูปแบบข้อมูล 29

รูปที่ 4.4 แปลงข้อมูลที่ NaN ให้เป็นค่า 0.0 30

รูปที่ 4.5 นำข้อมูลเข้าอัลกอริทึม NMF 30

รูปที่ 4.6 การคำนวณหาค่าเมทริกซ์ตัวที่ 1 31

รูปที่ 4.7 การคำนวณหาค่าเมทริกซ์ตัวที่ 2 31

รูปที่ 4.8 การคำนวณเมทริกซ์ตัวที่ 1 กับ เทริกซ์ตัวที่ 2.....32

รูปที่ 4.9 ผลลัพธ์ที่ได้จากเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม 32

รูปที่ 4.10 เพิ่มข้อมูลลูกค้าที่จะนำมาทดสอบ 33

รูปที่ 4.11 จัดการสร้างเมทริกซ์และจัดรูปแบบข้อมูล.....34

รูปที่ 4.12 ทำการคัดลอกตารางเดิมมาใช้งาน.....34

รูปที่ 4.13 ทำการลบข้อมูลข้อมูลในแถว 35

รูปที่ 4.14 ทำการดึงข้อมูลลูกค้าที่สมมุติมาใส่ในตาราง 35

รูปที่ 4.15 ผลลัพธ์ที่ได้จากเทคนิคการกรองข้อมูลแบบอิงเนื้อหา 36

รูปที่ ก.1 ออกแบบ UX UI หน้าเว็บไซต์ 40

รูปที่ ก.2 เขียนโค้ดระบบแนะนำสินค้า 41

รูปที่ ก.3 ลองใช้เครื่องมือต่างๆในการวิเคราะห์ข้อมูล 41

รูปที่ ก.4 เลือกข้อมูลจาก Database.....42

รูปที่ ก.5 คัดเลือกข้อมูลลง Excel 42

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เนื่องจากบริษัทยูไนเต็ดฟูดส์เป็นบริษัทในเครือที่ใช้สำหรับขายสินค้าของบริษัทยูไนเต็ดในรูปแบบแคชแวนหรือพอยท์ ออฟ เซลล์ โดยการขับรถไปหาลูกค้า ซึ่งภายในหน่วยงานมีพนักงานและข้อมูลลูกค้า อยู่เป็นจำนวนมาก จากที่ได้ร่วมงานกับแผนก ไอที และได้พบปัญหาต้องการเพิ่มยอดขาย เพื่อเป็นการเพิ่มยอดขายให้กับผู้ประกอบการและนำเสนอผลิตภัณฑ์ที่มีอยู่ให้ตรงตามความต้องการลูกค้าและช่วยให้ลูกค้าตัดสินใจซื้อผลิตภัณฑ์ได้ง่ายขึ้น ระบบแนะนำสินค้าจึงเป็นวิธีที่จะช่วยให้ผู้ประกอบการสามารถเพิ่มรายได้และช่วยให้ผู้บริโภคได้รับสินค้าตรงตามความชื่นชอบหรือสนใจ

ดังนั้นผู้จัดทำจึงมีแนวความคิดที่จะสร้างระบบแนะนำ (Recommendation System) คือระบบที่ทำการแนะนำสินค้าหรือบริการช่วยเพิ่มความสะดวกให้กับผู้ใช้งานในการเลือกซื้อสินค้าและบริการสามารถเพิ่มยอดขายให้กับทางบริษัทโดยระบบจะแนะนำสินค้าที่คาดว่าผู้ใช้ให้ความสนใจ ระบบแนะนำจึงมีความสำคัญในการกระตุ้นยอดขายซึ่งถูกนำมาใช้อย่างแพร่หลายบนเว็บไซต์หรือแอปพลิเคชัน

จากปัญหาที่กล่าวมาข้างต้นผู้จัดทำจึงนำเอา ระบบแนะนำสินค้า เข้ามาประยุกต์ใช้กับข้อมูลลูกค้าและข้อมูลสินค้าภายในบริษัทยูไนเต็ดฟูดส์ ทำให้สามารถแนะนำสินค้าให้กับลูกค้าได้ง่าย สะดวกรวดเร็วยิ่งขึ้น ดังนั้นระบบแนะนำสินค้าจึงถูกสร้างขึ้นเพื่ออำนวยความสะดวกให้แก่พนักงานขายภายในบริษัทยูไนเต็ดฟูดส์

1.2 วัตถุประสงค์ของโครงการ

1.2.1 เพื่อเพิ่มความสะดวกสบายในการขายของพนักงานขาย แนะนำสินค้าให้ลูกค้าใหม่และเก่าได้อย่างรวดเร็ว

1.2.2 เพื่อเพิ่มยอดขายจากการแนะนำสินค้าที่ลูกค้าไม่เคยซื้อ

1.3 ขอบเขตของโครงงาน

1.3.1 ผู้ใช้สามารถแนะนำสินค้าให้กับลูกค้าประจำได้

1.3.2 ผู้ใช้สามารถแนะนำสินค้าให้กับลูกค้าใหม่ได้

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 เพิ่มยอดขายให้กับบริษัทและพนักงานสามารถทำงานได้อย่างสะดวก และง่ายขึ้นกว่าระบบเดิม

1.4.2 นำระบบแนะนำสินค้าไปประยุกต์ใช้ในการเพิ่มยอดขายสินค้าในรายการอื่นได้

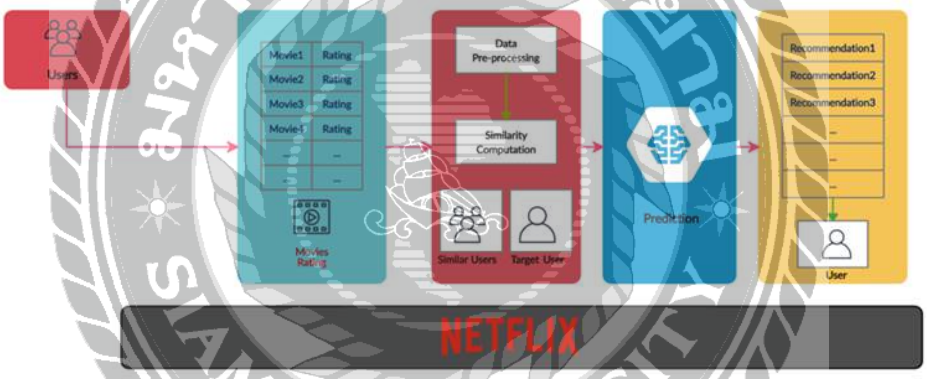
มหาวิทยาลัยสยาม SIAM UNIVERSITY

บทที่ 2
แนวคิดทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2.1 แนวความคิด

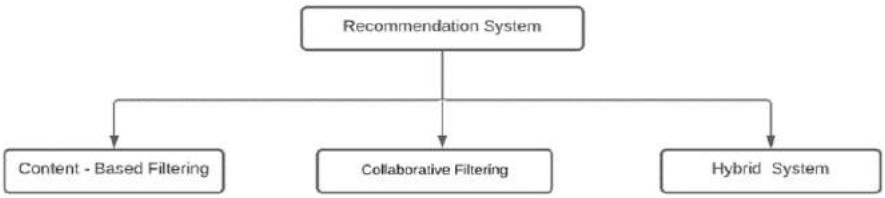
2.1.1 ระบบแนะนำ (Recommendation System)

ระบบแนะนำคือระบบที่มีไว้เพื่อแนะนำสินค้าหรือบริการให้กับผู้ใช้งานซึ่งจะช่วยแนะนำสินค้าให้กับผู้ใช้งานในการเลือกซื้อสินค้าและบริการต่าง ๆ ระบบจะทำหน้าที่ช่วยในการแนะนำสินค้าโดยอ้างอิงจากข้อมูลและประวัติการใช้งานของผู้ใช้บริการระบบจะทำการเรียนรู้จากพฤติกรรมประวัติการใช้งาน เพื่อที่จะทำการแนะนำสินค้าให้ตรงตามความต้องการของผู้ใช้งาน ทำให้เกิดประโยชน์ในด้านการใช้งานโดยที่ผู้ใช้งานจะไม่เสียเวลาในการค้นหาสินค้าหรือบริการ ซึ่งระบบแนะนำได้ถูกนำมาใช้งานเพื่อแนะนำสินค้าและบริการอย่างมาก (Saniya Parveez, 2020)



ภาพประกอบ 2.1 ระบบคำแนะนำของ Netflix

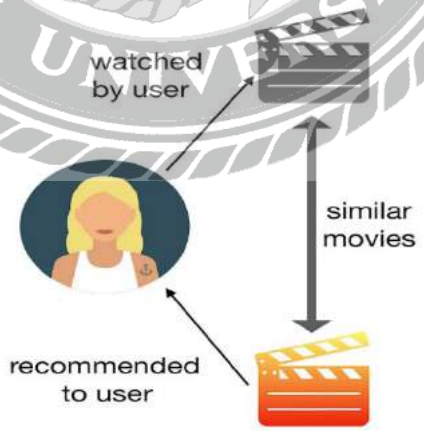
ที่มา : https ://pub.towardsai.net/recommendation-system-in-depth-tutorial-with-python-for-netflix-using-collaborative-filtering-533ff8a0e444



ภาพประกอบ 2.2 วิธีการสำหรับระบบแนะนำ

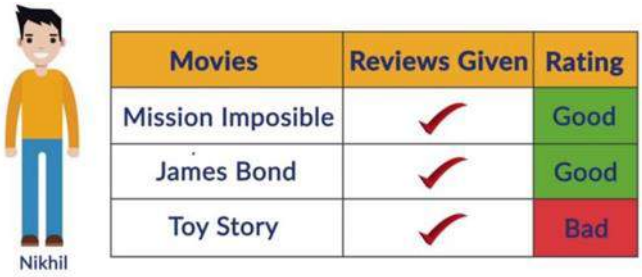
2.1.2 การกรองแบบอิงเนื้อหา (Content-Based Filtering)

เทคนิคการกรองแบบอิงเนื้อหาคือการแนะนำสินค้าหรือบริการที่ได้มาจากการพิจารณาความคล้ายคลึงของพฤติกรรมจากประวัติการใช้งานของผู้ใช้งาน โดยจะดูจากลักษณะของสินค้าที่จะแนะนำและจะทำการแนะนำสินค้าที่มีลักษณะใกล้เคียงที่ผู้ใช้งานเคยใช้หรือชื่นชอบ จากภาพประกอบ 4 ระบบจะทำการแนะนำเนื้อหาของหนังที่มีความคล้ายคลึงกับหนังที่ผู้ใช้งานเคยดูมาก่อนหน้านั้น ซึ่งข้อดีของเทคนิคการกรองแบบอิงเนื้อหานั้น คือไม่ต้องใช้ข้อมูลจำนวนมากในการประมวลผลเพื่อแนะนำสินค้า และการแนะนำสินค้าใหม่ที่ยังไม่มีข้อมูลสามารถทำได้เพราะจะพิจารณาจากความคล้ายคลึงกับลักษณะสินค้าประเภทเดิม ส่วนข้อเสียของเทคนิคการกรองแบบอิงเนื้อหาจะไม่สามารถแนะนำสินค้าที่มีความแตกต่างจากสินค้าประเภทที่ผู้ใช้งานเคยซื้อหรือใช้บริการทำให้ผู้ใช้งานไม่ได้รับการแนะนำสินค้าที่มีความหลากหลาย (Balu, 2019)



ภาพประกอบ 2.3 ตัวอย่างของเทคนิคการกรองแบบอิงเนื้อหา

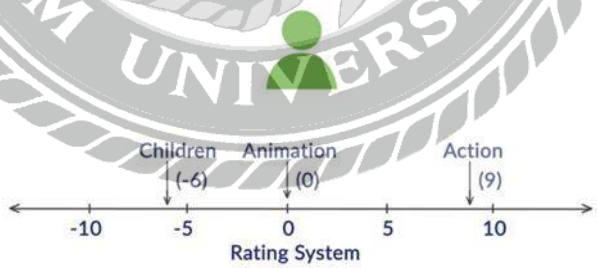
ที่มา : https ://towardsdatascience.com/how-to-build-from-scratch-a-content- based-movie-recommender-with-natural-language-processing-25ad400eb243



ภาพประกอบ 2.4 ตัวอย่างของคำแนะนำภาพยนตร์สำหรับผู้ใช้งาน Netflix

ที่มา : https ://medium.com/@bindhubalu/content-based-recommender-system-4db1b3de03e

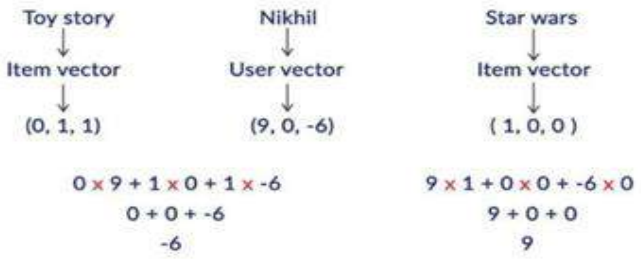
จากภาพประกอบ 2.4 ตัวอย่างของคำแนะนำภาพยนตร์สำหรับผู้ใช้งานใน Netflix โดยใช้ชื่อโปรไฟล์ Nikhil ในกรณีที่ Nikhil ให้คะแนนที่ดีกับหนังเรื่อง Mission Impossible และJames Bond ที่ถูกจัดเป็นประเภทหนังบู๊ และให้คะแนนต่ำกับหนังเรื่อง Toy Story ที่ถูกจัดเป็นประเภทหนังสำหรับเด็ก ก็จะสามารถทำการสร้างเวกเตอร์ของผู้ใช้งานสำหรับ Nikhil ตามการให้คะแนน 3 อันดับ ดังภาพประกอบ 6 ในระดับคะแนน -10 ถึง 10 เนื่องจาก Nikhil ชอบภาพยนตร์ประเภทหนังบู๊ จึงได้กำหนดค่า 9 ให้กับ ประเภทหนังบู๊ และNikhil ยังไม่ได้ดูภาพยนตร์แอนิเมชัน จึงกำหนด 0ให้กับประเภทแอนิเมชัน และเนื่องจาก Nikhil ได้ให้คำวิจารณ์ที่ไม่ดีสำหรับ ภาพยนตร์ที่ประเภทเด็กจึงกำหนด -6 ให้กับประเภทหนังสำหรับเด็ก (Prasad, 2020)



ภาพประกอบ 2.5 การสร้าง User Vector

ที่มา : https ://medium.com/@bindhubalu/content-based-recommender-system-4db1b3de03e7

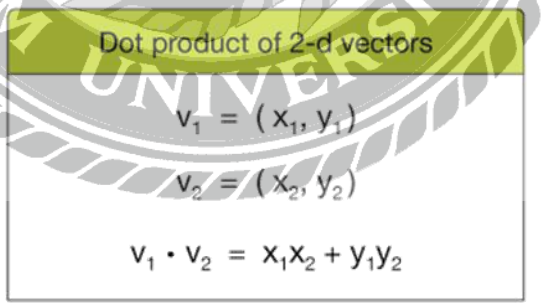
ดังนั้นเวกเตอร์ของ Nikhil จึงมีค่าเท่ากับ (9, 0, -6) ตามลำดับ (ประเภทหนังบู๊, ประเภทแอนิเมชัน, ประเภทหนังสำหรับเด็ก) เวกเตอร์สำหรับภาพยนตร์เรื่อง Toy Story คือ (0, 1, 1) และภาพยนตร์เรื่อง Star Wars คือ (1, 0, 0) ตามลำดับ ตามภาพประกอบ 2.6



ภาพประกอบ 2.6 เวกเตอร์รายการสำหรับภาพยนตร์

ที่มา : https ://medium.com/@bindhubalu/content-based-recommender-system-4db1b3de03e7

จากภาพประกอบ 2.7 การคูณเวกเตอร์ 2 มิติสองตัว คือเวกเตอร์ของรายการและเวกเตอร์ของผู้ใช้งาน ดังนั้นการคูณเวกเตอร์ของ Toy Story คือ-6 และStar Wars คือ 9 จึงได้แนะนำหนังเรื่อง Star Wars ให้กับ Nikhil ซึ่งตรงกับความชอบของ Nikhil ที่ชอบดูหนังประเภทหนังบู๊และไม่ชอบภาพยนตร์สำหรับเด็ก ในลักษณะที่คล้ายกันนั้นนสามารถคำนวณการคูณเวกเตอร์ของเวกเตอร์รายการทั้งหมดของภาพยนตร์ทั้งหมดในร้านค้าและแนะนำภาพยนตร์ 10 อันดับแรกให้กับ Nikhil



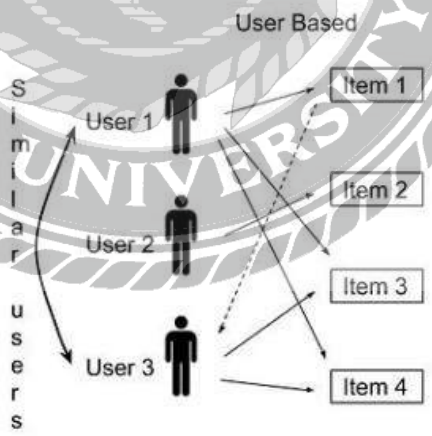
ภาพประกอบ 2.7 การคูณเวกเตอร์ 2 มิติ

ที่มา : https ://medium.com/@bindhubalu/content-based-recommender-system-4db1b3de03e7

2.1.3 การกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering)

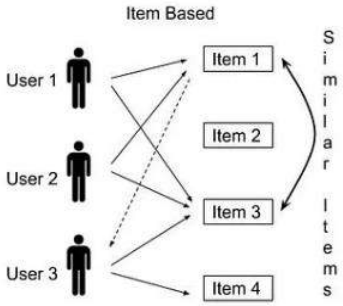
การกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม คือการแนะนำสินค้าหรือบริการโดยพิจารณาจากความคล้ายคลึงกันกับผู้ใช้งานรายอื่นมาประกอบคำแนะนำด้วย โดยนำข้อมูลจากผู้ใช้งานรายอื่นมาทำการคาดเดาว่าผู้ใช้งานรายนี้จะมีความชื่นชอบสินค้าประเภทใด ซึ่งจะใช้หลักการของ The Wisdom of the Clouds คือฐานข้อมูลจะต้องมีค่า rating ของสินค้าหรือบริการของผู้ใช้งานแต่ละคนที่ผ่านมา

ข้อมูล rating สามารถนำมาใช้ในการแนะนำสินค้าได้ 2 วิธี ได้แก่ 1. User-Based คือ ผู้ใช้งานที่ถูกจัดอยู่ในกลุ่มประเภทเดียวกัน ซึ่งจะมีความชื่นชอบสินค้าชนิดเดียวกันจากภาพประกอบ 2.8 ถ้าผู้ใช้ 1 และผู้ใช้ 3 ชอบสินค้ารายการ3 และ4 เหมือนกัน ในกรณีที่ผู้ใช้ 1 สั่งสินค้ารายการ1 เพิ่มเติม ระบบแนะนำก็จะทำการแนะรายการสินค้า 1 ให้กับ ผู้ใช้ 3 ด้วยเช่นกัน 2. Item-Based คือสินค้าที่มีลักษณะใกล้เคียงกัน เช่น ถ้าสินค้ารายการ1 และ3 มีความคล้ายคลึงกัน ในกรณีที่ผู้ใช้ 1 ตัดสินใจเลือกซื้อสินค้ารายการ1 โอกาสที่ผู้ใช้ 3 จะตัดสินใจเลือกซื้อสินค้ารายการ 1 ก็มีเช่นกัน โดยดูตัวอย่างได้จากภาพประกอบ 2.8 (Gross, 2023)



ภาพประกอบ 2.8 ตัวอย่างการแนะนำแบบ User-based

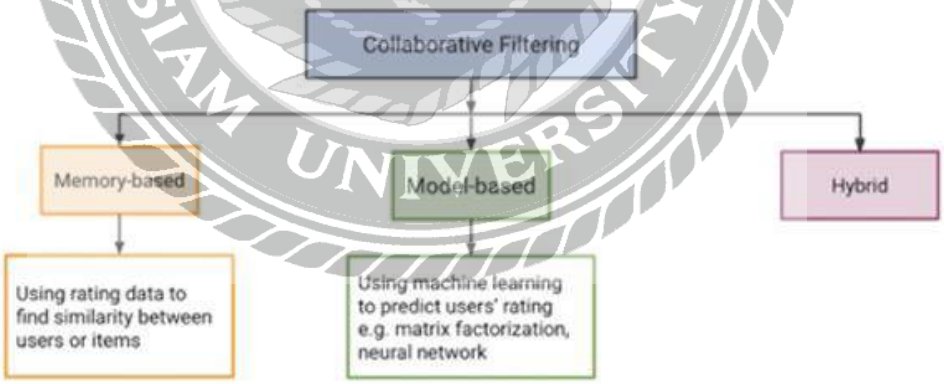
ที่มา : https ://levelup.gitconnected.com/the-mathematics-of-recommendation-systems-e8922a50bd



ภาพประกอบ 2.9 ตัวอย่างการแนะนำแบบ Item-based

ที่มา : https ://levelup.gitconnected.com/the-mathematics-of-recommendation-systems-e8922a50bdea

เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม มีการนำข้อมูลจากผู้ใช้งานรายอื่นมาช่วยในการแนะนำ ว่าผู้ใช้งานรายนี้จะชื่นชอบสินค้าประเภทใด โดยรูปแบบของเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม จะมีทั้งหมด 3 ประเภท ตามภาพประกอบ 2.10 ซึ่งงานวิจัยนี้จะมีการสร้างเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมจากประเภทของ Model-base

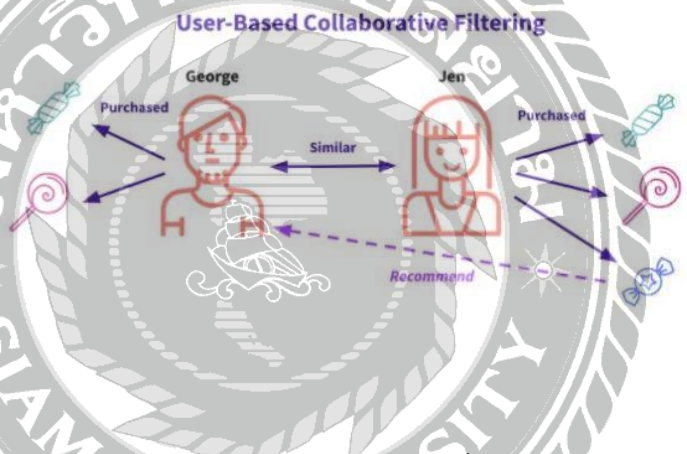


ภาพประกอบ 2.10 เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมแต่ละประเภท

ที่มา : https ://www.medium.com/@sinart.t/recommendation-system- แบบสรุปเอาเอง-ce6246f49754

2.1.4 การพิจารณาแบบจดจำ (Memory-based)

ดูจากข้อมูลและหาความสัมพันธ์ระหว่างผู้ใช้งานกับสินค้าการใช้ Nearest Neighbor ในการคาดเดา rating จากผู้ใช้งาน (user-based) หรือสินค้า (item-based) ที่มีความใกล้เคียงกัน จากภาพประกอบ 2.11 ข้อมูลจากผู้ใช้งาน (user-based) ถ้าคุณ George และคุณ Jen มีการใช้งานที่คล้ายคลึงกัน ในกรณีที่คุณ Jen ทำการสั่งซื้อสินค้าประเภทอื่นเพิ่มเติม ระบบแนะนำก็จะทำการแนะนำสินค้าชิ้นนี้ให้กับคุณ George ด้วยเช่นกัน และจากภาพประกอบ 2.12 ข้อมูลจากสินค้า เช่น ถ้าสินค้าเบอร์ 1 มีความคล้ายคลึงกับสินค้าเบอร์ 2 ในกรณีที่ผู้ใช้งานเลือกซื้อสินค้าเบอร์ 1 ก็อาจจะซื้อสินค้าเบอร์ 2 ด้วย (Tangruamsub, 2019)



ภาพประกอบ 2.11 ความสัมพันธ์ระหว่างผู้ใช้งาน

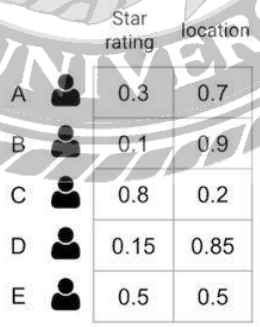
ที่มา : https ://blog.dataiku.com/recommendation-engines-how-they-work-in-plain-english

ภาพประกอบ 2.12 ความสัมพันธ์ระหว่างสินค้า

ที่มา : https ://blog.dataiku.com/recommendation-engines-how-they-work-in-plain-english

2.1.5 การพิจารณาแบบจำลอง (Model-based)

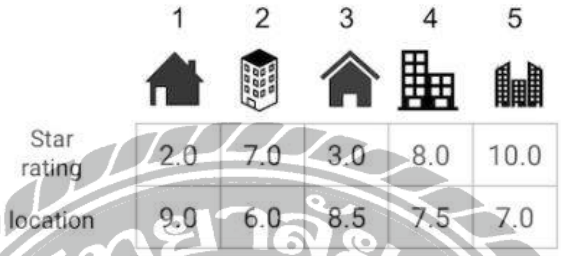
มีการใช้ Machine Learning เพื่อหา user embedding และ item embedding มาทำนาย rating ที่ผู้ใช้งานจะให้คะแนนกับสินค้า Model-based ส่วนใหญ่จะอยู่ในรูปแบบการใช้แบบจำลองหา user weight และ item weight เพื่อให้สามารถประเมินความชื่นชอบของผู้ใช้งานได้จาก ภาพประกอบ 2.13 ตัวอย่างเว็บไซต์การจองโรงแรม ในการเลือกโรงแรมนั้นผู้ใช้งานมักจะมีเกณฑ์ในการตัดสินใจเลือกโรงแรมที่ต่างกัน เช่น บางคนให้ความสำคัญกับตำแหน่งที่ตั้งโรงแรมหรือบางคนอาจจะให้ความสำคัญของรีวิวที่พักมากกว่า เป็น (Prasad, 2020)



ภาพประกอบ 2.13 เวกเตอร์แต่ละผู้ใช้งานที่ให้ความสำคัญกับคุณลักษณะของโรงแรม

ที่มา : https ://www.medium.com/@sinart.t/recommendation-system-แบบสรุปเอาเอง-ce6246f49754

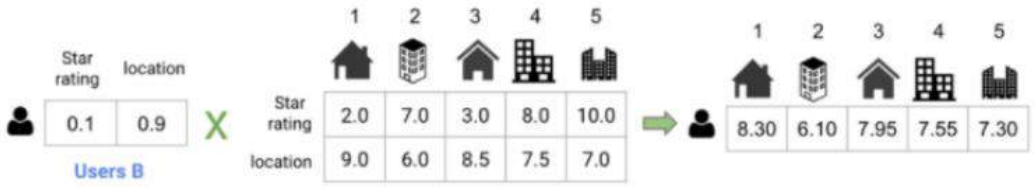
hotel features คือคุณลักษณะของโรงแรมแต่ละโรงแรมว่ามีคะแนนในแต่ละด้านเป็นอย่างไร จากภาพประกอบ 2.14 ตัวเลขในแต่ละช่อง คือคะแนนของแต่ละหมวดหมู่ เช่น โรงแรมหมายเลข 5 มีคะแนนของ star rating 10 คะแนน หรือโรงแรมหมายเลข 10 นั้นมีคะแนนในส่วนของ star rating อยู่ที่ 2 คะแนน ซึ่งไม่ดีมากนัก แต่ว่ามีคะแนนในส่วนของ location ที่ดี



ภาพประกอบ 2.14 คุณลักษณะของโรงแรม

ที่มา : https ://www.medium.com/@sinart.t/recommendation-system-แบบสรุปเอาเอง-ce6246f49754

จากข้อมูลเบื้องต้นนั้นสามารถทำการคาดเดาได้ว่าผู้ใช้งานแต่ละคนนั้นจะมีความชื่นชอบโรงแรมใดบ้าง เช่น ผู้ใช้งาน B มีความชื่นชอบโรงแรมหมายเลข 1เนื่องจากผู้ใช้งาน B ให้ความสำคัญกับ location และโรงแรมหมายเลข 1 ก็เป็นโรงแรมที่มี location ดีมาก ส่วนผู้ใช้งาน C อาจจะมีความชื่นชอบโรงแรมหมายเลข 5 เนื่องจากผู้ใช้งาน C ชอบโรงแรมที่มีคะแนน star rating ที่สูง ซึ่งการคาดเดาลักษณะแบบนี้สามารถนำ weight ของ users มาคูณกับคะแนนด้านต่าง ๆ ของโรงแรมได้เช่น เมื่อนำข้อมูลผู้ใช้งาน B มาทำการคาดเดาว่าในบรรดาทั้ง 5โรงแรมนั้นผู้ใช้งาน B จะมีความชื่นชอบโรงแรมใดมากที่สุด และนำ weight vector ของผู้ใช้งาน B มาคูณกับ hotel features ตามภาพประกอบ 2.15 (Tangruamsub, 2019)



ภาพประกอบ 2.15 ตัวอย่างการค้าดเดาผู้ใช้งาน B จะมีความชื่นชอบโรงแรมใด

ที่มา : https

://www.medium.com/@sinart.t/recommendation-system-แบบสรุปเอา

เอง-ce6246f49754

2.1.6 การแยกตัวประกอบเมทริกซ์ (Matrix Factorization)

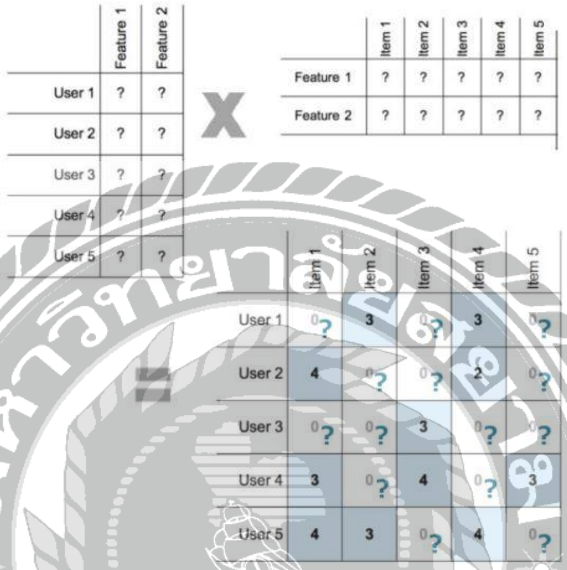
Matrix Factorization คือการแยกตัวประกอบของเมทริกซ์ให้ได้เมทริกซ์ย่อยลงไป โดยที่ผลคูณของเมกทริกซ์ย่อยนั้นจะได้เท่ากับเป็นเมทริกซ์ตั้งต้น โดยทั่วจะใช้เป็นเมทริกซ์ของrating หรือการให้คะแนน rating ของแต่ละผู้ใช้งาน แต่ละสินค้านั้นจะทำการแยกเมทริกซ์ rating นี้ออกเป็น 2 ส่วน ได้แก่ ส่วนที่เป็นเมทริกซ์ของผู้ใช้งาน X และเมทริกซ์ของสินค้า Y โดยเมื่อนำเมทริกซ์ทั้ง X และY มาคูณกันก็จะได้กลับมาเป็นเมทริกซ์ rating ตั้งต้น (Bailey, 2016)

	Item 1	Item 2	Item 3	Item 4	Item 5
User 1	0	3	0	3	0
User 2	4	0	0	2	0
User 3	0	0	3	0	0
User 4	3	0	4	0	3
User 5	4	3	0	4	0

ภาพประกอบ 2.16 เมทริกซ์ของผู้ใช้งานและการให้คะแนน

ที่มา : http ://katbailey.github.io/post/matrix-factorization-with-tensorflow/

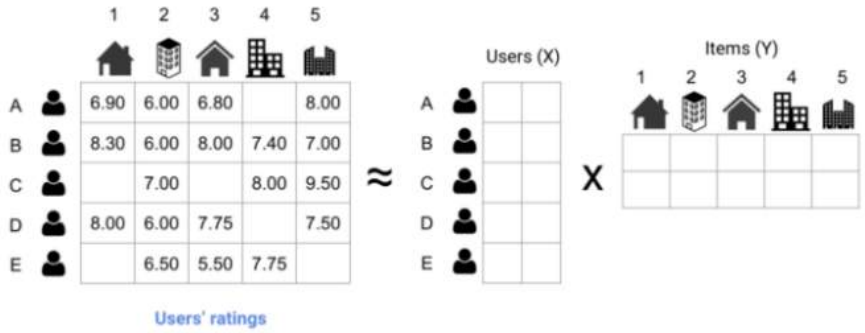
จากภาพประกอบ 2.16 คือเมทริกซ์ของการให้คะแนนที่มีผู้ใช้งานเป็นแถวและรายการเป็นคอลัมน์ ซึ่งการแยกตัวประกอบเมทริกซ์ของเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมนั้นจะมีเมทริกซ์ 2 เมทริกซ์คือ users และitems โดยที่เมทริกซ์ของ user กับ item จะมี ลักษณะที่จะบ่งบอกถึงความิชอบหรือความใกล้เคียง และเมื่อนำเมทริกซ์มาคูณกันแล้วจะได้ความชอบของ user ที่มีต่อ item นั้น ตามภาพประกอบ 2.17



ภาพประกอบ 2.17 ตัวอย่างการแยกตัวประกอบของเมทริกซ์

ที่มา : http ://katbailey.github.io/post/matrix-factorization-with-tensorflow/

จากภาพประกอบ 2.18 ในการเรียนรู้จะพยายามหาเมทริกซ์ของผู้ใช้งาน X และเมทริกซ์ของสินค้า Y เพื่อให้ได้ผลคูณของ X และY นั้นให้มีความใกล้เคียงกับเมทริกซ์ rating ของจริงให้ได้มากที่สุด โดยวิธีการนั้นก็สามารถใช้ SVD, PCA ,NMF หรือ Gradient Descent ในส่วนของข้อดีนั้นการใช้ user และitem embedding ทำให้แบบจำลองสามารถเรียนรู้หาการแทนข้อมูลที่เหมาะสมได้และยังช่วยลดเรื่องมิติของข้อมูลได้อีกด้วย ในส่วนข้อเสียยังคงมีปัญหาสำหรับผู้ใช้งานใหม่หรือสินค้าใหม่ เช่น ในกรณีที่จะทำระบบแนะนำโรงแรมให้ผู้ใช้งาน จะไม่สามารถใช้ข้อมูลอื่น ๆ ที่ผู้ใช้งานจะใช้ในการค้นหาที่พัก ไม่ว่าจะเป็นจำนวนคนเข้าพัก เมืองที่ผู้ใช้ค้นหาหรือจำนวนวันเข้าพักมาร่วมใช้ในการทำนาย (Tangruamsub, 2019)

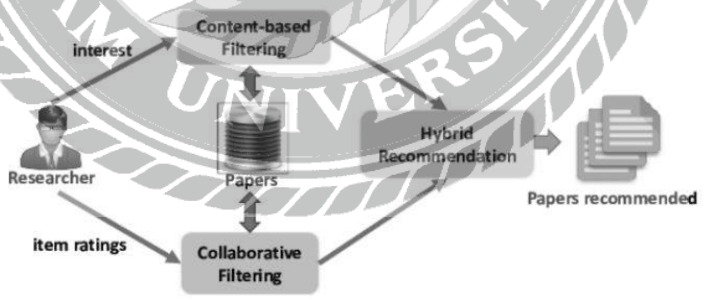


ภาพประกอบ 2.18 การหาเมทริกซ์ของผู้ใช้งาน X และเมทริกซ์ของสินค้า Y

ที่มา : https ://www.medium.com/@sinart.t/recommendation-system- แบบสรุปเอาเอง-ce6246f49754

2.1.7 ระบบแนะนำแบบผสมผสาน (Hybrid Recommendation System)

ระบบแนะนำแบบผสมผสานคือการใช้เทคนิคของระบบแนะนำสินค้ามากกว่า 2 เทคนิคขึ้นไป จากภาพประกอบ 2.19 ได้นำเทคนิคการกรองแบบอิงเนื้อหา และเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม มาทำงานร่วมกันเพื่อทำให้ระบบแนะนำมีประสิทธิภาพการทำงานที่ดีขึ้น โดยจะนำจุดเด่นของเทคนิคหนึ่งไปปรับปรุงจุดด้อยของอีกเทคนิคหนึ่ง เช่น เมื่อมีสินค้าใหม่เข้ามาและยังไม่เคยมีประวัติการให้คะแนนทำให้ไม่สามารถแนะนำด้วยเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมได้ดังนั้นจึงนำวิธีการเทคนิคการกรองแบบอิงเนื้อหามาช่วยสร้างโอกาสในการแนะนำสินค้าที่เข้ามาใหม่ (Kong, 2019)



ภาพประกอบ 2.19 ตัวอย่างของระบบแนะนำแบบผสมผสาน

ที่มา : https ://www.researchgate.net/figure/A-hybrid-paper-recommendation-system_fig5_330077673

2.1.8 วิธีการวัดความคล้ายคลึง

2.1.8.1 ความคล้ายคลึงแบบโคไซน์ (Cosine Similarity)

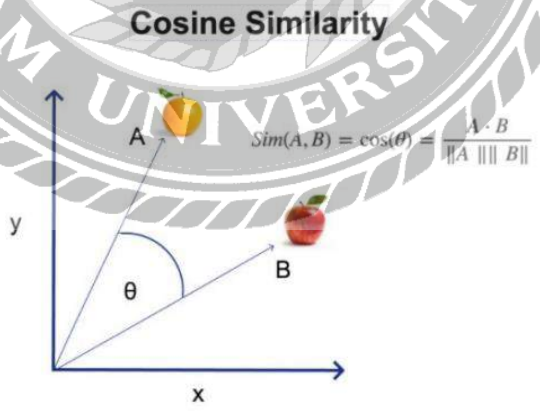
ความคล้ายคลึงแบบโคไซน์ คือการหาความเหมือนของข้อมูล โดยใช้ค่าโคไซน์จากการวัดค่าข้อมูลเป็นลายลักษณ์อักษรในเทคนิคการกรองแบบอิงเนื้อหา (Content-Based Filtering) จะถูกใช้เพื่อวัดความเหมือนระหว่างเวกเตอร์ของน้ำหนัก ค่าโคไซน์จะมีค่าอยู่ระหว่าง 0 ถึง 1 ยกตัวอย่างเช่น ค่า 1 หมายถึง ข้อมูล A และข้อมูล B มีลักษณะข้อมูลที่เหมือนกัน และส่วนของค่า 0 นั้นหมายถึงข้อมูล A และข้อมูล B มีลักษณะข้อมูลที่ไม่เหมือนกันดังสมการ (1) โดยมีที่มาจากกฎสามเหลี่ยมคือ $\cos(\theta) =$ ชิด/ฉาก โดยที่A และB คือรายการที่ต่างกัน ตามภาพประกอบ 21 (Lin, 2018)

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}, \quad (1)$$

เมื่อ A = รายการA

B = รายการB

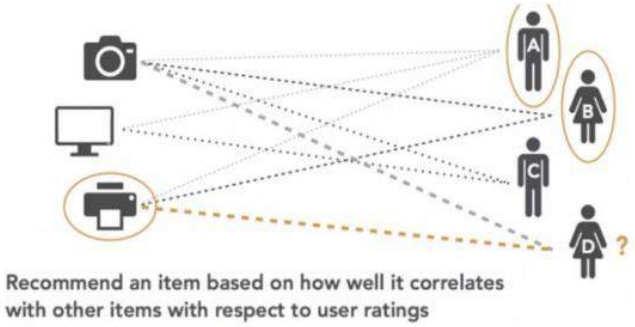
n = จำนวนคู่ของรายการA และรายการB ทั้งหมด



ภาพประกอบ 2.20 การหาความคล้ายคลึงแบบโคไซน์

ที่มา : https ://www.linkedin.com/pulse/cosine-similarity-classification-michael-lin

จากภาพประกอบ 2.21 การประยุกต์ใช้การหาความคล้ายคลึงแบบโคไซน์ การทำระบบแนะนำเพื่อดูว่าผู้ใช้งานส่วนใหญ่เมื่อซื้อสินค้า A แล้วจะมีโอกาสซื้อสินค้า B หรือ C เท่าไร



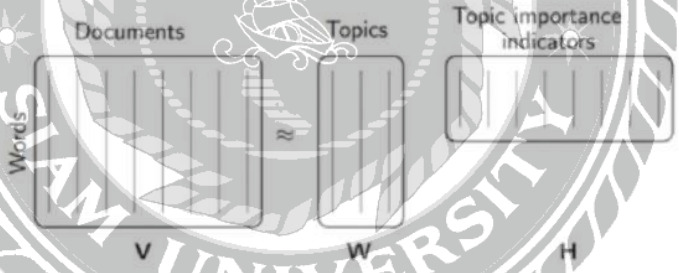
ภาพประกอบ 2.21 การประยุกต์ใช้การหาความคล้ายคลึงแบบโคไซน์

ที่มา :https ://www.softnix.co.th/2019/05/29/similarity-ความเหมือนที่แตกต่าง

2.1.9 อัลกอริทึมในการสร้างแบบจำลองเทคนิคการกรองข้อมูลแบบพึ่ง พาผู้ใช้ร่วม

2.1.9.1 Non-negative Matrix Factorization (NMF)

Non-negative Matrix Factorization (NMF) ทำหน้าที่ลดมิติของข้อมูลโดยที่ข้อมูลจะไม่มีค่าเป็นลบ ผลลัพธ์จะได้เมทริกซ์ที่เป็นค่าบวกเท่านั้น ดังภาพประกอบ 2.22 (Kanoktipsatharporn, 2019)



ภาพประกอบ 2.22 ตัวอย่างการลดมิติของวิธี Non-negative Matrix Factorization

ที่มา : https ://www.bualabs.com/archives/2971/lsa-latent-semantic-analysis-text-classification-singular-value-decomposition-svd-non-negative-matrix-factorization-nmf-nlp-ep-4

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้ได้ทำการศึกษาค้นคว้างานวิจัยที่เกี่ยวข้องกับเทคนิคการกรองแบบอิงเนื้อหา (Content-Based Filtering) และเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) โดยงานวิจัยที่เกี่ยวข้องมีรายละเอียดดังนี้

(1) บทความวิจัยเรื่อง Generating Items Recommendations by Fusing Content and User-Item based Collaborative Filtering (Tewari, 2020)

ปัจจุบันอีคอมเมิร์ซได้แพร่กระจายไปทั่วโลก ร้านค้าอิเล็กทรอนิกส์สามารถมีสินค้าได้หลายร้อย หรือหลายพันรายการโดยไม่มีขอบเขต ซึ่งข้อมูลผลิตภัณฑ์เหล่านี้ทั้งหมดมีอยู่บนอินเทอร์เน็ต อีกทั้งยังมีข้อมูลของลูกค้าเป็นจำนวนมากระบบแนะนำจะค้นหาความสนใจของผู้ใช้งาน และแนะนำรายการที่ตรงกับความต้องการของผู้ใช้งานมากที่สุด เทคนิคที่ใช้ในการแนะนำสินค้าคือเทคนิคการกรองแบบอิงเนื้อหาและเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมแนะนำรายการที่มีลักษณะคล้ายคลึงกับสินค้าที่ผู้ใช้งานเคยใช้ในอดีต ในส่วนของเทคนิค เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม สร้างกลุ่มผู้ใช้ที่คล้ายกันและแนะนำรายการให้กับผู้ใช้งาน จุดประสงค์หลักของงานวิจัยนี้คือเพื่อลดจำนวนของสินค้าที่จะนำมาแนะนำลง เพื่อช่วยในการจัดวางสินค้าที่มีความสัมพันธ์กันให้ดียิ่งขึ้น ซึ่งผลการทดลองแสดงให้เห็นว่างานที่ทำขึ้นนั้น ดีกว่าวิธีการแนะนำแบบทั่วไป

(2) บทความวิจัยเรื่อง Combining Content-based and Collaborative Filtering for Personalized Sports News Recommendations (Lenhart Philip,Daniel Herzog, 2016)

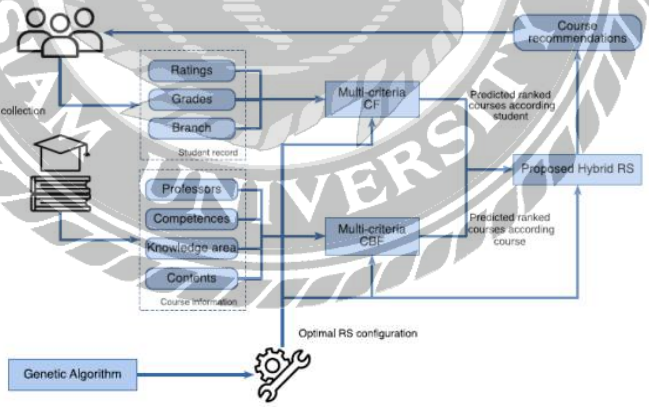
ในการนำเสนอข่าวกีฬา ผู้อ่านมักจะมาพร้อมกับอารมณ์ที่แนบแน่นกับทีมกีฬาหรือผู้เล่น นอกจากนี้ความสนใจในหัวข้อของข่าวนั้นสามารถเปลี่ยนแปลงได้อย่างรวดเร็ว เช่น มีการแข่งขันกีฬาที่สำคัญเกิดขึ้นในงานนี้ ได้นำเสนอระบบผู้แนะนำข่าวกีฬาแบบไฮบริดที่รวมคำแนะนำตามเนื้อหาเข้ากับการกรองแบบร่วมมือมีพัฒนาแดชบอร์ดผู้แนะนำและรวมเข้ากับเว็บไซต์ Sport1.de จากภาพประกอบ 2.23 การพัฒนาวิดเจ็ตแดชบอร์ดซึ่งสามารถรวมเข้าด้วยกันได้ในเว็บไซต์ที่มีอยู่เพื่อให้คำแนะนำข่าวกีฬาส่วนบุคคล โดยแสดงภาพหน้าจอปัจจุบันของแดชบอร์ดผู้แนะนำของเรา มีการนำเสนอคำแนะนำเการายการในครั้งเดียว เมื่อเลื่อนเมาส์ไปเหนือบทความหนึ่ง ผู้ใช้งานสามารถอ่าน ("Ansehen") หรือปฏิเสธคำแนะนำ ("Entfernen")

ภาพประกอบ 2.23 แดชบอร์ดของระบบแนะนำ

ที่มา : (Lenhart Philip,Daniel Herzog, 2016)

(3) บทความวิจัยเรื่อง Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization. (Esteban,Zafra,Romero, 2020)

การศึกษาในระดับมหาวิทยาลัยมักจะมีหลักสูตรรายวิชาเลือกวิชาต่าง ๆ ซึ่งจะถูกเลือกจากนักเรียนจำนวนมากในมหาวิทยาลัย นักเรียนจะใช้เวลาจำนวนมากในการค้นหาข้อมูลเกี่ยวกับหลักสูตรที่แตกต่างกันเพื่อการตัดสินใจเลือกรายวิชาให้เหมาะสมกับตัวเองมากที่สุด ในการตัดสินใจของนักเรียนอาจจะไม่ดีพอเนื่องจากนักเรียนมีข้อมูลไม่เพียงพอดังนั้นโดยทั่วไปนักเรียนอาจได้รับอิทธิพลจากความคิดเห็นของนักศึกษารายอื่น หรืออาจจะมีความชอบและความสนใจของนักเรียนมาเป็นเกณฑ์ในการพิจารณาร่วมด้วย วิธีในการแก้ไขปัญหานี้คือการใช้ระบบการแนะนำ (Recommendation System) ที่ช่วยให้นักเรียนตัดสินใจได้ดีโดยปรับตามความต้องการเฉพาะของตนเอง เทคนิคในการทำระบบแนะนำที่นิยมใช้คือเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วมและเทคนิคการกรองแบบอิงเนื้อหาจากภาพประกอบ 2.24 คือขั้นตอนของระบบการแนะนำหลักสูตรแบบผสมผสาน โดยเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม แนะนำรายการที่อิงตามการให้คะแนนของผู้ใช้ที่คล้ายกันในขณะที่เทคนิคการกรองแบบอิงเนื้อหาแนะนำรายการที่อ้างอิงเนื้อหาของรายการที่คล้ายกันในโปรไฟล์ผู้ใช้ โดยจะนำทั้งสองอัลกอริทึมนี้มาทำการแนะนำแบบผสมผสานเพื่อเพิ่มประสิทธิภาพในการแนะนำหลักสูตรรายวิชาเรียนให้กับนักเรียนได้ดีขึ้น

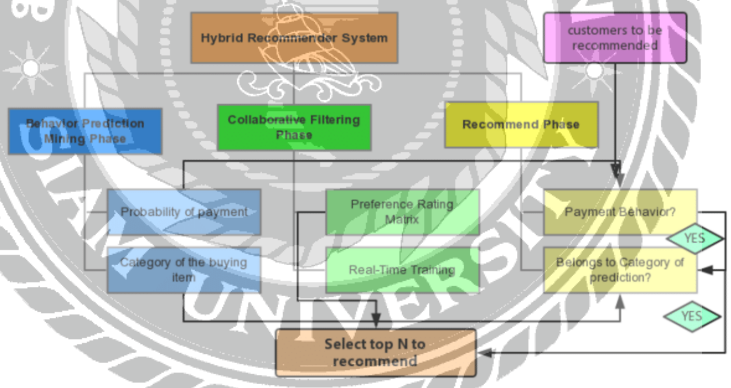


ภาพประกอบ 2.24 ขั้นตอนระบบการแนะนำหลักสูตรแบบผสมผสาน

ที่มา : (Esteban,Zafra,Romero, 2020)

(4) บทความวิจัยเรื่อง Hybrid Recommender System Based on Personal Behavior Mining (Fang, 2016)

ระบบแนะนำเป็นที่รู้จักอย่างแพร่หลายจากเว็บไซต์อีคอมเมิร์ซ และส่วนใหญ่จะเป็นในรูปแบบของแบบจำลอง อัลกอริทึมที่ใช้แนะนำความชอบส่วนบุคคล ทีรวมถึงการแนะนำแบบพึ่งพาผู้ใช้ร่วมจาก item-based ที่ใช้ใน Amazon การแนะนำแบบพึ่งพาผู้ใช้ร่วมจาก matrix factorization based ที่ใช้กับ Netflix และวิธีอื่นอีกมากมาย ในบทความนี้ผู้วิจัยจะทำการรวมแบบจำลองแบบเดิม ๆ กับวิธีแบบ pattern extraction ซึ่งใช้ข้อมูล desensitized mobile transaction record ที่ได้มาจาก T-mal ของ Alibaba จากภาพประกอบ 2.25 จะทำการบันทึกการทำธุรกรรมซึ่งจัดทำโดย T-mall, Alibaba ซึ่งมีข้อมูลการซื้อในอดีตของผู้ใช้งาน โดยจะแบ่งพฤติกรรมที่แตกต่างกันของผู้บริโภคออกเป็น 4 ประเภทได้แก่ คลิก เลือก หยิบใส่รถเข็นและชำระเงิน เพื่อสร้างระบบแนะนำแบบผสมผสานที่แนะนำได้อย่างหลากหลาย sequential pattern mining มีจุดประสงค์เพื่อหาความถี่ของ รูปแบบตามลำดับใน ลำดับของฐานข้อมูล และจะนำไปใช้ในแบบจำลองแบบผสมผสานเพื่อใช้ทำนายพฤติกรรมการจ่ายเงินของลูกค้า และมีส่วนในการเพิ่มความถูกต้องของแบบจำลอง

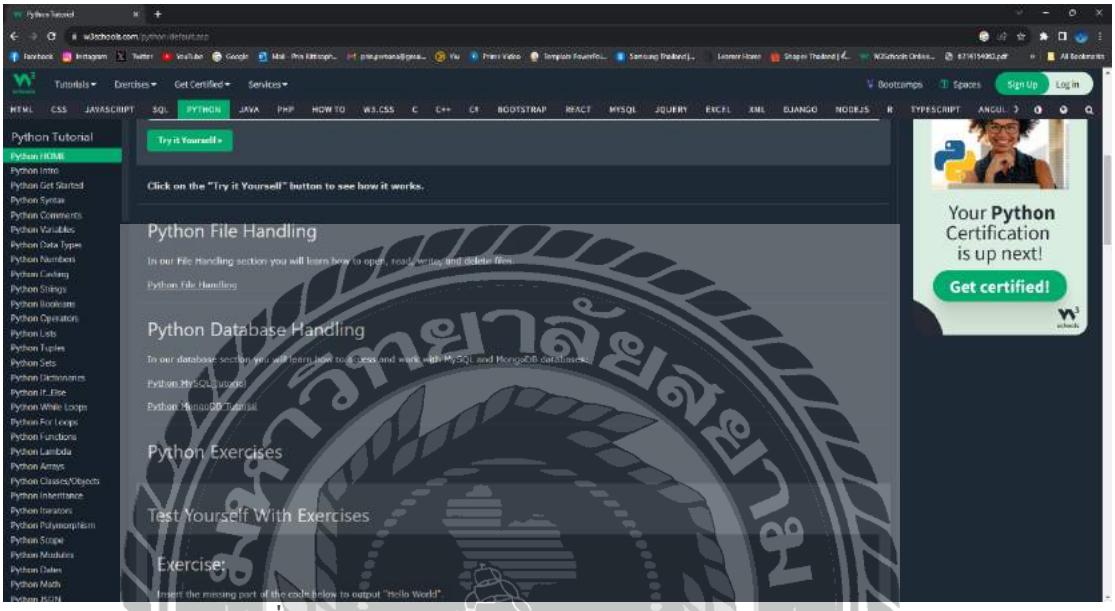


ภาพประกอบ 2.25 โครงสร้างของระบบแนะนำแบบผสมผสาน

ที่มา : (Fang, 2016)

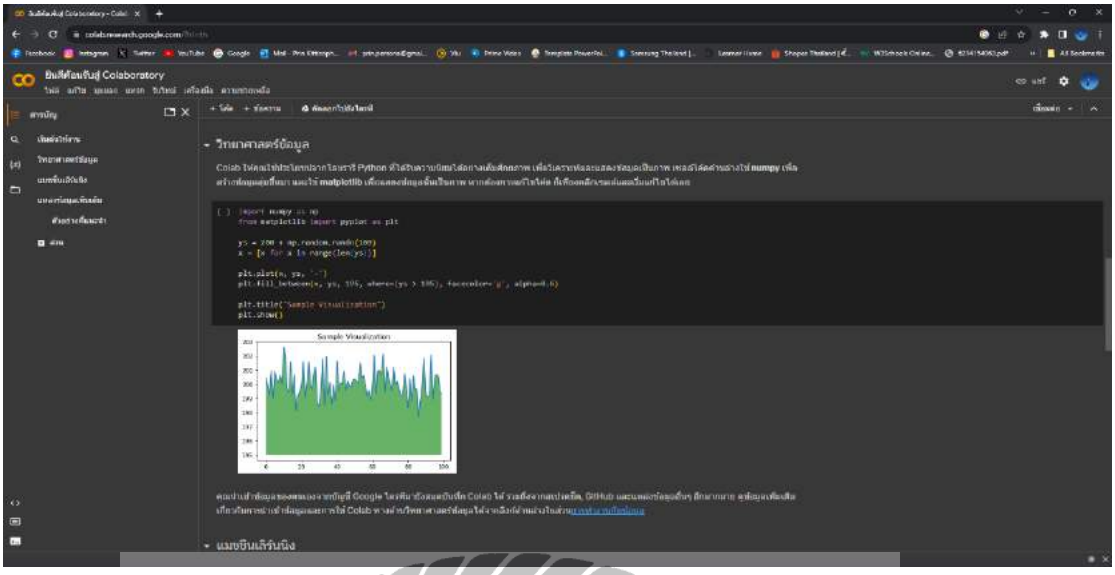
2.3 เว็บไซต์ที่เกี่ยวข้อง

ผู้จัดทำได้ทำการค้นหาเว็บไซต์ที่เกี่ยวข้องกับการพัฒนาระบบ เนื่องจากมีประโยชน์ต่อการดำเนินโครงงานในด้านการออกแบบเว็บไซต์ด้านการออกแบบฐานข้อมูลและการเขียนโปรแกรม โดยค้นหาจากเว็บไซต์ดังต่อไปนี้



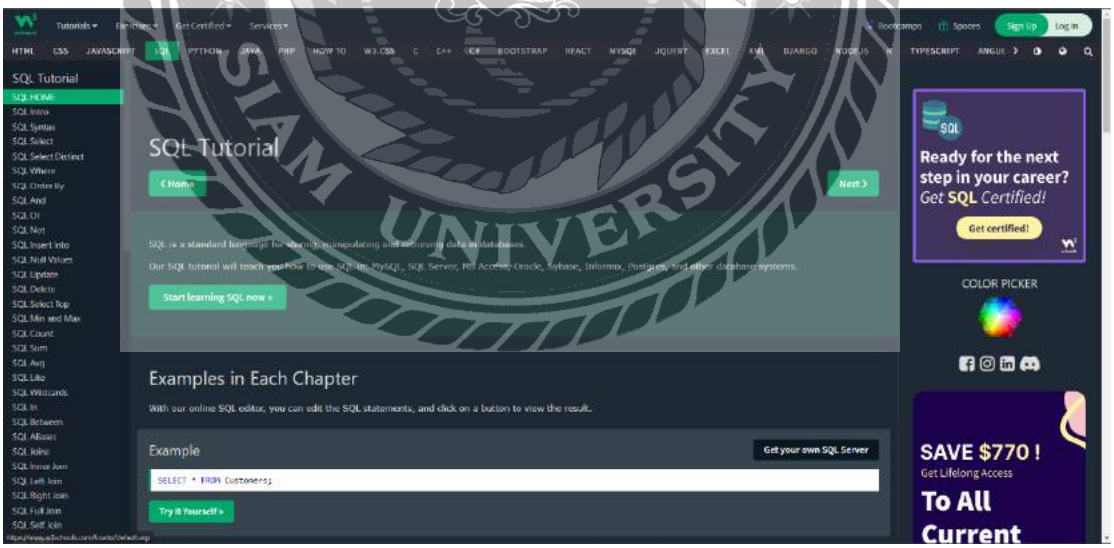
รูปที่ 2.26 https://www.w3schools.com/python/default.asp

เป็นเว็บไซต์ที่สอนเกี่ยวกับการเขียนภาษา Python ไซต์ด้วยการเขียนคำสั่ง จะแสดงทั้งคำสั่งและผลลัพธ์ซึ่งสามารถเลือกดูได้ว่าต้องการแบบใด และเว็บไซต์นี้ยังสามารถเรียนรู้ได้หลายภาษา หลายระดับความยาก ตั้งแต่ขั้นพื้นฐานจนถึงขั้นสูง จึงเป็นที่นิยมมากในการศึกษาการเขียนโปรแกรมในภาษาต่างๆ



รูปที่ 2.27 https://colab.research.google.com/?hl=th

เป็นเว็บไซต์ที่ให้คุณใช้ประโยชน์จากไลบรารี Python ที่ได้รับความนิยมได้อย่างเต็มศักยภาพ เพื่อวิเคราะห์และแสดงข้อมูลเป็นภาพ เซลล์โค้ดด้านล่างใช้ Numpy เพื่อสร้างข้อมูลสุ่มขึ้นมา และใช้ matplotlib เพื่อแสดงข้อมูลนั้นเป็นภาพ หากต้องการแก้ไขโค้ด ก็เพียงคลิกเซลล์และเริ่มแก้ไขได้เลย



รูปที่ 2.28 https://www.w3schools.com/sql/default.asp

เป็นเว็บไซต์ที่สอนเกี่ยวกับการเขียนภาษา SQL ด้วยการเขียนคำสั่ง จะแสดงทั้งคำสั่งและผลลัพธ์ซึ่งสามารถเลือกดูได้ว่าต้องการแบบใด และเว็บไซต์นี้ยังสามารถเรียนรู้ได้หลายภาษาหลายระดับความยาก ตั้งแต่ขั้นพื้นฐานจนถึงขั้นสูง จึงเป็นที่นิยมมากในการศึกษาการเขียนโปรแกรมในภาษาต่างๆ

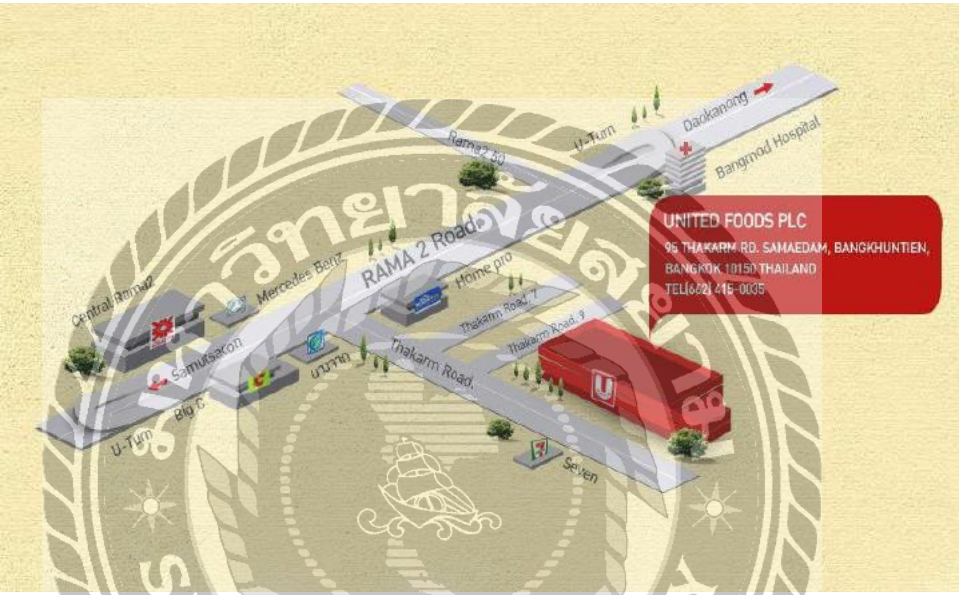
บทที่ 3
รายละเอียดการปฏิบัติงาน

3.1 ชื่อและที่ตั้งของสถานประกอบการ

บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน)

95 ถ. ท่าข้าม แขวงแสมดำ เขตบางขุนเทียน กรุงเทพมหานคร 10150

โทรศัพท์ 0-2849-6600 โทรสาร 0-2849-6666

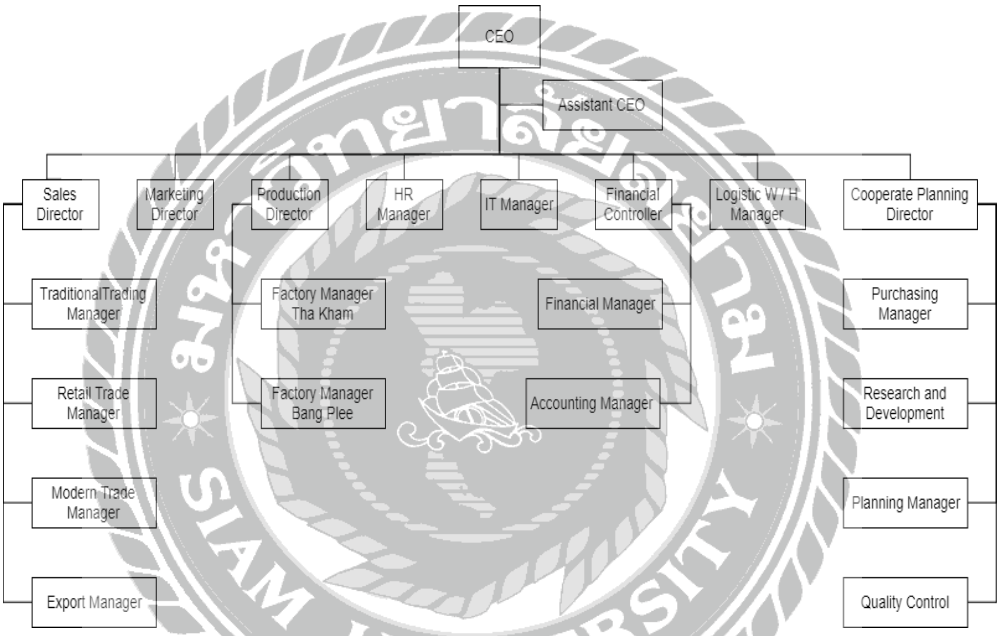


รูปที่ 3.1 แผนที่บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน)

3.2 ลักษณะการประกอบและการให้บริการหลัก

บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน) เป็นธุรกิจขนมขนาดใหญ่ ที่เป็นผู้ผลิต และจัดจำหน่าย แก่กลุ่มลูกค้าที่มีความสนใจ เช่น ห้างสรรพสินค้า ร้านขายของชำ เป็นต้น โดยมีศูนย์กระจายสินค้าอยู่ในจังหวัดต่างๆในประเทศ และบริษัทได้มีการติดต่อกับลูกค้าในต่างประเทศหลายๆ ประเทศ

3.3 รูปแบบการจัดองค์การและการบริหารงานขององค์กร



รูปที่ 3.2 โครงสร้างการจัดองค์การและการบริหารงานขององค์กร

งานแผนก IT

หน่วยงานแผนก ITเป็นหน่วยงานที่ให้บริการด้านสารสนเทศแก่งานของบริษัทยูไนเต็ดฟูดส์ ทั้งงานด้านข้อมูลลูกค้า ข้อมูลการขาย การบำรุงรักษาซ่อมแซมอุปกรณ์คอมพิวเตอร์และงานด้านระบบเครือข่ายหลักทั้งหมดรวมถึงการดูแล และพัฒนาเว็บไซต์การใช้งานด้านต่างๆ จึงเป็นหน่วยงานหลักที่มีประสิทธิภาพ และพร้อมให้บริการ

3.4 ตำแหน่งและลักษณะงานที่ได้รับมอบหมาย

นายปรินทร์ กิตติโสภณธรรม ตำแหน่ง หน่วยงาน IT

งานที่ได้รับมอบหมาย คือ จัดการข้อมูลของหน่วยงานและนำเสนอข้อมูลให้กับผู้จัดการฝ่าย IT การทำโมเดลวิเคราะห์ข้อมูล การไปตรวจคลังสินค้า การเข้าประชุมบริษัท เป็นต้น

3.5 ชื่อและตำแหน่งงานของพนักงานที่ปรึกษา

คุณปรีชา ตั้งเกียรติกำจาย ผู้จัดการฝ่าย IT

นายชิติวุฒ สีม่วง ตำแหน่ง พนักงาน IT

3.6 ระยะเวลาที่ปฏิบัติงาน

22 พฤษภาคม 2566 – กันยายน 2566

3.7 ขั้นตอนและวิธีการดำเนินงาน

ตารางที่ 3.1 แสดงระยะเวลาและขั้นตอนในการดำเนินงาน

ขั้นตอนการดำเนินงาน	พ.ค. 66	มิ.ย. 66	ก.ค. 66	ส.ค. 66
1.รวบรวมความต้องการและศึกษาความต้องการของโครงงาน	←	→		
2.จัดทำระบบ ปรับปรุงแก้ไขข้อบกพร่อง		←	→	
3.นำระบบไปทดลองใช้งานจริง			←	→
4.สรุปผลความพึงพอใจ			←	→
5.ปรับแก้ไขตามความต้องการ			←	→
6.จัดทำเอกสาร	←	—	—	→

3.8 อุปกรณ์และเครื่องมือที่ใช้

3.8.1 คุณสมบัติของฮาร์ดแวร์ผู้พัฒนาระบบ

1.ฮาร์ดแวร์สำหรับผู้พัฒนาระบบ

2.CPU CORE i5-1135G7 2.40 GHz

3.Hard Disk 1 TB.

4.DDR4 16.0 GB.

5.Mouse

6.Keyboard

3.8.2 ฮาร์ดแวร์สำหรับผู้ใช้ระบบ

1.CPU CORE i5-1135G7 2.40 GHz

2.Hard Disk 512 GB.

3.DDR4 16.0 GB.

4.Mouse

5.Keyboard

3.9 คุณสมบัติของซอฟต์แวร์

3.9.1 ซอฟต์แวร์สำหรับผู้พัฒนาระบบ

1.Microsoft Windows 11

2.Microsoft Excel 2019

3.Microsoft SQL Server 2019

4.Google Chrome Version 116.0.5845.111 (64-bit)

3.9.2 ซอฟต์แวร์สำหรับผู้ใช้ระบบ

1.Microsoft Windows 11

2.Microsoft Excel 2019

3.Google Chrome Version 116.0.5845.111 (64-bit)

บทที่ 4
ผลการปฏิบัติงานตามโครงงาน

4.1 รายละเอียดของโครงงาน

ภาพรวมเป็นการวิเคราะห์ข้อมูลธุรกิจเป็นการนำข้อมูลของลูกค้ามาวิเคราะห์เพื่อนำไปสู่การแนะนำสินค้าและสนับสนุนการตัดสินใจต่อยอดในแผนกลยุทธ์ทางธุรกิจต่อไปตามความต้องการของลูกค้าโดยการพัฒนาโครงงานนี้เป็นการนำเสนอให้กับทีมงานในบริษัทเพื่อที่จะนำไปต่อยอดในเชิงธุรกิจต่อไปได้โปรแกรมหลักที่ใช้ในการทำงานคือ Google Colab ในการทำ Data Science แล้วนำการแนะนำสินค้ามานำเสนอโดยนำข้อมูลมาจาก Database ของบริษัท มาทำการวิเคราะห์ ใช้เทคนิคการกรองแบบอิงเนื้อหา (Content-Based Filtering) เพื่อทำการแนะนำสินค้าให้ลูกค้าใหม่และใช้เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) เพื่อทำการแนะนำสินค้าให้ลูกค้าประจำในบางอย่างที่เขาไม่เคยซื้อ

4.2 ขั้นตอนการวิเคราะห์ข้อมูล

เนื่องจากงานที่ได้รับมอบหมายนั้นมีการใช้ข้อมูลลูกค้าจริงไม่สามารถนำข้อมูลออกมาเผยแพร่ได้จึงสามารถแสดงตัวอย่างงานได้บางส่วนเท่านั้นโดยงานที่ได้รับมอบหมายมีลักษณะขั้นตอนทำงานดังนี้

4.2.1 ศึกษาเครื่องมือการทำงาน

ผู้จัดทำได้ศึกษาภาษา Python และวิธีใช้ การทำงานต่าง ๆ ของ Google Colab

4.2.2 กำหนดเป้าหมายในการวิเคราะห์ข้อมูล (Research Understanding)

เพื่อเพิ่มยอดขายและแนะนำสินค้าให้กับลูกค้าใหม่และลูกค้าประจำโดยดูจากความสัมพันธ์ของข้อมูล

4.2.3 รวบรวมข้อมูล (Data Understanding)

ผู้จัดทำได้ศึกษาชุดข้อมูลที่ได้รับมอบหมายจากบริษัทข้อมูลชุดนี้เป็นข้อมูลการซื้อขายสินค้าโดยข้อมูลชุดนี้นำเข้าจาก Database และเปลี่ยนแปลงเป็นไฟล์ CSV เพื่อให้นำข้อมูลเข้าโมเดลใน Google Colab ในขั้นตอนถัดไป

```
[ ] datarecom = pd.read_csv('/data_recom.csv') #นำเข้าไฟล์ csv
    datarecom.head() #แสดงข้อมูล
```

	CustomerID	ID	TotalOrderQty
0	103110301682	1	1
1	206101102034	1	2
2	1031101120477	1	1
3	206104602037	1	4
4	206104201277	1	3

```
itemrecom = pd.read_csv('/data_item.csv') #นำเข้าไฟล์ csv
itemrecom.head()
```

	ProductID	ID	ProductRefCode	ProductName
0	30000048	1	30000048	ลูกอมรสบ๊วย 100 เม็ด 1x24x100
1	30000815	2	30000815	ลูกอม 123 ซ่าส์ รวมรส 1x24x100
2	30001003	3	30002926	โยโย โคลา 20g. 1x12x12
3	30001021	4	30001021	แก๊ป ชอกโกแลต 40g 1x6x12
4	30001080	5	30001080	แก๊ป สตรอเบอรี่ 40g 1x6x12

รูปที่ 4.1 ข้อมูลลูกค้าและรายการสินค้าการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

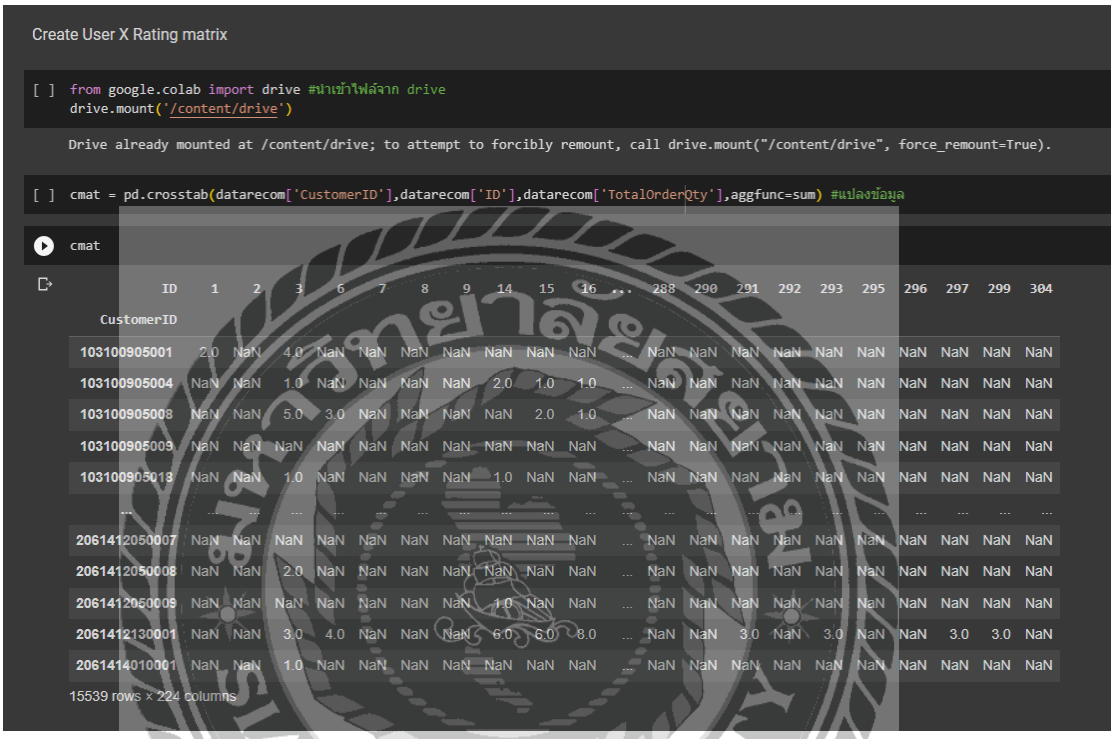
```
[ ] df.head()
```

	CustomerID	ID	TotalOrderQty
0	01111111111	1	1
1	01111111111	2	1
2	01111111111	3	2
3	01111111111	6	3
4	01111111111	7	4

รูปที่ 4.2 ข้อมูลลูกค้าการกรองแบบอิงเนื้อหา

4.2.4 เทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม (Collaborative Filtering)

สำหรับเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม ใช้ไลบรารี Pandas และ Numpy มาช่วยในการสร้างแบบจำลองการแนะนำสินค้า ซึ่งใช้อัลกอริทึม NMF โดยพิจารณาจากค่าความคล้ายคลึงแบบโคไซน์



Create User X Rating matrix

```
[ ] from google.colab import drive #นำเข้าไฟล์จาก drive
    drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

```
[ ] cmat = pd.crosstab(datarecom['CustomerID'],datarecom['ID'],datarecom['TotalOrderQty'],aggfunc=sum) #แปลงข้อมูล
```

```
cmat
```

ID	1	2	3	6	7	8	9	14	15	16	...	288	290	291	292	293	295	296	297	299	304
CustomerID																					
103100905001	2.0	NaN	4.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
103100905004	NaN	NaN	1.0	NaN	NaN	NaN	NaN	2.0	1.0	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
103100905008	NaN	NaN	5.0	3.0	NaN	NaN	NaN	NaN	2.0	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
103100905009	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
103100905018	NaN	NaN	1.0	NaN	NaN	NaN	NaN	1.0	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
20614120500007	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20614120500008	NaN	NaN	2.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20614120500009	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.0	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20614121300001	NaN	NaN	3.0	4.0	NaN	NaN	NaN	6.0	6.0	8.0	...	NaN	NaN	3.0	NaN	3.0	NaN	NaN	3.0	3.0	NaN
20614140100001	NaN	NaN	1.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

15539 rows × 224 columns

รูปที่ 4.3 การสร้างเมทริกซ์และจัดรูปแบบข้อมูล

จากรูปที่ 4.3 ทำการจัดรูปแบบข้อมูลให้อยู่ในการเข้าอัลกอริทึม NMF

```
[ ]  cmat = cmat.fillna(0) #เปลี่ยน NaN เป็น 0
     cmat
```

CustomerID	ID	1	2	3	6	7	8	9	14	15	16	...	288	290	291	292	293	295	296	297	299	304
103100905001		2.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905004		0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	1.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905008		0.0	0.0	5.0	3.0	0.0	0.0	0.0	0.0	2.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905018		0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...		...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2061412050007		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050008		0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412130001		0.0	0.0	3.0	4.0	0.0	0.0	0.0	6.0	6.0	8.0	...	0.0	0.0	3.0	0.0	3.0	0.0	0.0	3.0	3.0	0.0
2061414010001		0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

15539 rows x 224 columns

รูปที่ 4.4 แปลงข้อมูลที่ NaN ให้เป็นค่า 0.0

จากรูปที่ 4.4 ทำการแปลงข้อมูล NaN ให้เป็นค่า 0.0 เพื่อที่จะนำไปคูณกันกับข้อมูลในเมทริกซ์

```
[ ]  from sklearn.decomposition import NMF #นำเข้า libraries NMF
     nmf = NMF(50) #แบ่ง 50 กลุ่มที่คล้ายคลึงกัน
     nmf.fit(cmat) #คำนวณจาก cmat ตัวแปรที่ประกาศ
```

```
NMF(n_components=50)
```

```
[ ]  cmat
```

CustomerID	ID	1	2	3	6	7	8	9	14	15	16	...	288	290	291	292	293	295	296	297	299	304
103100905001		2.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905004		0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	1.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905008		0.0	0.0	5.0	3.0	0.0	0.0	0.0	0.0	2.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905018		0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...		...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2061412050007		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050008		0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412130001		0.0	0.0	3.0	4.0	0.0	0.0	0.0	6.0	6.0	8.0	...	0.0	0.0	3.0	0.0	3.0	0.0	0.0	3.0	3.0	0.0
2061414010001		0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

15539 rows x 224 columns

รูปที่ 4.5 นำข้อมูลเข้าอัลกอริทึม NMF

จากรูปที่ 4.5 แบ่งกลุ่มลูกค้าออกมา 50 กลุ่มที่มีรูปแบบการซื้อสินค้าคล้ายคลึงกัน

```
reconstructed = pd.DataFrame(np.round(np.dot(W,H),2), columns=cmat.columns) #คำนวณหาค่าเมทริก W และ H นำมาคูณกัน
reconstructed.index = cmat.index #เติมค่าลงใน  cmat
reconstructed
```

ID	1	2	3	6	7	8	9	14	15	16	...	288	290	291	292	293	295	296	297	299	304
CustomerID																					
103100905001	0.61	0.40	3.11	0.56	0.43	0.00	0.00	0.09	0.08	0.22	...	0.02	0.00	0.01	0.00	0.04	0.01	0.00	0.01	0.09	0.02
103100905004	0.01	0.09	0.24	0.00	0.03	0.00	0.00	3.08	1.25	0.56	...	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.02	0.01	0.01
103100905008	1.21	0.65	2.90	0.70	0.63	0.00	0.00	0.02	2.00	1.12	...	0.08	0.00	1.83	0.00	1.34	0.03	0.00	0.81	1.00	0.16
103100905009	0.00	0.04	0.11	0.03	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00	0.00	0.04	0.00	0.01	0.00	0.00	0.01	0.06	0.01
103100905018	0.30	0.20	0.12	0.01	0.21	0.00	0.00	0.00	0.04	0.00	...	0.00	0.00	0.03	0.00	0.01	0.00	0.00	0.00	0.03	0.02
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2061412050007	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2061412050008	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2061412050009	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2061412130001	0.64	0.48	3.34	1.78	0.27	0.54	0.34	6.67	7.84	3.86	...	0.39	0.01	10.77	0.01	8.03	0.12	0.24	4.42	4.81	0.79
2061414010001	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

15539 rows × 224 columns

รูปที่ 4.8 ผลการคำนวณเมทริกซ์ตัวที่ 1 กับ เทริกซ์ตัวที่ 2

```
import re

def recomendation(uid,topk=5): #หา Top 5 index
  res = list(reconstructed.T[uid].sort_values(ascending=False)[0:topk].index)
  res = itemrecom[itemrecom['ID'].isin(res)]
  mList = datarecom[datarecom['CustomerID']==uid][['ID','TotalOrderQty']] #หาข้อมูลการซื้อสินค้าของลูกค้าตามสินค้าที่หาได้
  # #return mList, res
  res = res.merge(mList,on='ID',how='left') #จะมีสินค้าที่เขาเคยซื้อกับไม่เคยซื้อ
  # #return res
  res = res[['ID','ProductID','ProductName','TotalOrderQty']]
  return res
```

```
res = recomendation(2061414010001,topk=10)
res
```

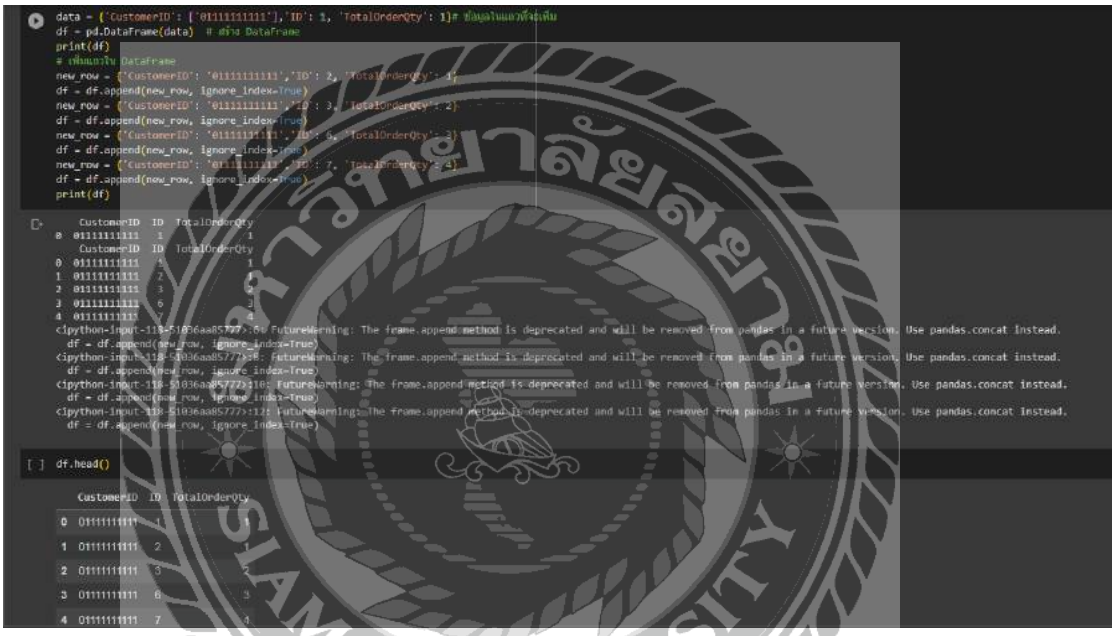
	ID	ProductID	ProductName	TotalOrderQty
0	24	30001781	โยโย องุ่น 20g 1x12x12 Pro	1.0
1	25	30001782	โยโย รวมรส 20g 1x12x12 Pro	NaN
2	26	30001825	เชียงไฮ เคลือบช็อกโกแลต 6g 1x36x15	1.0
3	27	30001827	เชียงไฮ เคลือบวานิลา 6g 1x36x15	1.0
4	28	30001829	เชียงไฮ เคลือบสตรอเบอร์รี 6g 1x36x15	1.0
5	30	30001831	เชียงไฮ เคลือบนม 6g 1x36x15	1.0
6	45	30001975	ไทม ช็อกโกแลต 30g 1x12x12	1.0
7	47	30001978	ไทม นม 30 g. 1x12x12	NaN
8	48	30001979	ไทม สตรอเบอร์รี 30g 1x12x12	NaN
9	50	30002455	เชียงไฮ ซุปเปอร์จัมโบ้ รสช็อกโกแลต 34 g. 1x12x12	1.0

รูปที่ 4.9 ผลลัพธ์ที่ได้จากเทคนิคการกรองข้อมูลแบบพึ่งพาผู้ใช้ร่วม

จากรูปที่ 4.9 จะทำการเลือกรหัสลูกค้า 2061414010001 มาทดสอบโดยเลือกสินค้าที่จะทำการแนะนำ 10 รายการ โดยจะแนะนำสินค้าที่ลูกค้าเคยซื้อและไม่เคยซื้อ

4.2.5 เทคนิคการกรองแบบอิงเนื้อหา (Content-Based Filtering)

สำหรับเทคนิคการกรองข้อมูลแบบองเนื้อหา ใช้ไลบรารี Pandas และ Numpy มาช่วยในการสร้างแบบจำลองการแนะนำสินค้าให้กับลูกค้าใหม่ซึ่งใช้อัลกอริทึม Cosine Math และทำการจำลองข้อมูลลูกค้าขึ้นมาใหม่และนำไปเทียบกับประวัติการซื้อของลูกค้าคนอื่นๆ โดยพิจารณาจากค่าความคล้ายคลึงแบบโคไซน์



รูปที่ 4.10 เพิ่มข้อมูลลูกค้าที่จะนำมาทดสอบ

จากรูปที่ 4.10 ทำการเพิ่มข้อมูลลูกค้าใหม่โดยสมมุติข้อมูลการซื้อสินค้า 5 ชนิด

```
[ ]  cmat1 = pd.crosstab(df['CustomerID'],df['ID'],df['TotalOrderQty'],aggfunc=sum)

[ ]  cmat1
```

CustomerID	ID	1	2	3	6	7
01111111111		1	1	2	3	4

รูปที่ 4.11 จัดการสร้างเมทริกซ์และจัดรูปแบบข้อมูล

จากรูปที่ 4.11 ทำการจัดรูปแบบข้อมูลให้อยู่ในการเข้าอัลกอริทึม Cosine Math

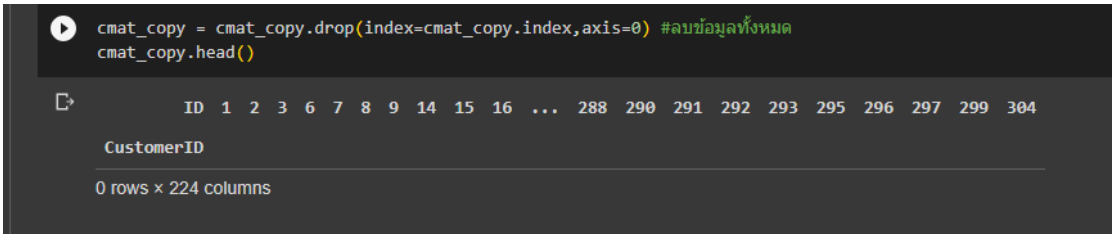
```
cmat_copy = cmat.copy()
cmat_copy
```

CustomerID	ID	1	2	3	6	7	8	9	14	15	16	...	288	290	291	292	293	295	296	297	299	304
103100905001		2.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905004		0.0	0.0	1.0	0.0	0.0	0.0	0.0	2.0	1.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905008		0.0	0.0	5.0	3.0	0.0	0.0	0.0	0.0	2.0	1.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
103100905018		0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...		...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2061412050007		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050008		0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412050009		0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2061412130001		0.0	0.0	3.0	4.0	0.0	0.0	0.0	6.0	6.0	8.0	...	0.0	0.0	3.0	0.0	3.0	0.0	0.0	3.0	3.0	0.0
2061414010001		0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

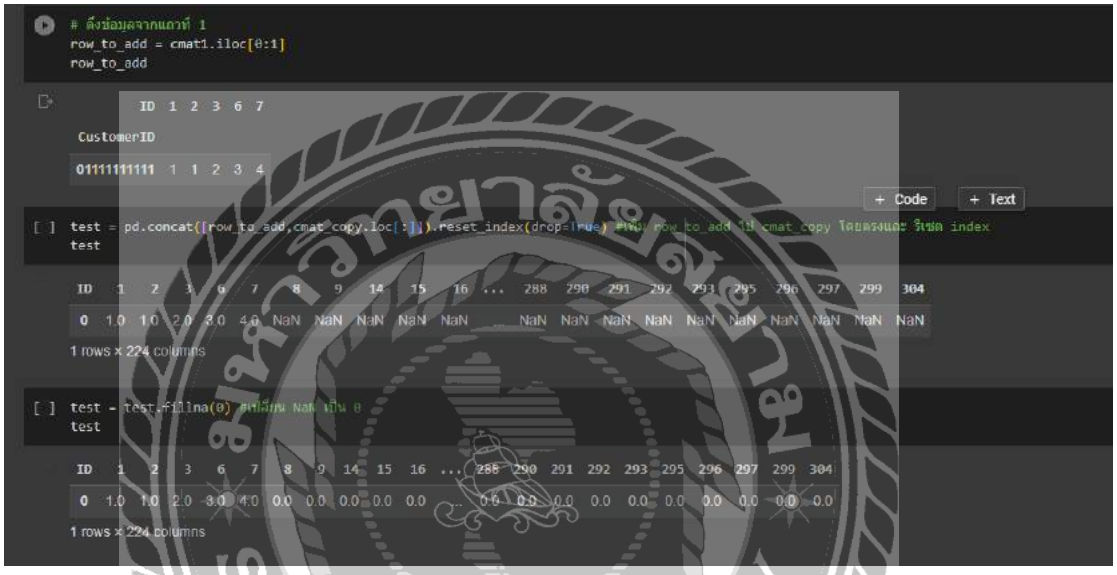
15539 rows × 224 columns

รูปที่ 4.12 ทำการคัดลอกตารางเดิมมาใช้งาน

จากรูปที่ 4.12 ทำการคัดลอกข้อมูลจากตารางเดิมเพื่อนำคอลัมน์ให้มีจำนวนเท่ากันเพื่อนำมาคูณเมทริกซ์ได้

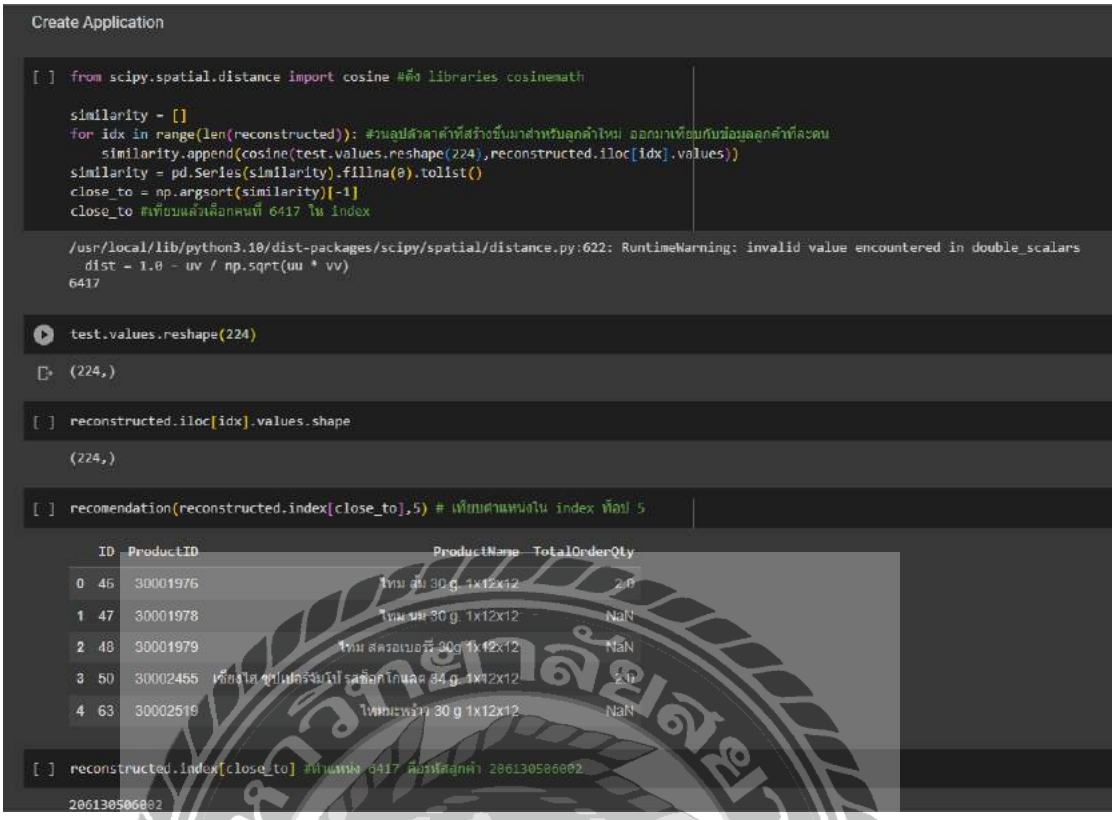


รูปที่ 4.13 ทำการลบข้อมูลข้อมูลในแถว



รูปที่ 4.14 ทำการดึงข้อมูลลูกค้าที่สมมุติมาใส่ในตาราง

จากรูปที่ 4.14 ทำการดึงข้อมูลลูกค้าที่สมมุติมาใส่ในตาราง ทำการจัดรูปแบบและแปลงค่า NaN เป็น 0.0



รูปที่ 4.15 ผลลัพธ์ที่ได้จากเทคนิคการกรองข้อมูลแบบอิงเนื้อหา

จากรูปที่ 4.15 จะทำการเลือกรหัสลูกค้า 01111111111 มาทดสอบโดยเลือกสินค้าที่จะทำการแนะนำ 5 รายการ โดยจะแนะนำสินค้าที่ลูกค้าเคยซื้อและ ไม่เคยซื้อซึ่งจะทำการเทียบโคไซน์ว่ามีการซื้อคล้ายคลึงกับลูกค้าคนไหนตามรูปคือ ตำแหน่งที่ 6417 คือรหัสลูกค้า 206130506002 จึงทำการแนะนำสินค้าที่รหัสลูกค้า 206130506002 เคยซื้อให้กับรหัสลูกค้า 01111111111

บทที่ 5
สรุปผลและข้อเสนอแนะ

5.1 สรุปผลโครงงาน

5.1.1 สรุปผลโครงงาน

คณะนักศึกษาสหกิจศึกษาได้พัฒนาระบบแนะนำสินค้าของ บริษัท ยูไนเต็ดฟูดส์ จำกัด (มหาชน) เพื่อนำมาใช้งานในการแนะนำสินค้าให้กับลูกค้าใหม่และลูกค้าประจำ เพื่อเพิ่มยอดขายและความรวดเร็วให้กับพนักงานขาย เช่น พนักงานขายต้องการแนะนำสินค้าให้ลูกค้าใหม่โดยดูจากรายการที่เขาลองสั่งซื้อและแนะนำสินค้าที่เขาไม่เคยซื้อและลูกค้าประจำที่มีประวัติการซื้อสินค้านำมาเปรียบเทียบกับลูกค้าคนอื่นๆที่มีลักษณะการซื้อที่คล้ายกันนำมาแนะนำสินค้าที่เขายังไม่เคยซื้อ

5.1.2 ข้อจำกัดหรือปัญหาของโครงงาน

5.1.2.1 นักศึกษายังไม่มีความชำนาญในการเขียนโปรแกรมภาษา Python จึงต้องใช้เวลาศึกษาเพิ่มเติม

5.1.2.2 นักศึกษายังไม่เข้าใจในระบบงานที่จะทำจึงต้องใช้เวลาในการศึกษา

5.1.3 ข้อเสนอแนะ

5.1.3.1 ควรศึกษาด้านการพัฒนาโปรแกรมภาษา Python เพิ่มเติม

5.2 สรุปผลการปฏิบัติงานสหกิจศึกษา

5.2.1 ข้อดีของการปฏิบัติงานสหกิจศึกษา

5.2.1.1 ได้เรียนรู้ประสบการณ์ในการปฏิบัติงานจริงในการทำงานในสถานประกอบการใหญ่ๆ

5.2.1.2 พัฒนาศักยภาพด้านการพัฒนาโปรแกรมภาษา Python

5.2.1.3 ทำให้เข้าใจถึงกระบวนการทางธุรกิจ

5.2.2 ปัญหาที่พบของการปฏิบัติงานสหกิจศึกษา

5.2.2.1 วันที่เริ่มปฏิบัติงานแรกๆไม่มีความรู้ในการแก้ไขปัญหา

5.2.2.2 ขาดประสบการณ์ในการปฏิบัติงานในรูปแบบขององค์กร

5.2.2.3 ขาดประความรู้ความเข้าใจทางด้านกระบวนการทางธุรกิจ

5.2.3 ข้อเสนอแนะ

5.2.3.1 ควรศึกษาจากการปฏิบัติงานสหกิจศึกษา เพื่อเป็นประสบการณ์ในการทำงาน

5.2.3.2 ควรศึกษากระบวนการทางธุรกิจให้เพิ่มมากขึ้น

บรรณานุกรม

Anand Shanker Tewari. (2020). *Generating Items Recommendations by Fusing Content and User-Item based Collaborative Filtering*. https://www.sciencedirect.com/science/ article/pii/ S1877050920306803

Ankita Prasad. (2020). *The Mathematics of Recommendation Systems*. https://levelup .gitconnected.com /the-mathematics-of-recommendation-systems-e8922a50bdea

Bindhu Balu. (2019). *Content-Based Recommendation System*. https://medium.com/ @bindhubalu/ content-based-recommender-system-4db1b3de03e7

CHAKRIT. (2019). *Similarity*. https ://www.softnix.co.th/2019/05/29/similarity-ความเหมือนที่แตกต่าง

Esteban,Zafra,Romero. (2020). *Helping university students to choose elective courses by using a hybrid multi-criteria recommendation system with genetic optimization*. https:// www.sciencedirect. com/science/article/abs/pii/S0950705119306306

Google Colab. (2023). https://colab.research.google.com/?hl=th

Katherine Bailey. (2016). *Matrix Factorization with Tensorflow*. https://katbailey.github.io/ post/matrix-factorization-with-tensorflow/

Katie Gross. (2023). *Recommendation Engines*. https://blog.dataiku.com/recommendation-engines-how-they-work-in-plain-english

Lenhart Philip,Daniel Herzog. (2016). *Combining Content-based and Collaborative Filtering*. http://www.di.uniba.it/~lops/CBRecSys/CBRecSys2016_proceedings.pdf#page=8

Michael Lin. (2018). *Cosine Similarity & Classification*. https://www.linkedin.com/pulse/cosine - similarity-classification-michael-lin

My SQL Database. (2023). https://www.w3schools.com/sql/default.asp

Python. (2023). https://www.w3schools.com/python/default.asp

Roberto Iriondo Saniya Parveez. (2020). *Recommendation System Tutorial with Python using Collaborative Filtering*. https://pub.towardsai.net/recommendation-system-in- depth-tutorial-with-python-for-netflix-using-collaborative-filtering-533ff8a0e444

Sirinart Tangruamsub. (2019). *Recommendation System (แบบสรุปไม่ค่อยสั้น)*. https:// www.medium.com/@sinart.t/recommendation-system-แบบสรุปเอา

Surapong Kanoktipsatharporn. (2019). *Latent Semantic Analysis (LSA) คืออะไร Text Classification ด้วย Singular Value Decomposition (SVD), Non-negative Matrix Factorization (NMF) – NLP*. https://www.bualabs.com/archives/2971/lsa-latent-semantic-analysis-text-classification-singular-value-decomposition-svd-non-negative-matrix-factorization-nmf-nlp-ep-4/

Xiangjie Kong. (2019). *A hybrid paper recommendation system*. https://www. researchgate.net /figure/A-hybrid-paper-recommendation-system_fig5_330077673

Zhiyuan Fang. (2016). *Hybrid Recommender System Based on Personal Behavior Mining*. https://www.researchgate.net/publication/305186147_Hybrid_Recommender_System_Based_on_Personal_Behavior_Mining

ภาคผนวก

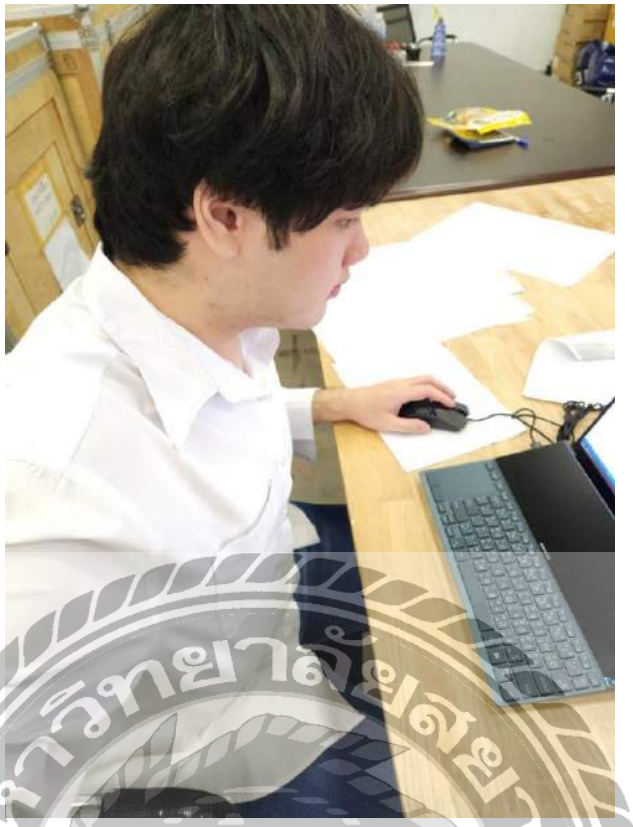
ภาคผนวก ก

ภาพระหว่างปฏิบัติงาน

รูปที่ ก.1 ออกแบบ UX UI หน้าเว็บไซต์

รูปที่ ก.2 เขียนโค้ดระบบแนะนำสินค้า

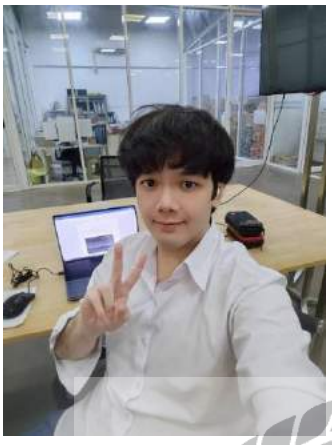
รูปที่ ก.3 ลองใช้เครื่องมือต่างๆในการวิเคราะห์ข้อมูล



รูปที่ ก.4 เลือกข้อมูลจาก Database

รูปที่ ก.5 คัดเลือกข้อมูลลง Excel

ประวัตินักศึกษาสหกิจศึกษา



รหัสนักศึกษา	:	6305000007
ชื่อ-นามสกุล	:	นายปรินทร์ กิตติโสภณธรรม
ที่อยู่	:	หมู่บ้านวินนิ่ง 98/8 ซ.ศาลธนบุรี44/2 ถ.กำนันแม้น แขวงบางหว้า เขตภาษีเจริญ กรุงเทพฯ 10160
เบอร์โทรศัพท์	:	090-992-1096
E-Mail	:	printae_ss@hotmail.com
ระดับปริญญาตรี	:	มหาวิทยาลัยสยาม

https://drive.google.com/drive/folders/1FnGHp9BcsuYCg-VGH6pG0ykV6AiSbV9b?usp=sharing

รายงานการปฏิบัติงานสหกิจศึกษา

ระบบแนะนำสินค้า

PRODUCT RECOMMENDATION SYSTEM

โดย

นายปรินทร์ กิตติโสภณธรรม 6305000007

รายงานนี้เป็นส่วนหนึ่งของวิชา สหกิจศึกษา

ภาควิชาธุรกิจดิจิทัล

คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

ภาคการศึกษาที่ 3 ปีการศึกษา 2565