



Sales Promotion System for Bookstores based on Face Recognition and
Book Recommendation

ระบบส่งเสริมการขายสำหรับร้านหนังสือโดยใช้การจดจำใบหน้าและการแนะนำหนังสือ

สุนิตา ขจรวิทยา
Sunita Kachornvitaya

กาญจนา ดอนไพรที
Kanjana Donpraitee

โครงการนี้เป็นส่วนหนึ่งของการศึกษาหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ปีการศึกษา 2561



คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

โครงงาน

ภาษาไทย : ระบบส่งเสริมการขายสำหรับร้านหนังสือ โดยใช้การจดจำใบหน้าและ
การแนะนำหนังสือ

ภาษาอังกฤษ : Sales Promotion System for Bookstores based on Face
Recognition and Book Recommendation

นิสิต

นางสาวสุนิตา ขจรวิทยา 58102010214

นางสาวกาญจนา ดอนไพรठी 58102010801

ปริญญา วิทยาศาสตร์บัณฑิต (วท.บ.)

ภาควิชา วิทยาการคอมพิวเตอร์

อาจารย์ที่ปรึกษาโครงงาน ผู้ช่วยศาสตราจารย์ รุ่งระพี กรานคำยี่

ลงชื่อ.....

(ผู้ช่วยศาสตราจารย์ รุ่งระพี กรานคำยี่)

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ โดยใช้เทคโนโลยีจดจำใบหน้า มาใช้ร่วมกับระบบการแนะนำหนังสือ ซึ่งการดำเนินงานวิจัยนี้ประกอบไปด้วยการพัฒนาทั้งหมด 2 โมเดล และการพัฒนาเว็บแอปพลิเคชันแนะนำหนังสือ ซึ่งโมเดลแรกคือ โมเดลสำหรับระบบจดจำใบหน้า จะใช้วิธี Deep Convolutional Neural Network มีขั้นตอนการทำงาน คือ Face detection, Face Landmark Estimation, Encoding Faces และ Classify model สำหรับการแยกแยะลักษณะใบหน้าของแต่ละบุคคล จะใช้อัลกอริทึม K-Nearest Neighbor, Gaussian Naive Bayes, Random Forest และ DecisionTree ซึ่งค่าความแม่นยำในการแบ่งกลุ่มของวิธี K-Nearest Neighbor มีความแม่นยำสูงสุดคือ 100% ในส่วนของโมเดลที่ 2 คือ โมเดลสำหรับระบบแนะนำหนังสือ จะใช้วิธีการกรองข้อมูลแบบผสม Hybrid Filtering โดยการรวม Content-based และ Collaborative Filtering เข้าด้วยกัน ซึ่งใช้ข้อมูลจาก Goodreads (เว็บไซต์สำหรับรีวิวหนังสือทั่วโลก) ในการทำนายการแนะนำหนังสือจากวิธี Collaborative Filtering จะใช้อัลกอริทึม Singular Value Decomposition (SVD), SVD++, BaselineOnly, CoClustering, K Nearest Neighbors (KNN) Basic, KNNWithMeans, KNNWithZScore, Non-negative Matrix Factorization, SlopeOne และ NormalPredictor ซึ่ง SVD เป็นอัลกอริทึมที่เหมาะสม และที่มีค่า Root Mean Square Error (RMSE) อยู่ที่ 0.925 ในส่วนของแอปพลิเคชันแนะนำหนังสือสามารถตรวจจับและระบุตัวบุคคลจากใบหน้าของลูกค้าได้ และสามารถแนะนำหนังสือได้ตามความสนใจของลูกค้า

Abstract

This research aims to develop the sales promotion application for bookstores using face recognition and book recommendation technology by developing 2 models and a book recommendation application. The first model is for face recognition. It uses Deep Convolutional Neural Network which consists of Face Detection, Face Landmark Estimation, Encoding Faces, and Classify model for extracts each face. Multiple algorithm including K-Nearest Neighbor (KNN), Gaussian Naive Bayes, Random Forest, and DecisionTree. The KNN is the most appropriate algorithm to build the model that has 100% accuracy. The second model is for the book recommendation. It uses Hybrid filtering combining Content-based and Collaborative Filtering, which based on training data from Goodreads (Website for worldwide books review) to recommend books. Multiple algorithm including Singular Value Decomposition (SVD), SVD++, BaselineOnly, CoClustering, K Nearest Neighbors (KNN) Basic, KNNWithMeans, KNNWithZScore, Non-negative Matrix Factorization, SlopeOne, and NormalPredictor are tested on these models. The SVD is the most appropriate algorithm to build the model that RMSE and MAE give the least error results of 0.9117 and 0.7288 respectively. The book recommendation application can detect and identify the face of the customer and recommend books according to the customer's interest.

กิตติกรรมประกาศ

โครงการระบบส่งเสริมการขายสำหรับร้านหนังสือ โดยการใช้การจดจำใบหน้าและการแนะนำหนังสือ (Sales Promotion System for Bookstores based on Face Recognition and Book Recommendation) สามารถสำเร็จลุล่วงได้ด้วยความสำเร็จจาก ผู้ช่วยศาสตราจารย์ รุ่งระพี กรานคำ ยี อาจารย์ที่ปรึกษาโครงการ ที่กรุณาเสียสละเวลาเพื่อคอยช่วยเหลือ ให้คำปรึกษา คำแนะนำแนวทาง ในการแก้ปัญหาและแก้ไขข้อบกพร่องที่เกิดขึ้นระหว่างการทำงาน คณะผู้จัดทำกราบขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ทุกๆ ท่านในภาควิชาวิทยาการคอมพิวเตอร์ที่กรุณาให้ความรู้ และสามารถนำความรู้ที่ได้เรียนมาประยุกต์ใช้ในการดำเนินโครงการ รวมทั้งให้คำปรึกษาและคำแนะนำที่เป็นประโยชน์ ทำให้คณะผู้จัดทำสามารถแก้ไขโครงการได้สมบูรณ์มากยิ่งขึ้น

ขอขอบพระคุณมหาวิทยาลัยศรีนครินทรวิโรฒ, คณะวิทยาศาสตร์ และภาควิชาวิทยาการคอมพิวเตอร์ ที่เป็นแหล่งศึกษาหาความรู้ ให้การสนับสนุนสถานที่และอุปกรณ์ต่างๆ สำหรับการดำเนินโครงการ นอกจากนี้ยังรวมถึงประสบการณ์ในการทำกิจกรรมต่างๆ เพื่อนำไปประยุกต์ใช้ในการทำงาน และการดำเนินชีวิตต่อไป

ขอขอบพระคุณบิดา มารดา รวมถึงครอบครัวที่ได้ให้การช่วยเหลือและสนับสนุน ตลอดจนถึงความห่วงใย และยังเป็นกำลังใจสำคัญในการทำโครงการเสมอมา

ขอขอบคุณเพื่อนๆ ที่ช่วยสละเวลาให้ความร่วมมือในการเก็บรวบรวมข้อมูล ตลอดจนถึงให้คำแนะนำ คำปรึกษา และคอยช่วยเหลือ ทำให้โครงการมีความสมบูรณ์มากยิ่งขึ้น

สุดท้ายขอขอบคุณผู้มีส่วนเกี่ยวข้องทุกท่านที่ทำให้โครงการนี้สำเร็จลุล่วงไปได้ด้วยดี และหวังว่าโครงการระบบส่งเสริมการขายสำหรับร้านหนังสือ โดยการใช้การจดจำใบหน้าและการแนะนำหนังสือ (Sales Promotion System for Bookstores based on Face Recognition and Book Recommendation) จะเป็นประโยชน์ต่อผู้ที่สนใจศึกษา และสามารถเป็นส่วนหนึ่งสำหรับผู้ที่สนใจจะนำไปต่อยอดในบริบทต่างๆต่อไป

สุนิตา ขจรวิทยา

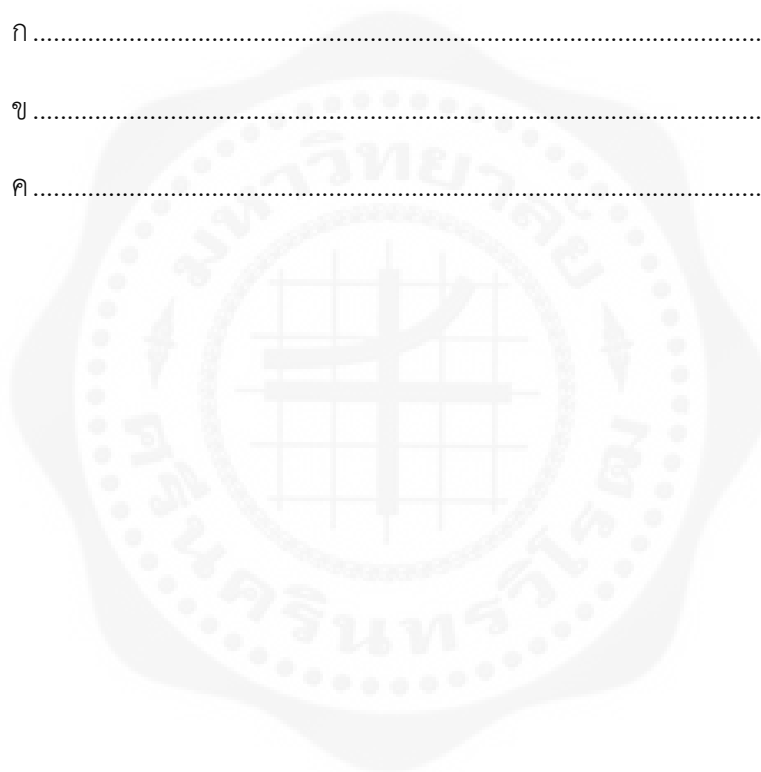
กาญจนา ดอนไพรที้

สารบัญ

บทคัดย่อ.....	ก
Abstract.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง	ช
สารบัญรูปภาพ	ซ
บทที่ 1 บทนำ.....	1
1.1. ที่มาและความสำคัญของโครงการ.....	1
1.2. วัตถุประสงค์ของโครงการ.....	1
1.3. ขอบเขตของโครงการ	2
1.4. ประโยชน์ที่คาดว่าจะได้รับ	2
บทที่ 2 องค์ความรู้ที่เกี่ยวข้อง	3
2.1 องค์ความรู้เกี่ยวกับ Face Recognition [1].....	3
2.2 องค์ความรู้เกี่ยวกับ Recommendation Systems [2]	3
2.3 องค์ความรู้เกี่ยวกับ Content Based Filtering [2].....	4
2.4 องค์ความรู้เกี่ยวกับ Collaborative Filtering [2].....	5
2.5 องค์ความรู้เกี่ยวกับ Hybrid Filtering [3]	6
2.6 งานวิจัยที่เกี่ยวข้อง	8
2.6.1 งานวิจัยเรื่อง“OpenFace: A general-purpose face recognition library with mobile applications” [5].....	8

2.6.2 งานวิจัยเรื่อง “A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System” [6]	9
2.6.2 งานวิจัยเรื่อง “Book Recommendation System Based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining” [7]....	9
บทที่ 3 วิธีการดำเนินการศึกษาค้นคว้า	10
3.1 แผนการดำเนินงาน.....	10
3.1.1 การเริ่มต้นและวางแผนโครงการ.....	11
3.1.2 การรู้จำใบหน้า (Face Recognition).....	11
3.1.3 การแนะนำหนังสือ (Book Recommendation).....	12
3.1.4 การพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ	13
3.2 แผนการดำเนินงานตลอดโครงการ	14
3.3 อุปกรณ์และเครื่องมือที่ใช้ในโครงการ	15
3.3.1 ฮาร์ดแวร์ (Hardware).....	15
3.3.2 ซอฟต์แวร์ (Software).....	15
3.3.3 ภาษาที่ใช้	15
3.4 ชุดข้อมูล.....	16
3.4.1 Goodbooks-10k.....	16
บทที่ 4 ผลการดำเนินงาน	17
4.1 ผลลัพธ์จากการเก็บข้อมูลในส่วนของการรู้จำใบหน้า (Face Recognition)	17
4.2 ผลลัพธ์จากการพัฒนาแบบจำลอง (Modeling) ในส่วนของการรู้จำใบหน้า (Face Recognition)	18
4.3 ผลลัพธ์จากการเก็บข้อมูลในส่วนของการแนะนำหนังสือ (Book Recommendation)	19
4.4 ผลลัพธ์จากการพัฒนาแบบจำลอง (Modeling) ในส่วนของการแนะนำหนังสือ (Book Recommendation).....	19

4.5 ผลลัพธ์จากการพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ.....	21
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะ	24
5.1 สรุปผลการดำเนินโครงการ	24
5.2 ปัญหาและอุปสรรคในการดำเนินโครงการ.....	24
5.3 ข้อเสนอแนะ	25
บรรณานุกรม	26
ภาคผนวก	27
ภาคผนวก ก	28
ภาคผนวก ข	31
ภาคผนวก ค	34



สารบัญตาราง

ตารางที่ 1 ผลการทดสอบความแม่นยำของการรู้จำใบหน้า (Face Recognition).....	18
--	----



สารบัญรูปภาพ

ภาพที่ 1	สถาปัตยกรรมพื้นฐานของระบบให้การแนะนำ	4
ภาพที่ 2	ภาพแสดงลักษณะการทำงานของ Content Based Filtering	5
ภาพที่ 3	ภาพแสดงลักษณะการทำงานของ Collaborative Filtering.....	6
ภาพที่ 4	ภาพแสดงลักษณะการทำงานของ Hybrid Recommendation	8
ภาพที่ 5	ภาพแสดงกระบวนการ Train ของ OpenFace	8
ภาพที่ 6	แผนการดำเนินงานโดยภาพรวม	10
ภาพที่ 7	ขั้นตอนการ Training Data By Face Recognition	11
ภาพที่ 8	วิธีการกรองข้อมูลแบบผสม Hybrid Filtering	12
ภาพที่ 9	ตัวอย่างเว็บแอปพลิเคชัน	13
ภาพที่ 10	ภาพแสดง Cloumn ของ books.csv	16
ภาพที่ 11	ภาพแสดง Cloumn ของ book_tags.csv.....	16
ภาพที่ 12	ภาพแสดง Cloumn ของ ratings.csv.....	16
ภาพที่ 13	ภาพแสดง Cloumn ของ tags.csv	16
ภาพที่ 14	ตัวอย่างของรูปถ่ายที่ไม่สามารถเก็บข้อมูลได้.....	17
ภาพที่ 15	ตัวอย่างของรูปถ่ายที่สามารถเก็บข้อมูลได้.....	17
ภาพที่ 16	ตัวอย่างข้อมูลของรูปถ่ายที่สามารถเก็บข้อมูลได้	18
ภาพที่ 17	ตัวอย่างของการเก็บข้อมูลของการแนะนำหนังสือ.....	19
ภาพที่ 18	ตัวอย่างของ Vector Representations โดยใช้ TfidfVectorizer ใน scikit-learn.....	19
ภาพที่ 19	ตัวอย่างของ Cosine Similarity โดยใช้ cosine_similarity ใน scikit-learn	20
ภาพที่ 20	กราฟแสดงผลทดลองกับอัลกอริทึมทั้ง 10 แบบของ Collaborative Filtering.....	20
ภาพที่ 21	ภาพตัวอย่างของการแนะนำหนังสือโดยใช้การกรองข้อมูลแบบผสม Hybrid Filtering	21
ภาพที่ 22	ภาพตัวอย่างการรู้จำใบหน้าผ่าน Video Streaming	21
ภาพที่ 23	ภาพตัวอย่างห็นด้านข้าง โดยเห็นดวงตาข้างเดียว	22
ภาพที่ 24	ภาพตัวอย่างห็นด้านข้าง โดยเห็นดวงตาทั้งสองข้าง	22
ภาพที่ 25	ภาพตัวอย่างการจับภาพใบหน้าในระยะ 1.5 เมตร	22
ภาพที่ 26	ภาพตัวอย่างการจับภาพใบหน้าในระยะ 2.4 เมตร	23
ภาพที่ 27	ภาพตัวอย่างเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ.....	23

บทที่ 1

บทนำ

1.1. ที่มาและความสำคัญของโครงการ

เนื่องจากมนุษย์แต่ละบุคคลมีพฤติกรรมด้านสังคมแตกต่างกัน นับตั้งแต่ลักษณะการพูดจา สื่อสาร การแต่งกาย การคบเพื่อน และบุคลิกภาพทางสังคมอื่นๆ ทั้งนี้เพราะแต่ละบุคคลมาจากสังคมที่แตกต่างกัน เช่น มาจากครอบครัวที่แตกต่างกัน ซึ่งหมายถึงได้รับการอบรมเลี้ยงดูที่แตกต่างกัน บิดามารดามีอาชีพการศึกษา ฐานะทางเศรษฐกิจและลักษณะอื่นๆ ที่แตกต่างกัน ย่อมส่งผลให้บุคคลมีลักษณะสังคมที่ไม่เหมือนกัน นอกจากครอบครัวแล้วยังมีหน่วยสังคมอื่นๆ ที่มีอิทธิพลทำให้บุคคลมีความแตกต่างกันทางด้านสังคม เช่น กลุ่มเพื่อน ผู้ร่วมงาน โรงเรียน ชุมชนที่บุคคลอาศัยอยู่ และบุคคลที่เกี่ยวข้อง ฯลฯ และความแตกต่างทางด้านสังคมดังกล่าวจะส่งผลให้บุคคลมีแนวโน้มที่จะมีลักษณะของความสนใจ ความต้องการ ค่านิยมและแรงจูงใจในการทำพฤติกรรมต่างๆ แตกต่างกันไปด้วย

ความแตกต่างของแต่ละบุคคลนี้สามารถนำไปประยุกต์เข้ากับการแนะนำสินค้า และเทคโนโลยีระบบจดจำใบหน้า ซึ่งการแนะนำสินค้านั้นมีจุดมุ่งหมายเพื่อให้ผู้เสนอสินค้าสามารถส่งออกสินค้าที่ตนนำเสนอได้ดียิ่งขึ้น และให้ผู้ซื้อได้สินค้าที่เหมาะสมกับตนเอง ในปัจจุบันทั้งร้านค้าและผู้ซื้อมีความเชื่อมต่อกันมากขึ้น ข้อมูลจำนวนมากทั้งของร้านค้า และผู้ซื้อสินค้าได้ถูกบันทึกไว้บนฐานข้อมูล ทำให้มีการสร้างระบบการแนะนำสินค้าที่แม่นยำและเฉพาะบุคคลมากยิ่งขึ้น และเทคโนโลยีระบบจดจำใบหน้า (Face Recognition) ซึ่งเป็นกระบวนการนำภาพไปตรวจจับและประมวลผล แล้วนำไปใช้ในระบบการรู้จำใบหน้าซึ่งเป็นส่วนหนึ่งของ เทคโนโลยีปัญญาประดิษฐ์ในส่วนของการรับรู้ของเครื่อง (Machine perception) เพื่อระบุว่าใบหน้านั้นตรงกับบุคคลใด แล้วจึงนำมาใช้งานในด้านต่างๆ เช่น ระบบการยืนยันตัวตน และ ระบบการติดตามบุคคล เป็นต้น

ผู้วิจัยจึงเล็งเห็นว่าการนำเทคโนโลยีจดจำใบหน้าที่จะทำการระบุตัวบุคคล มาใช้ร่วมกับระบบการแนะนำหนังสือที่จะทำการแนะนำหนังสือที่เหมาะสมกับแต่ละบุคคล โดยทำให้เกิดเป็นระบบส่งเสริมการขายสำหรับร้านหนังสือ ซึ่งจะเป็นตัวช่วยในการขายสินค้าให้ตรงกับความต้องการของแต่ละบุคคลมากขึ้น

1.2 วัตถุประสงค์ของโครงการ

ในการวิจัยครั้งนี้ผู้วิจัยได้ตั้งวัตถุประสงค์ในการการทำงานวิจัยไว้ดังนี้

- 1.2.1 เพื่อพัฒนาเว็บแอปพลิเคชันแนะนำหนังสือให้กับลูกค้าภายในร้านหนังสือ
- 1.2.2 เพื่อสามารถแนะนำรายการหนังสือจากประวัติความสนใจของลูกค้า
- 1.2.3 เพื่อเพิ่มคุณภาพด้านการขายจากการรู้จำใบหน้าและการแนะนำหนังสือ

1.3 ขอบเขตของโครงการ

- 1.3.1 พัฒนาเว็บแอปพลิเคชันแนะนำหนังสือให้กับลูกค้าภายในร้านหนังสือ
- 1.3.2 ตรวจจับใบหน้าและระบุข้อมูลลูกค้าที่เคยใช้งานได้ โดยใช้ Face recognition
- 1.3.3 แนะนำรายการหนังสือได้ โดยใช้ Book recommendation ด้วยวิธีการกรองข้อมูลแบบผสม (Hybrid Filtering)
- 1.3.4 ใช้ Face recognition ในการเข้าถึงข้อมูลลูกค้าที่เคยใช้งาน เพื่อใช้ในการ Recommendation

1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1.4.1 สามารถพัฒนาเว็บแอปพลิเคชันเพื่อการพัฒนาคุณภาพด้านการขาย จากการจดจำใบหน้าและการแนะนำหนังสือ
- 1.4.2 สามารถตรวจจับใบหน้าและข้อมูลลูกค้าที่เคยใช้งานได้ จากการใช้ Face recognition
- 1.4.3 สามารถแนะนำรายการสินค้าได้ โดยใช้ Book recommendation ด้วยวิธีการกรองข้อมูลแบบผสม (Hybrid Filtering)

บทที่ 2

องค์ความรู้ที่เกี่ยวข้อง

2.1 องค์ความรู้เกี่ยวกับ Face Recognition [1]

ระบบการรู้จำใบหน้าหรือ ระบบการจดจำใบหน้า (facial recognition system) คือระบบการตรวจหาใบหน้าของมนุษย์และปรับภาพใบหน้าโดยอัตโนมัติ กรอบจะปรากฏขึ้นบนใบหน้าที่ถูกตรวจจับและโฟกัส สี และค่าการวัดแสงจะถูกปรับโดยอัตโนมัติ นอกจากนั้นเมื่อบันทึกด้วยคุณภาพแบบ HD เทคโนโลยีการบีบอัดจะจัดสรรความจุของข้อมูลให้ลดลง แต่ได้ข้อมูลที่เป็นประโยชน์มากขึ้นเพื่อปรับคุณภาพของภาพ ข้อมูลที่ได้จะถูกนำไปเปรียบเทียบกับข้อมูลตัวอย่างที่เก็บบันทึกไว้ อาจจะทั้งใบหน้าหรือเพียงบางส่วน ขึ้นกับชนิดของวิธีแยกเอกลักษณ์ใบหน้า ระบบการรู้จำใบหน้าเป็นส่วนหนึ่งของเทคโนโลยีปัญญาประดิษฐ์ในส่วนเนื้อหาของเรื่อง การรับรู้ของเครื่อง (Machine perception)

2.2 องค์ความรู้เกี่ยวกับ Recommendation Systems [2]

ระบบให้การแนะนำ เป็นเทคโนโลยีสมัยใหม่ที่ถูกนำมาใช้ในการแนะนำข้อมูลต่างๆ ที่คาดว่าผู้ใช้น่าจะสนใจ หรืออาจจะเป็นข้อมูลที่ผู้ใช้ต้องการ และเหมาะสำหรับการดำเนินธุรกิจแบบพาณิชย์อิเล็กทรอนิกส์ (e-Commerce) โดยทั่วไประบบให้การแนะนำประกอบด้วย 3 ส่วน ดังต่อไปนี้ คือ

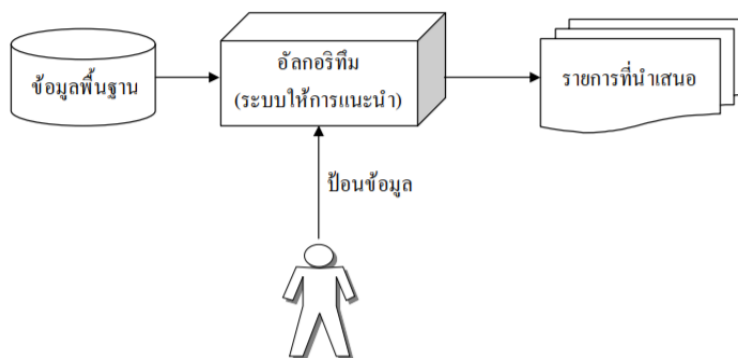
2.2.1 ส่วนข้อมูลพื้นฐานที่จำเป็นต้องใช้ในการประมวลผล เช่น โปรไฟล์ของผู้ใช้แต่ละคน

2.2.2 ส่วนการป้อนข้อมูล เป็นข้อมูลที่ได้จากการป้อนข้อมูลเข้ามาของผู้ใช้ เช่น การให้คะแนน Rating ซึ่งมีอยู่ 2 แบบ ดังนี้

2.2.2.1 แบบชัดเจน (Explicit) Rating แบบชัดเจนจะแสดงอยู่ในรูปของจำนวนตัวเลขตามระดับความนิยมตั้งแต่ 1 ถึง 5, 1 ถึง 10 หรือระดับอื่นๆ ขึ้นอยู่กับการใช้งาน

2.2.2.2 แบบไม่ชัดเจน (Implicit) ส่วน Rating แบบไม่ชัดเจนได้มาจากพฤติกรรมการใช้งานของผู้ใช้ต่างๆ เช่น ประวัติการซื้อสินค้าหรือประวัติการเข้ามาใช้งานของผู้ใช้ในอดีตที่ผ่านมา

2.2.3 ส่วนอัลกอริทึมเป็นส่วนสำคัญที่สุดที่ใช้ประมวลข้อมูลเพื่อให้การแนะนำขึ้นข้อมูลออกมาดังแสดงไว้ในภาพที่ 1



ภาพที่ 1 สถาปัตยกรรมพื้นฐานของระบบให้คำแนะนำ

2.3 องค์ความรู้เกี่ยวกับ Content Based Filtering [2]

Content-based เป็นเทคนิคหนึ่งในการนำเสนอระบบการให้คำแนะนำ ที่ใช้พื้นฐานจากข้อมูลจากรายละเอียดเนื้อหาของข้อมูลนั้น ๆ Algorithm การเรียนรู้ของเครื่องในการแนะนำความชอบของผู้ใช้ ซึ่งจาก ลักษณะต่างๆ ที่เป็นรายละเอียดของข้อมูลนั้นๆ สามารถบ่งชี้ได้ว่าผู้ที่มีความชอบในลักษณะของเนื้อหาข้อมูล แบบใดก็อาจจะมีความชอบในเนื้อหาของข้อมูลที่มีลักษณะที่คล้ายคลึงกัน Algorithm ที่ใช้ในการวิเคราะห์ เนื้อหาของเอกสารนั้น เป็นการคำนวณหาความคล้ายคลึงระหว่างเอกสาร กับ Profile ของผู้ใช้ Algorithm ที่นิยมนำมาใช้ในการจัดกลุ่มเอกสารนั้น โดยทั่วไปแล้วมีด้วยกันหลายวิธี แต่วิธีที่นิยมใช้ในปัจจุบันนั้นมีด้วยกัน 4 วิธีซึ่งสามารถสรุป ได้ดังนี้

2.3.1 Naïve Bayes อาศัยหลักการของความน่าจะเป็น Byes' Theorem ง่ายต่อการสร้างตัวจำแนกหมวดหมู่และจำแนกได้เร็วความถูกต้องของผลลัพธ์น้อยกว่าวิธีการอื่น

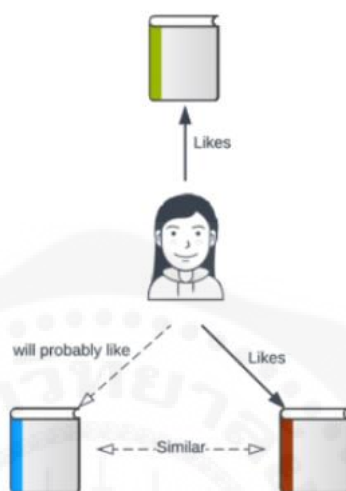
2.3.2 k-Nearest Neighbor ใช้การวัดระยะห่างระหว่างเอกสารด้วยวิธี EuclidianDistance ผลลัพธ์มีความถูกต้องสูงและเรียนรู้เร็ว รองรับการทำงานกับข้อมูลจำนวนมากหากมีคุณลักษณะที่ไม่เกี่ยวข้องต่อการจำแนกหมวดหมู่ของเอกสารมากจะทำให้ความถูกต้องลดลง

2.3.3 Rocchio สร้างตัวแยกเอกสาร โดยคติน้ำหนักของเทอม กับกลุ่ม ci ง่ายต่อการสร้าง ตัวจำแนกหมวดหมู่ถ้าเอกสารในกลุ่มมีแนวโน้มไม่อยู่ในกลุ่ม จะทำให้การจำแนกหมวดหมู่นั้นผิดพลาด

2.3.4 Support Vector Machine หรือ SVM สร้างเส้นตรงเพื่อแบ่งเขต ข้อมูล 2 กลุ่มออกจากกัน รองรับจำนวนคุณลักษณะได้มาก และมีความถูกต้องสูงต้องเลือก Kernel Function ที่เหมาะสม

Content-based เป็นวิธีการหนึ่งของการทำ Information Filtering โดยวิธีการทำงานของเทคนิคนี้ คือจะให้ความสนใจเนื้อหาของข้อมูลเป็นสำคัญ เช่น คุณลักษณะ (Feature) เพื่อค้นหาข้อมูล ที่ผู้ใช้นั้น สนใจ ซึ่งวิธีการของเทคนิคดังกล่าวจะไม่ประสบกับปัญหาการให้ Rating ต่อชิ้นข้อมูลที่ไม

ทั่วถึง และปัญหาขึ้นข้อมูลที่ยังไม่ได้ให้ Rating ดังนั้นวิธีการนี้จะต้องมีการคำนวณหาความคล้ายคลึงกัน (Similarity Measure) ระหว่างข้อมูล หรือสินค้าที่ระบบมีอยู่กับการต้องการในตัวข้อมูลหรือสินค้าที่ลูกค้าต้องการ ซึ่ง Algorithm ที่นิยมใช้มาก ที่สุดคือ การทำนายด้วยวิธีการหาสมาชิกที่ใกล้เคียงกัน ที่สุด K-NN (Knearest Neighbor) ซึ่งจะต้องมีการคำนวณหาค่าระยะห่างระหว่างข้อมูล (Standard Euclidean Distance) ข้อดีคือให้ผลลัพธ์ที่มีความถูกต้อง สูง เรียนรู้เร็ว และรองรับข้อมูลจำนวนมาก



ภาพที่ 2 ภาพแสดงลักษณะการทำงานของ Content Based Filtering

(ที่มา : <https://www.researchgate.net/figure/3-Recommendation-approaches-a-Collaborative-filtering>)

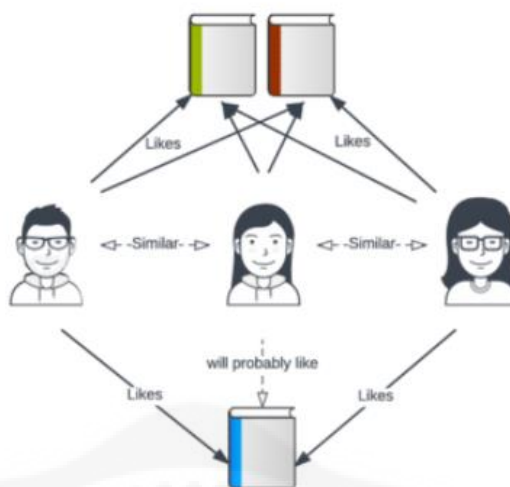
2.4 องค์ความรู้เกี่ยวกับ Collaborative Filtering [2]

ระบบแนะนำแบบ Collaborative Filtering จะทำการแนะนำวัตถุโดยใช้ข้อมูลในระบบเป็นจำนวนมากที่มีความคล้ายคลึงหรือใกล้เคียงกันกับผู้ใช้เช่น ความชอบ ความเคลื่อนไหว กิจกรรมของผู้ใช้คนอื่น ๆ ในระบบ ซึ่งจะเรียกเฉพาะเจาะจงได้ว่าระบบแนะนำแบบคัดกรองผู้ใช้ร่วม (User-based Collaborative Filtering Recommender System) แต่ถ้าใช้ข้อมูลลักษณะหรือคุณสมบัติของวัตถุอื่น ๆ ในระบบ ก็จะใช้เรียกอีกแบบหนึ่งว่าระบบแนะนำแบบคัดกรองสิ่งของร่วม (Item-based Collaborative Filtering Recommender System) กล่าวคือใช้ข้อมูลของสิ่งอื่น ๆ ในระบบมาช่วยในการคำนวณด้วย โดยปกติระบบแนะนำประเภทนี้จะมีขั้นตอนในการประมวลผล 3 ขั้นตอนใหญ่ๆ ดังนี้

2.4.1 สร้างประวัติผู้ใช้หรือวัตถุตามแต่ข้อมูลที่จะใช้เป็นพื้นฐานของระบบ

2.4.2 คัดเลือกผู้ใช้ที่เคยเลือกวัตถุนั้น (co-rated item) หรือวัตถุที่เคยถูกเลือกจากผู้ใช้คนดังกล่าว (couser rate) ที่มีความใกล้เคียงหรือคล้ายคลึงกันขึ้นมาตามจำนวนที่กำหนดไว้โดยการเปรียบเทียบประวัติผู้ใช้ หรือวัตถุตามแต่ข้อมูลที่ใช้เป็นพื้นฐานของระบบ ซึ่งอาจใช้ค่าความเหมือนโคไซน์ (Cosine Similarity) หรือระยะทางแบบยูคลิด (Euclidean Distance)

2.4.3 คำนวณว่าผู้ใช้และวัตถุนั้นมีความเหมาะสมกันมากน้อยเพียงใด โดยอาศัยข้อมูลที่หาได้ในขั้นตอนที่ 2.4.2



ภาพที่ 3 ภาพแสดงลักษณะการทำงานของ Collaborative Filtering

(ที่มา : <https://www.researchgate.net/figure/3-Recommendation-approaches-a-Collaborative-filtering>)

2.5 องค์ความรู้เกี่ยวกับ Hybrid Filtering [3]

ระบบแนะนำมากมายใช้วิธีการกรองข้อมูลแบบผสม (Hybrid Filtering) โดยการรวม Content-based และ Collaborative Filtering เข้าด้วยกัน ซึ่งช่วยหลีกเลี่ยงข้อจำกัดของ Content-based และ Collaborative Filtering ได้แต่อาจทำให้มีความซับซ้อนและใช้ทรัพยากรในการแนะนำสูง โดยหลายผลงานแสดงให้เห็นว่ามีความแม่นยำมากกว่า Content-based Filtering และ Collaborative Filtering

Hybrid Filtering จะรวมข้อมูลคุณลักษณะของวัตถุกับผู้ใช้และข้อมูล Rating เพื่อให้ได้ผลประโยชน์จากข้อมูลทั้งสองแบบ ระบบแนะนำแบบไฮบริดทำการรวมอัลกอริทึม A เข้ากับ B โดยพยายามที่จะใช้ข้อดีของ A ในการแก้ปัญหาของ B เช่น Collaborative Filtering จะมีปัญหาเรื่องการแนะนำวัตถุใหม่ เพราะวัตถุใหม่ไม่มีข้อมูล Rating แต่เรื่องนี้ไม่ใช่ปัญหาสำหรับ Content-based Filtering เพราะว่าการทำนายวัตถุใหม่มาจากตัววัตถุเอง (คุณลักษณะ) ซึ่งมีให้อยู่แล้วเมื่อวัตถุใหม่เข้าสู่ระบบ ซึ่งการสร้างระบบแนะนำแบบไฮบริด โดยการรวม Collaborative Filtering เข้ากับ Content-based Filtering เราสามารถก้าวข้ามข้อจำกัดต่างๆของแต่ละอัลกอริทึมได้ เช่น ปัญหา Cold-Start และ Popularity Bias

วิธีการที่จะนำ Content-Based Filtering และ Collaborative Filtering มารวมกันเป็นระบบแนะนำแบบ Hybrid Filtering สามารถแบ่งได้ ดังนี้

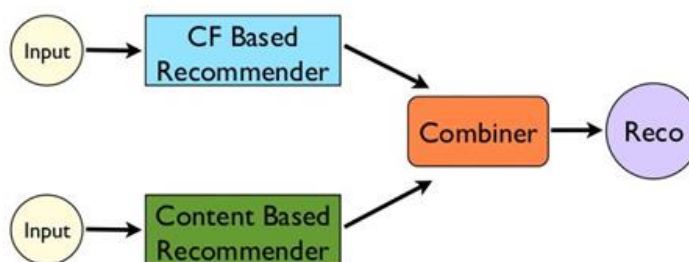
2.5.1 Combining Separate Recommenders เป็นทางเลือกหนึ่งในการสร้างระบบแนะนำแบบ Hybrid ซึ่งก็คือการใช้ Content-based และ Collaborative Filtering แยกกันคนละส่วน จากนั้นจะมีผลลัพธ์ออกมาสองแบบ ซึ่งแบบแรกสามารถรวมผลลัพธ์ที่ได้จากระบบแนะนำแต่ละแบบแล้วให้การแนะนำเดียวกัน โดยจะใช้วิธีการรวมแบบ Linear หรือการ Vote หรืออีกทางหนึ่งคือใช้ผลลัพธ์จากระบบแนะนำใดระบบแนะนำหนึ่งเท่านั้น ซึ่งจะเลือกจากผลลัพธ์ที่ดีที่สุดโดยอ้างอิงจากเมตริกซ์คุณภาพในการแนะนำ

2.5.2 Adding Content-based Characteristics to Collaborative Models เป็นการนำคุณสมบัติเฉพาะบางอย่างของวิธี Content-based มาใส่ไว้ในวิธี Collaborative Filtering ซึ่งมีหลายระบบที่ทำระบบแนะนำแบบ Hybrid ในรูปแบบนี้เช่น Fab หรือ Collaborative via Content ซึ่งใช้ Collaborative Filtering แบบดั้งเดิมข้อมูลความชอบของผู้ใช้จะเป็นแบบ Content-Based ซึ่งเป็นการแก้ปัญหาข้อมูลของผู้ใช้ที่มีการ Rating อยู่น้อยเกินไปซึ่งทำให้เกิดข้อจำกัดแบบ Sparsity

2.5.3 Adding Collaborative Characteristics to Content-based Models เป็นการนำคุณสมบัติเฉพาะบางอย่างของวิธี Collaborative Filtering มาใส่ไว้ในวิธี Content-based วิธีที่นิยมใช้กันคือ การใช้เทคนิคการลดมิติของกลุ่มข้อมูลผู้ใช้ที่อ้างอิงตาม Content ซึ่งตัวอย่างของวิธีการรูปแบบนี้ได้แก่ Latent Semantic Indexing ซึ่งเป็น Algorithm ที่ใช้ในการสร้างมุมมองแบบ Collaborative ให้กับ User Profile ที่แสดงในรูปเทอมของเวกเตอร์ ซึ่งจากผลลัพธ์จากวิธี Hybrid ในรูปแบบนี้ทำให้ระบบแนะนำมี ประสิทธิภาพเพิ่มขึ้น

2.5.4 Developing a Single Unifying Recommendation Model เป็นการสร้างโมเดลทางคณิตศาสตร์ จากลักษณะเฉพาะของวิธี Content-based และ Collaborative Filtering เช่น Bayesian Mixed-Effects Regression Models ซึ่งใช้วิธีการของ Markov Chain Monte Carlo ในการประมาณค่าพารามิเตอร์แล้วคาดคะเนหาค่า Rating มากที่มีความคล้ายคลึงหรือใกล้เคียงกันกับผู้ใช้ เช่น ความชอบ ความเคลื่อนไหว กิจกรรมของผู้ใช้คนอื่น ๆ ในระบบ ซึ่งจะเรียกเฉพาะเจาะจงได้ว่าระบบแนะนำแบบคัดกรองผู้ใช้ร่วม (User-based Collaborative Filtering Recommender System)

Hybrid Recommendations



ภาพที่ 4 ภาพแสดงลักษณะการทำงานของ Hybrid Recommendation
(ที่มา : <https://www.marutitech.com/recommendation-engine-benefits/>)

2.6 งานวิจัยที่เกี่ยวข้อง

2.6.1 งานวิจัยเรื่อง“OpenFace: A general-purpose face recognition library with mobile applications” [5]

Brandon Amos และคณะ ได้นำเสนอว่า OpenFace คือ Face recognition library ที่อยู่ภายใต้ใบอนุญาต Apache 2.0 ซึ่งใช้ FaceNet’s architecture และ dlib ในการตรวจจับใบหน้า ที่มีความแม่นยำมากกว่า openCV โดย OpenFace จะใช้ชุดข้อมูลเพียง 500,000 ข้อมูลในการ Training ซึ่งมีขนาดข้อมูลที่น้อยกว่า DeepFace และ FaceNet ที่เป็นชุดข้อมูลที่ได้รับการเผยแพร่ และยังมี ความแม่นยำในการประมวลผลการทำงานแม้จะมีชุดข้อมูลขนาดเล็กในการ Training โดยใช้เกณฑ์มาตรฐานการตรวจสอบของ LFW

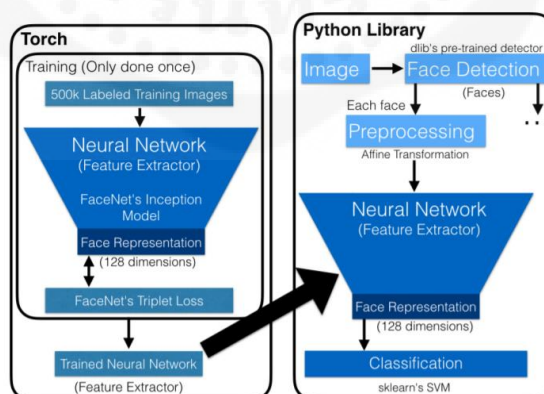


Figure 4: OpenFace's project structure.

ภาพที่ 5 ภาพแสดงกระบวนการ Train ของ OpenFace

(ที่มา : <http://reports-archive.adm.cs.cmu.edu/anon/anon/2016/CMU-CS-16-118.pdf>)

2.6.2 งานวิจัยเรื่อง “A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System” [6]

Geetha G และคณะ ได้เสนอระบบแนะนำภาพยนตร์ที่มีความสามารถในการแนะนำภาพยนตร์ให้กับผู้ใช้ใหม่เช่นเดียวกับผู้ใช้เก่า โดยการทำเหมืองฐานข้อมูลภาพยนตร์เพื่อรวบรวมข้อมูลสำคัญ เช่น ความนิยม หรือ ความน่าดึงดูด ซึ่งจำเป็นสำหรับการใช้ในการแนะนำ แล้วนำวิธี Hybrid filtering มาทำ Content-Based , Collaborative Filtering เพื่อนำระบบไปใช้ วิธีนี้จะเอาชนะข้อเสียของอัลกอริทึมแต่ละตัวและปรับปรุงประสิทธิภาพของระบบ การใช้เทคนิค เช่น Clustering , Similarity และ Classification จะทำให้ได้รับคำแนะนำที่ดีขึ้นซึ่งจะเป็นการเพิ่มความแม่นยำ

2.6.2 งานวิจัยเรื่อง “Book Recommendation System Based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining” [7]

Anand Shanker Tewari และคณะ ได้นำเสนอเป้าหมายของระบบแนะนำคือ การทำนายความสนใจของผู้ซื้อและทำการแนะนำหนังสือ ซึ่งการแนะนำหนังสือได้พิจารณาจากพารามิเตอร์ต่างๆ เช่น Content Based Filtering และคุณภาพของหนังสือโดยทำ Collaborative Filtering การให้คะแนนโดยผู้ซื้อรายอื่น ระบบผู้แนะนำนี้ยังใช้รูปแบบ Associative Model เพื่อให้การแนะนำดีขึ้น โดยระบบนี้ไม่มีปัญหาด้านประสิทธิภาพเนื่องจากการสร้างคำแนะนำแบบออฟไลน์

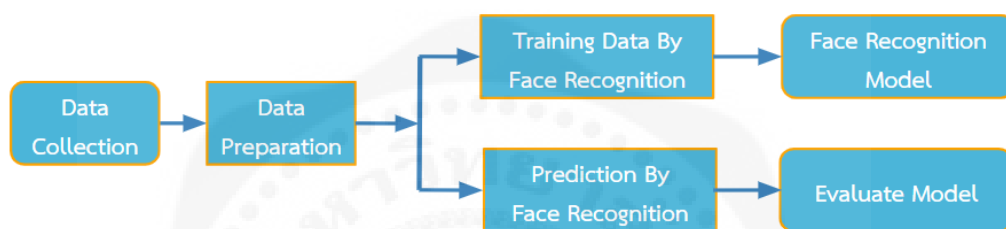
บทที่ 3

วิธีการดำเนินการศึกษาค้นคว้า

3.1 แผนการดำเนินงาน

ทางคณะผู้จัดทำได้แบ่งแผนการดำเนินงานออกเป็น 4 ส่วน คือ การเริ่มต้นและวางแผนโครงการ , การรู้จำใบหน้า (Face Recognition) , การแนะนำหนังสือ (Book Recommendation) และการพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ ดังที่แสดงในภาพที่ 6

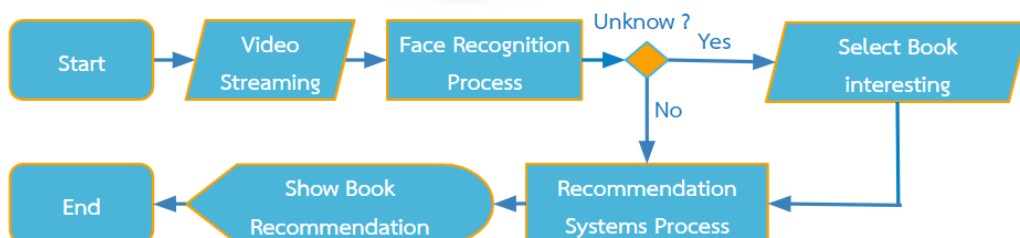
» การรู้จำใบหน้า (Face Recognition)



» การแนะนำหนังสือ (Book Recommendation)



» การพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ



ภาพที่ 6 แผนการดำเนินงานโดยภาพรวม

3.1.1 การเริ่มต้นและวางแผนโครงการ

→ วางแผนจัดทำโครงการ

การกำหนดหัวข้อโครงการที่ทางผู้จัดทำสนใจ

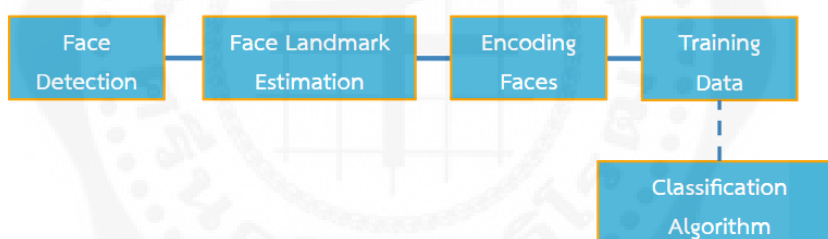
→ หาข้อมูลศึกษาแนวคิดและทฤษฎีที่เกี่ยวข้องกับงานวิจัย ค้นคว้าอุปกรณ์ที่เกี่ยวข้องต่างๆ

การศึกษางานวิจัย หลักการ ทฤษฎี เตรียมอุปกรณ์ เครื่องมือ โปรแกรม ที่ใช้ในการดำเนินโครงการ

3.1.2 การรู้จำใบหน้า (Face Recognition)

→ ศึกษาและวิเคราะห์ข้อมูลที่ได้จากอัลกอริทึมต่างๆ

การพัฒนาการจดจำใบหน้าโดยใช้ Python Library ที่ชื่อว่า face_recognition เป็นการหาลักษณะใบหน้าของแต่ละบุคคลจากรูปภาพให้เป็นตัวเลข เพื่อสามารถนำไปสร้าง Classification Model สำหรับการรู้จำใบหน้า ตามขั้นตอนดังที่แสดงในภาพที่ 7



ภาพที่ 7 ขั้นตอนการ Training Data By Face Recognition

→ การเก็บรวบรวมและเตรียมข้อมูล

เก็บข้อมูลด้วยวิธีการถ่ายภาพจากกล้อง web cam ของ iMac จะเก็บข้อมูลรูปภาพของนิสิตสาขาวิทยาการคอมพิวเตอร์จำนวน 20 คน โดยจะเก็บรูปถ่ายจำนวน 15 - 20 ภาพต่อคน

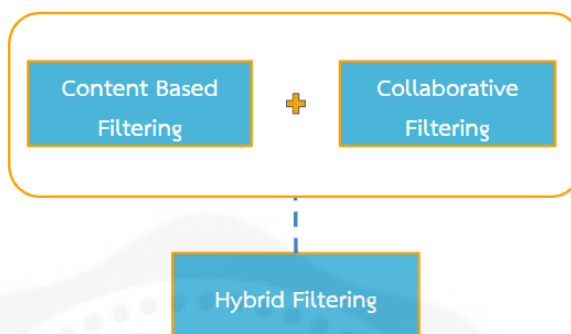
→ การพัฒนาแบบจำลอง (Modeling)

สร้าง Model Training ด้วย Classification Algorithm ทั้ง 4 แบบได้แก่ K-Nearest Neighbor, Gaussian Naive Bayes, Random Forest และ DecisionTree พร้อมทั้งทำการประเมินความถูกต้องด้วยการหาค่า Accuracy

3.1.3 การแนะนำหนังสือ (Book Recommendation)

→ ศึกษาและวิเคราะห์ข้อมูลที่ได้จากอัลกอริทึมต่างๆ

การพัฒนาระบบแนะนำหนังสือโดยใช้วิธีการกรองข้อมูลแบบผสม Hybrid Filtering โดยการรวม Content-based และ Collaborative Filtering เข้าด้วยกัน ดังที่แสดงในภาพที่ 8



ภาพที่ 8 วิธีการกรองข้อมูลแบบผสม Hybrid Filtering

→ การเก็บรวบรวมและเตรียมข้อมูล

เก็บข้อมูลด้วยวิธีการให้นักศึกษาวิทยาการคอมพิวเตอร์ เลือกหนังสือที่สนใจจาก Goodreads (เว็บไซต์สำหรับรีวิวหนังสือทั่วโลก) จำนวน 20 คน โดยจะเก็บบันทึกรายการหนังสือที่สนใจจำนวน 5 - 10 เล่มต่อคน

→ การพัฒนาแบบจำลอง (Modeling)

การพิจารณาเนื้อหา (Content-based) โดยใช้ Python Library ที่ชื่อว่า scikit-learn ในการสร้าง Vector Representations และ Cosine Similarity จะใช้ข้อมูลจาก ชื่อหนังสือ , ชื่อผู้แต่ง , ปีที่ตีพิมพ์ , ภาษา และ Tags ในการพิจารณาเพื่อหาความคล้ายคลึงกันของหนังสือนั้น ๆ

การพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) โดยใช้ Python Library ที่ชื่อว่า surprise ในการสร้าง Recommender Model สำหรับ Collaborative Filtering และสร้าง Model Training ด้วย Algorithm ทั้ง 10 แบบได้แก่ Singular Value Decomposition (SVD), SVD++, BaselineOnly, CoClustering, K Nearest Neighbors (KNN) Basic, KNNWithMeans, KNNWithZScore, Non-negative Matrix Factorization, SlopeOne และ NormalPredictor พร้อมทั้งทำการประเมินความถูกต้องด้วยการหาค่า Root Mean Square Error (RMSE)

จากการนำการทำงานของ Content-based และ Collaborative Filtering รวมเข้าด้วยกันจนเป็น Hybrid Filtering โดยจะใช้ Content-based ในการหาหนังสือที่คล้ายคลึงกับหนังสือที่ผู้ใช้งานเคยอ่าน จากนั้นจะใช้ Collaborative Filtering ในการคำนวณการคาดเดาเป็นค่า Estimated prediction ของหนังสือที่มีโอกาสที่ผู้ใช้งานจะชื่นชอบ แล้วจะนำค่าที่ได้มาเรียงลำดับ 10 ลำดับแรกของหนังสือที่ผู้ใช้งานอาจจะชื่นชอบ

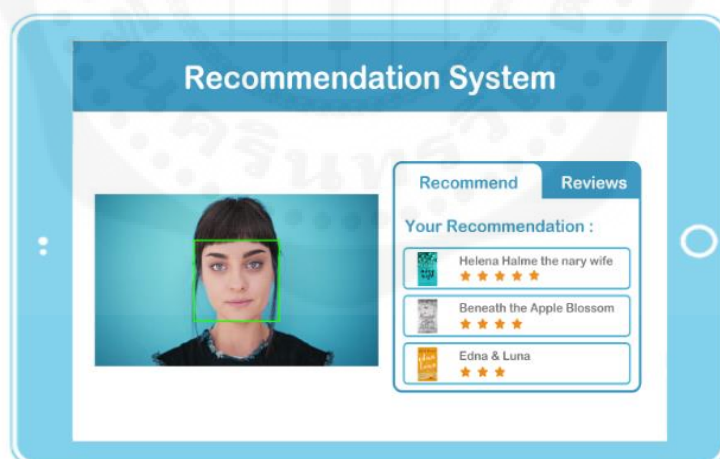
3.1.4 การพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ

→ การเตรียมข้อมูล

รวบรวม Face Recognition Model , Recommender Model และข้อมูลหนังสือ

→ การพัฒนาเว็บแอปพลิเคชัน

การพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ โดยใช้ Flask เป็น framework ที่ใช้ในการสร้าง Web Application ที่พัฒนาด้วยภาษา Python ซึ่งจะแบ่งการพัฒนาระบบเป็น 2 ส่วน คือการรู้จำใบหน้าผ่าน Video Streaming และการแนะนำหนังสือจากข้อมูลผู้ซื้อ



ภาพที่ 9 ตัวอย่างเว็บแอปพลิเคชัน

→ การประเมินผล

→ สรุปและเผยแพร่ผลงานวิจัย

3.3 อุปกรณ์และเครื่องมือที่ใช้ในโครงการ

3.3.1 ฮาร์ดแวร์ (Hardware)

- iMac
 - iMac (21.5-inch, Late 2015)
 - Processor 2.8 GHz Intel Core i5
 - RAM 8 GB
 - MacOS Mojave

3.3.2 ซอฟต์แวร์ (Software)

- OpenFace [4]

OpenFace เป็นการใช้งาน Python และ Torch ในการจดจำใบหน้าที่ด้วย Deep Neural Networks (DNN) จากงานวิจัย CVNet 2015 FaceNet: A Unified Embedding for Face Recognition and Clustering โดย Florian Schroff, Dmitry Kalenichenko และ James Philbin ที่ Google โดย Torch ช่วยให้ Neural Network สามารถดำเนินการบน CPU หรือกับ CUDA ได้
- dlib
- Anaconda
- Visual Studio Code (VS Code) เป็นโปรแกรมแก้ไขและปรับแต่งโค้ด (code optimized editor)
- Flask เป็น framework ที่ใช้ในการสร้าง Web Application ที่พัฒนาด้วยภาษา Python

3.3.3 ภาษาที่ใช้

- JavaScript
- Html
- Css
- Python

3.4 ชุดข้อมูล

3.4.1 Goodbooks-10k

- o เป็นข้อมูลจาก Goodreads (เว็บไซต์สำหรับรีวิวหนังสือทั่วโลก) แหล่งที่มา

<https://www.kaggle.com/zygmunt/goodbooks-10k>

โดยในข้อมูลประกอบด้วย

- o books.csv

id	int64
book_id	int64
best_book_id	int64
work_id	int64
books_count	int64
isbn	object
isbn13	float64
authors	object
original_publication_year	float64
original_title	object
title	object
language_code	object
average_rating	float64
ratings_count	int64
work_ratings_count	int64
work_text_reviews_count	int64
ratings_1	int64
ratings_2	int64
ratings_3	int64
ratings_4	int64
ratings_5	int64
image_url	object
small_image_url	object

ภาพที่ 10 ภาพแสดง Column ของ books.csv

- o book_tags.csv

goodreads_book_id	int64
tag_id	int64
count	int64

ภาพที่ 11 ภาพแสดง Column ของ book_tags.csv

- o ratings.csv

book_id	int64
user_id	int64
rating	int64

ภาพที่ 12 ภาพแสดง Column ของ ratings.csv

- o tags.csv

tag_id	int64
tag_name	object
..	..

ภาพที่ 13 ภาพแสดง Column ของ tags.csv

บทที่ 4

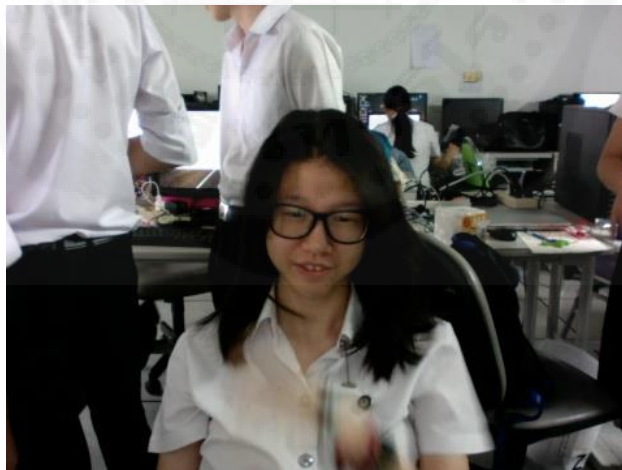
ผลการดำเนินงาน

4.1 ผลลัพธ์จากการเก็บข้อมูลในส่วนของการรู้จำใบหน้า (Face Recognition)

การเก็บข้อมูลด้วยวิธีการถ่ายภาพจากกล้อง web cam ของ iMac จะเก็บข้อมูลรูปภาพของนิสิตสาขาวิทยาการคอมพิวเตอร์จำนวน 20 คน โดยจะเก็บรูปถ่ายจำนวน 15 - 20 ภาพต่อคน จากนั้นใช้ Python Library ที่ชื่อว่า face_recognition หาลักษณะใบหน้าของแต่ละบุคคลจากรูปภาพให้เป็นตัวเลข จากการถ่ายภาพทั้งหมดได้จำนวน 394 รูปภาพ เนื่องจากรูปภาพมีความละเอียดและความสว่างของภาพไม่เพียงพอจึงเก็บข้อมูลได้เพียง 373 รูปภาพ



ภาพที่ 14 ตัวอย่างของรูปถ่ายที่ไม่สามารถเก็บข้อมูลได้



ภาพที่ 15 ตัวอย่างของรูปถ่ายที่สามารถเก็บข้อมูลได้

	X	Y
0	[-0.12503135204315186, 0.009990954771637917, 0...	lin
1	[-0.12827068567276, 0.03196568787097931, 0.040...	lin
2	[-0.09924603998661041, 0.04643559828400612, 0....	lin
3	[-0.12632755935192108, 0.05037033185362816, 0....	lin
4	[-0.13333402574062347, 0.01578657701611519, 0....	lin
5	[-0.09963660687208176, 0.03891250118613243, 0....	lin
6	[-0.09649359434843063, 0.027591507881879807, 0...	lin
7	[-0.09951993823051453, 0.026294175535440445, 0...	lin
8	[-0.10379182547330856, 0.026957053691148758, 0...	lin
9	[-0.10390745848417282, 0.01627863571047783, 0....	lin
10	[-0.09065660089254379, 0.007521618157625198, 0...	lin
11	[-0.11987138539552689, 0.01512322761118412, 0....	lin

ภาพที่ 16 ตัวอย่างข้อมูลของรูปถ่ายที่สามารถเก็บข้อมูลได้

4.2 ผลลัพธ์จากการพัฒนาแบบจำลอง (Modeling) ในส่วนของการรู้จำใบหน้า (Face Recognition)

จากข้อ 4.1 ข้อมูลที่สามารถเก็บได้จะนำมาแบ่งข้อมูลเป็น Train data 70% และ Test data 30% แล้วนำมาวิเคราะห์กับอัลกอริทึมกับ K-Nearest Neighbor, Gaussian Naive Bayes, Random Forest และ Decision Tree

จากการทดลองกับอัลกอริทึมทั้ง 4 แบบ ได้ผลการทดสอบออกมาดังตารางที่ ...

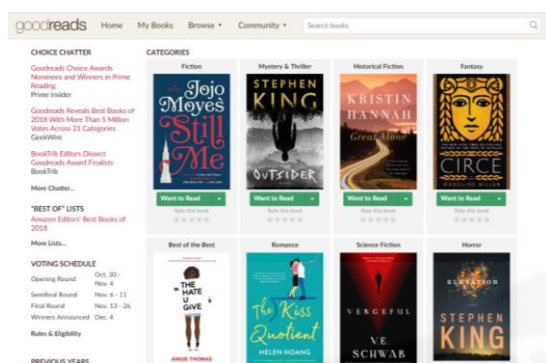
เทคนิคการแบ่งกลุ่ม	ความแม่นยำในการระบุบุคคลจาก ใบหน้า
K-Nearest Neighbor	100%
Gaussian Naive Bayes	99.1%
Random Forest	99.1%
Decision Tree	83.9%

ตารางที่ 1 ผลการทดสอบความแม่นยำของการรู้จำใบหน้า (Face Recognition)

จากตารางค่าความแม่นยำในการแบ่งกลุ่มของวิธี K-Nearest Neighbor มีความแม่นยำสูงสุด คือ 100%

4.3 ผลลัพธ์จากการเก็บข้อมูลในส่วนของการแนะนำหนังสือ (Book Recommendation)

การเก็บข้อมูลด้วยวิธีการให้นิสิตสาขาวิทยาการคอมพิวเตอร์ เลือกหนังสือที่สนใจจาก Goodreads (เว็บไซต์สำหรับรีวิวหนังสือทั่วโลก) จำนวน 20 คน โดยจะเก็บบันทึกรายการหนังสือที่สนใจจำนวน 5 - 10 เล่มต่อคน จากการเก็บบันทึกทั้งหมดได้จำนวน 160 ข้อมูล



user	user_id	book_id
nam	18052539	8337
nam	18052539	6310
nam	18052539	6319
nam	18052539	70720
nam	18052539	18601351
nam	18052539	19321
nam	18052539	39988
nam	18052539	3
nam	18052539	11127

ภาพที่ 17 ตัวอย่างของการเก็บข้อมูลของการแนะนำหนังสือ

4.4 ผลลัพธ์จากการพัฒนาแบบจำลอง (Modeling) ในส่วนของการแนะนำหนังสือ (Book Recommendation)

การพัฒนาระบบแนะนำหนังสือโดยใช้วิธีการกรองข้อมูลแบบผสม Hybrid Filtering โดยการรวม Content-based และ Collaborative Filtering เข้าด้วยกัน

การพิจารณาเนื้อหา (Content-based) ผลจากการสร้าง Vector Representations และ Cosine Similarity พบว่าใช้ Python Library ที่ชื่อว่า scikit-learn ในส่วนของ TfidfVectorizer และ cosine_similarity สามารถสร้างการพิจารณาเนื้อหาได้มีความคล้ายคลึงกันของหนังสือขึ้น ๆ

```

The Hunger Games (The Hunger Games, #1) Suzanne Collins 2008 eng
1997 0.000000
1997 eng 0.000000
2000 0.000000
2000 eng 0.000000
2000s 0.000000
2000s memoir 0.000000
2003 0.000000
2003 eng 0.000000
2006 0.000000
2006 popular 0.000000
2008 0.052075
2008 eng 0.052075
2009 0.000000
2009 eng 0.000000
2010 0.052075
2010 survival 0.052075
2011 0.000000

```

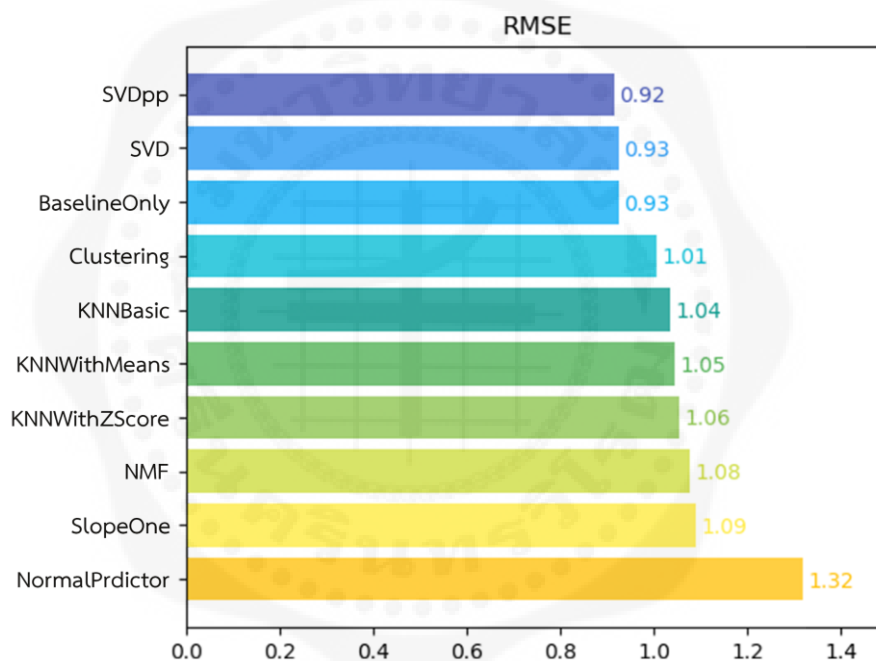
ภาพที่ 18 ตัวอย่างของ Vector Representations โดยใช้ TfidfVectorizer ใน scikit-learn

```
array([[1.          , 0.40101996, 0.2730969 , 0.13140014, 0.23248209],
       [0.40101996, 1.          , 0.27854185, 0.13360883, 0.25072769],
       [0.2730969 , 0.27854185, 1.          , 0.25654172, 0.45514957],
       [0.13140014, 0.13360883, 0.25654172, 1.          , 0.21671165],
       [0.23248209, 0.25072769, 0.45514957, 0.21671165, 1.          ]])
```

ภาพที่ 19 ตัวอย่างของ Cosine Similarity โดยใช้ cosine_similarity ใน scikit-learn

การพึ่งพาผู้ใช้ร่วม (Collaborative Filtering) ผลจากการสร้าง Recommender Model สำหรับ Collaborative Filtering นำมาวิเคราะห์กับอัลกอริทึมกับ Algorithm ทั้ง 10 แบบได้แก่ Singular Value Decomposition (SVD), SVD++, BaselineOnly, CoClustering, K Nearest Neighbors (KNN) Basic, KNNWithMeans, KNNWithZScore, Non-negative Matrix Factorization, SlopeOne และ NormalPredictor

จากการทดลองกับอัลกอริทึมทั้ง 10 แบบ ได้ผลการทดสอบออกมาดังรูปภาพที่ 20



ภาพที่ 20 กราฟแสดงผลทดลองกับอัลกอริทึมทั้ง 10 แบบของ Collaborative Filtering

ซึ่งเมื่อนำ SVD ไปปรับปรุง parameter ทำให้มีค่า Root Mean Square Error (RMSE) อยู่ที่ 0.92 ซึ่ง ใช้เวลาในการ train และ test ไม่นาน ดังนั้น SVD จึงเป็นอัลกอริทึมที่เหมาะสมในการสร้าง Recommender Model สำหรับ Collaborative Filtering

การพัฒนาระบบแนะนำหนังสือ ผลจากการสร้างโดยจะใช้วิธีการกรองข้อมูลแบบผสม (Hybrid Filtering) ในการนำ Content-based และ Collaborative Filtering มาใช้ร่วมกันในการคำนวณการคาดเดาเป็นค่า Estimated prediction สามารถแนะนำรายการหนังสือที่ผู้ใช้งานมีโอกาสชื่นชอบ

[] hybrid(7)

The Oresteia (Ορέστεια, #1-3)
White Teeth
Naked
Extremely Loud and Incredibly Close

	authors	book_id	est	language_code	original_publication_year	title
30	Stephen King	5091	4.046910	en-US	2004	The Dark Tower (The Dark Tower, #7)
23	William Shakespeare	8852	3.992864	eng	1606	Macbeth
89	Paolo Bacigalupi	6597651	3.899275	eng	2009	The Windup Girl
87	Lisa Scottoline	25773504	3.899275	eng	2016	Most Wanted
86	Sophia Amoruso	18667945	3.899275	eng	2014	#GIRLBOSS
85	Brandon Sanderson	16065004	3.899275	eng	2015	Shadows of Self (Mistborn, #5)
84	E.B. White, Garth Williams, Rosemary Wells	24178	3.899275	eng	1952	Charlotte's Web
83	Abbi Glines	22522202	3.899275	eng	2015	Until Friday Night (The Field Party, #1)
82	Mary Roach	7237456	3.899275	en-US	2010	Packing for Mars: The Curious Science of Life ...
81	Alex Berenson	341553	3.899275	en-US	2006	The Faithful Spy (John Wells, #1)

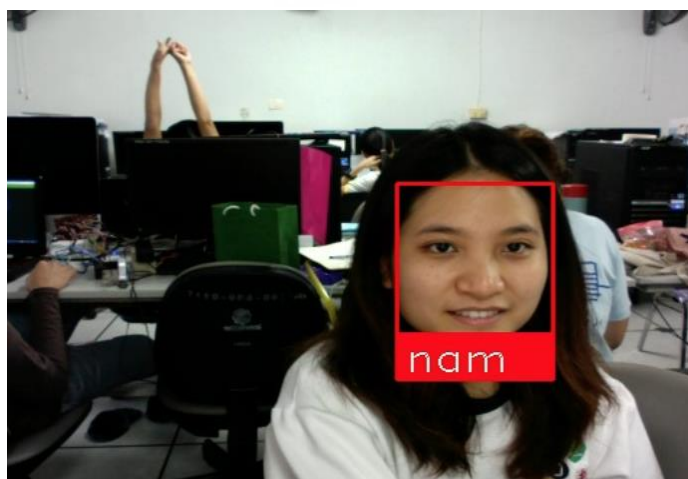
ภาพที่ 21 ภาพตัวอย่างของการแนะนำหนังสือโดยใช้การกรองข้อมูลแบบผสม Hybrid Filtering

4.5 ผลลัพธ์จากการพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ

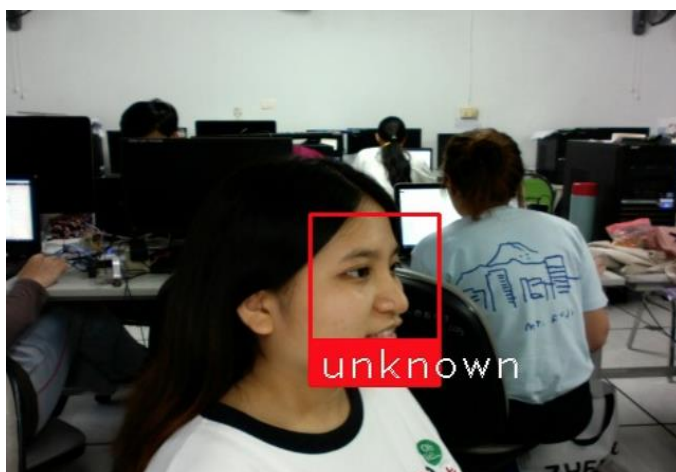
จากการพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ ซึ่งจะแบ่งการพัฒนา
ระบบเป็น 2 ส่วน คือการรู้จำใบหน้าผ่าน Video Streaming และการแนะนำหนังสือจากข้อมูลผู้ซื้อ

ผลลัพธ์ของการรู้จำใบหน้าผ่าน Video Streaming สามารถตรวจจับและระบุตัวบุคคลได้ แต่
ยังมีข้อจำกัด คือ

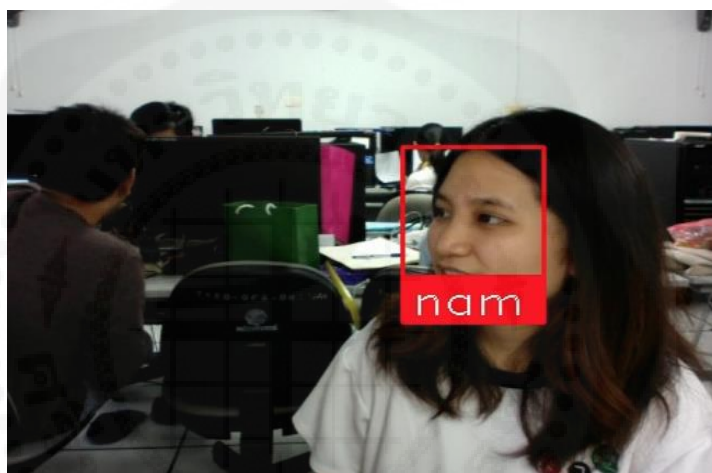
1. ระยะห่างจากกล้องและบุคคลโดยสามารถจับภาพใบหน้าได้ดีในระยะ 0.7 เมตร จับภาพ
ใบหน้าได้ในระดับปานกลางในระยะ 1.5 เมตร จับภาพใบหน้าได้ในระดับต่ำในระยะ 2.4 เมตร และ
ไม่สามารถจับภาพใบหน้าในระยะห่างมากกว่า 2.4 เมตร
2. กล้องไม่สามารถตรวจภาพของใบหน้าในมุมที่ไม่มีดวงตาทั้งสองข้าง
3. เนื่องจากตรวจจับใบหน้าและระบุตัวบุคคลด้วย frame ของ Video Streaming จะมีโอกาส
ที่ frame ที่ได้รับเป็นภาพที่ไม่สามารถนำไปตรวจจับได้ และในส่วนผลลัพธ์ของการแนะนำหนังสือจาก
ข้อมูลผู้ซื้อ สามารถแนะนำรายการสินค้าได้ตามความสนใจของลูกค้า



ภาพที่ 22 ภาพตัวอย่างการรู้จำใบหน้าผ่าน Video Streaming



ภาพที่ 23 ภาพตัวอย่างหันด้านข้าง โดยเห็นดวงตาข้างเดียว



ภาพที่ 24 ภาพตัวอย่างหันด้านข้าง โดยเห็นดวงตาทั้งสองข้าง

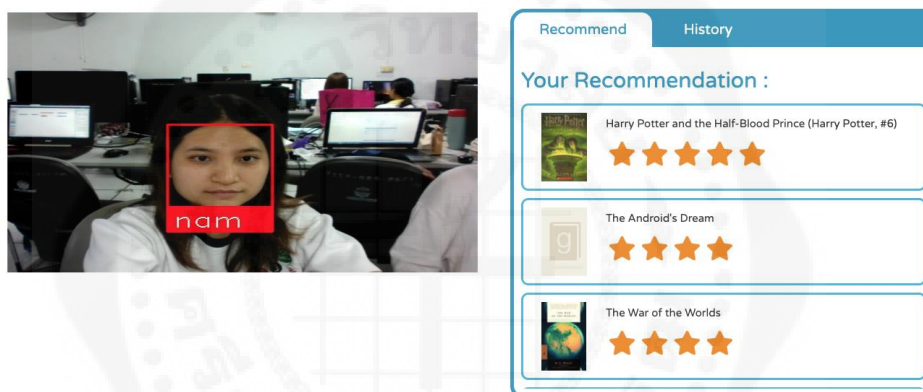


ภาพที่ 25 ภาพตัวอย่างการจับภาพใบหน้าในระยะ 1.5 เมตร



ภาพที่ 26 ภาพตัวอย่างการจับภาพใบหน้าในระยะ 2.4 เมตร

Recommendation System



ภาพที่ 27 ภาพตัวอย่างเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือ

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลการดำเนินโครงการ

จากผลลัพธ์ในส่วนของการรู้จำใบหน้า (Face Recognition) ขั้นตอนการเก็บข้อมูลถ่ายภาพจากกล้อง web cam ของ iMac แสดงให้เห็นว่าเนื่องจากชุดข้อมูลที่เป็นรูปภาพ มีความละเอียดและความสว่างของภาพไม่เพียงพอ จึงทำให้ไม่สามารถใช้ได้ทุกรูป และขั้นตอนการพัฒนาแบบจำลอง (Modeling) อัลกอริทึมในการแบ่งกลุ่มของวิธี K-Nearest Neighbor มีความแม่นยำสูงสุดคือ 100%

จากผลลัพธ์ในส่วนของการแนะนำหนังสือ (Book Recommendation) ขั้นตอนการเก็บข้อมูลให้บัณฑิตสาขาวิทยาการคอมพิวเตอร์เลือกหนังสือที่สนใจจาก Goodreads (เว็บไซต์สำหรับรีวิวหนังสือทั่วโลก) เพิ่มจากชุดข้อมูล Goodbooks-10k แสดงให้เห็นว่า รายการหนังสือที่ถูกแนะนำให้ความสนใจมากกว่าเดิม และขั้นตอนการพัฒนาแบบจำลอง (Modeling) เมื่อนำอัลกอริทึม SVD ไปปรับปรุง parameter ทำให้มีค่า Root Mean Square Error (RMSE) อยู่ที่ 0.92 ซึ่ง ใช้เวลาในการ train และ test ไม่นาน ดังนั้น SVD จึงเป็นอัลกอริทึมที่เหมาะสมในการสร้าง Recommender Model สำหรับ Collaborative Filtering โดยการสร้างระบบแนะนำหนังสือจะใช้วิธีการกรองข้อมูลแบบผสม (Hybrid Filtering) ซึ่งคือนำ Content-based และ Collaborative Filtering มาใช้ร่วมกันในการคำนวณการคาดเดาเป็นค่า Estimated prediction สามารถแนะนำรายการหนังสือที่ผู้ใช้งานมีโอกาสชื่นชอบ

จากผลลัพธ์ในส่วนของการพัฒนาเว็บแอปพลิเคชันระบบส่งเสริมการขายสำหรับร้านหนังสือสามารถแนะนำรายการสินค้าได้ตามความสนใจของลูกค้า และการรู้จำใบหน้าผ่าน Video Streaming สามารถตรวจจับและระบุตัวบุคคลได้ แต่ยังมีข้อจำกัดด้านระยะห่างของกล้องกับตัวบุคคล , มุมของใบหน้าขณะจับภาพ และ คุณภาพของ frame ใน Video Streaming

5.2 ปัญหาและอุปสรรคในการดำเนินโครงการ

5.2.1 การเก็บข้อมูลรูปภาพที่ใช้ในการรู้จำใบหน้าจำเป็นต้องใช้ความละเอียดเท่ากับกล้องที่ใช้ในระบบ

5.2.2 โมเดลที่ได้จาก OpenFace ไม่สามารถนำไปลงในเว็บแอปพลิเคชันได้

5.2.3 ในการทดสอบระบบจำเป็นต้องใช้คอมพิวเตอร์ที่มีสเปคสูง

5.2.4 ขั้นตอนการลง Dlib ในคอมพิวเตอร์ระบบปฏิบัติการ Window10 มีความยุ่งยาก

5.3 ข้อเสนอแนะ

ในอนาคตหากร้านหนังสือมีการเก็บข้อมูลของผู้ซื้อที่ประเทศไทยมากขึ้น จะทำให้ระบบแนะนำหนังสือสามารถแนะนำหนังสือได้เหมาะสมมากยิ่งขึ้น เนื่องจากชุดข้อมูลยังมีการรีวิวหนังสือภาษาไทยที่น้อย อาจยังไม่ตรงตามความสนใจของคนไทย



บรรณานุกรม

- [1] Mindphp.com. “เทคโนโลยีจดจำใบหน้า Face Recognition (เฟซ รีคognition)” . สืบค้นเมื่อ 20/11/18. <http://mindphp.com/บทความ/240-ai-machine-learning/5227-face-recognition.html>.
- [2] พันตรีหญิง รัฐวรรณ พันธุนิล, ผู้ช่วยศาสตราจารย์ ดร. นิเวศ จิระวิจิตชัย. “A PERSONALIZED RECOMMENDER SYSTEM USING USER-BASED AND ITEM-BASED COLLABORATIVE FILTERING METHOD” . สืบค้นเมื่อ 25/1/19. http://ereseach.library.ssru.ac.th/bitstream/123456789/252/6/ird_083_55%20%285%29.pdf.
- [3] MAYA.HRISTAKEVA. “A PRACTICAL GUIDE TO BUILDING RECOMMENDER SYSTEMS” .<https://buildingrecommenders.wordpress.com/2015/11/20/overview-of-recommender-algorithms-part-1/>. สืบค้นเมื่อ 25/1/19.
- [4] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. “OpenFace” . สืบค้นเมื่อ 25/1/19. <https://cmusatyalab.github.io/openface/#overview>.
- [5] Brandon Amos; Bartosz Ludwiczuk; Mahadev Satyanarayanan. (June 2016). “OpenFace: A general-purpose face recognition library with mobile applications”. The research was supported by the National Science Foundation (NSF) under grant number CNS-1518865,.
- [6] Geetha G, Safa M, Fancy C, Saranya D. (2018). “A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System” . National Conference on Mathematical Techniques and its Applications (NCMTA 18) ,IOP Publishing,.
- [7] Anand Shanker Tewari, Abhay Kumar, Asim Gopal Barman. (2014). “Book Recommendation System Based on Combine Features of Content Based Filtering, Collaborative Filtering and Association Rule Mining” . IEEE International Advance Computing Conference (IACC), IEEE,.

ภาคผนวก

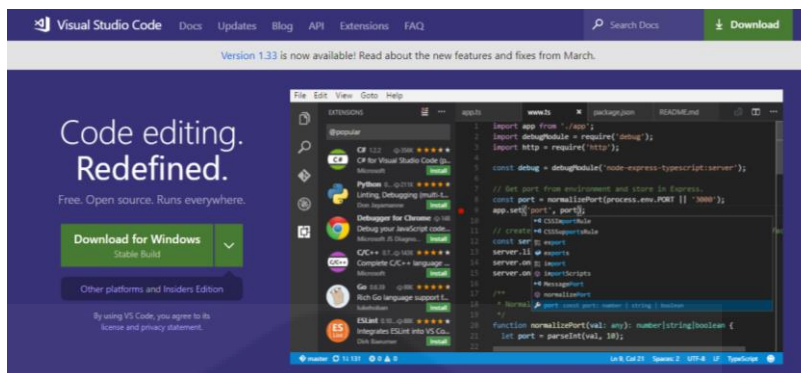


ภาคผนวก ก

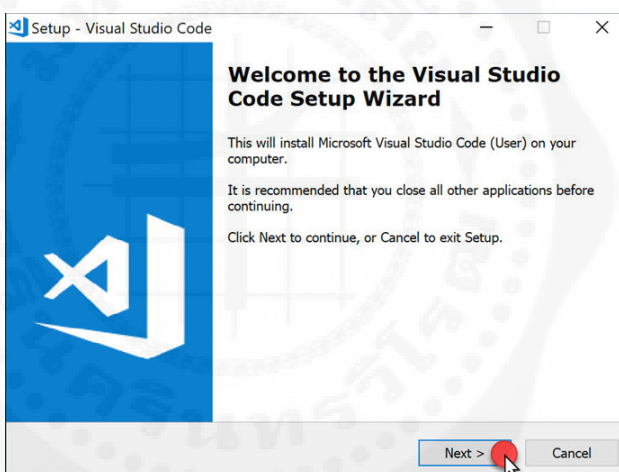
สอนวิธีการการติดตั้งโปรแกรมที่ใช้พัฒนา

Visual Studio Code

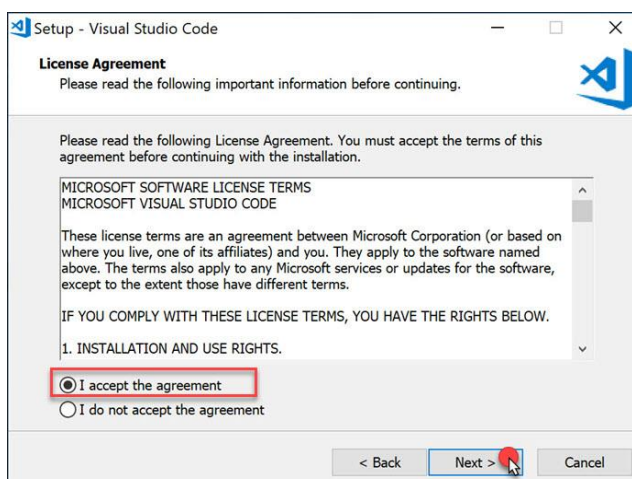
1. เข้าเว็บไซต์ <https://code.visualstudio.com> เพื่อติดตั้ง Visual Studio Code ดังรูป



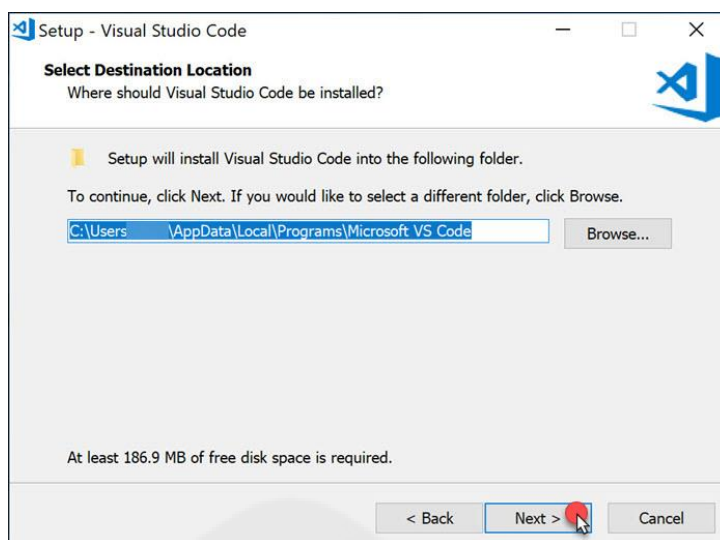
2. เมื่อเปิดตัวติดตั้งขึ้นมาแล้ว จะแสดงหน้าจอ Welcome to the Visual Studio Code Setup Wizard ให้กดปุ่ม Next >



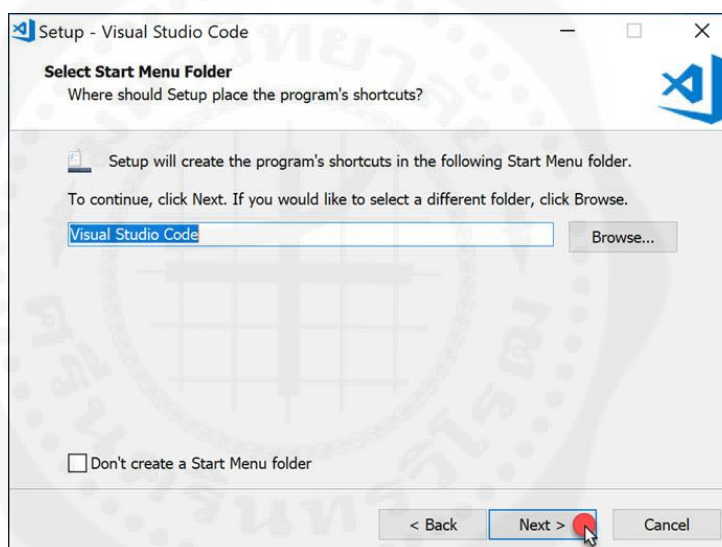
3. หลังจากนั้นจะแสดงหน้าจอ License Agreement ให้เลือก ☐ I accept the agreement แล้วกด Next >



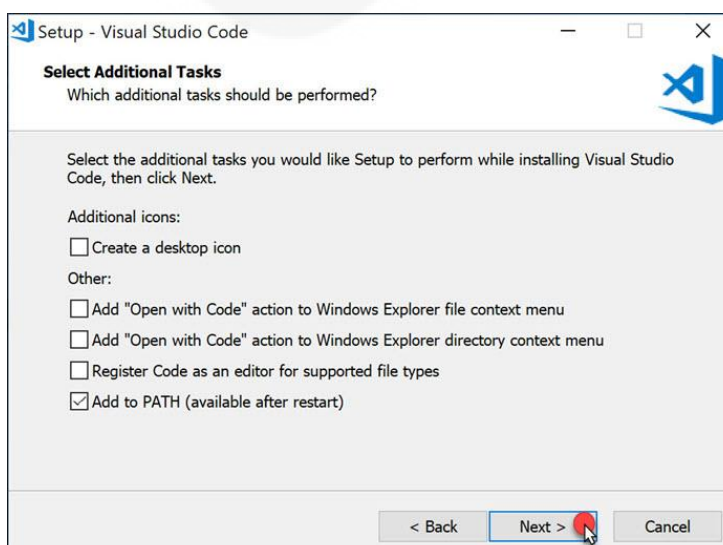
4. หน้าจอจะแสดง Select Destination Location ให้กดปุ่ม Next >



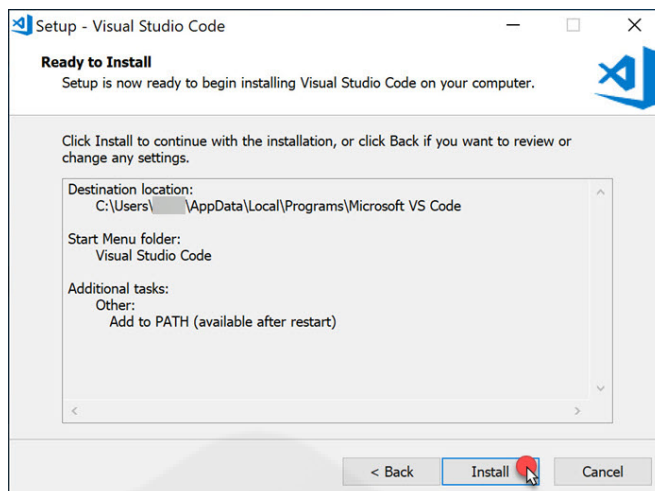
5. หน้าจอจะแสดง Select Start Menu Folder ให้กดปุ่ม Next >



6. หน้าจอจะแสดง Select Additional Tasks ให้กดปุ่ม Next >



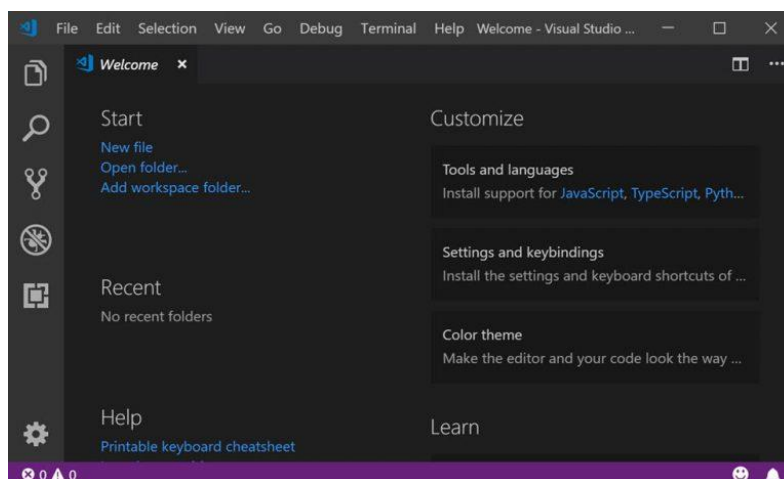
7. หน้าจอจะแสดง Ready to Install ให้กดปุ่ม Install



8. เมื่อติดตั้งเสร็จแล้วจะแสดงหน้าจอ Completing the Visual Studio Code Setup Wizard ให้กดปุ่ม Finish



9. จากนั้นโปรแกรม Visual Studio Code เปิดขึ้นมาก็จะเป็นการเสร็จสิ้นขั้นตอนในการติดตั้งโปรแกรม Visual Studio Code



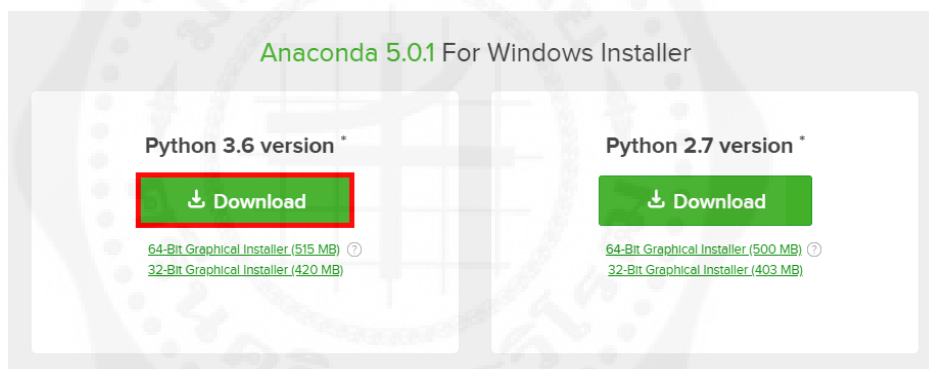
ภาคผนวก ข

Anaconda

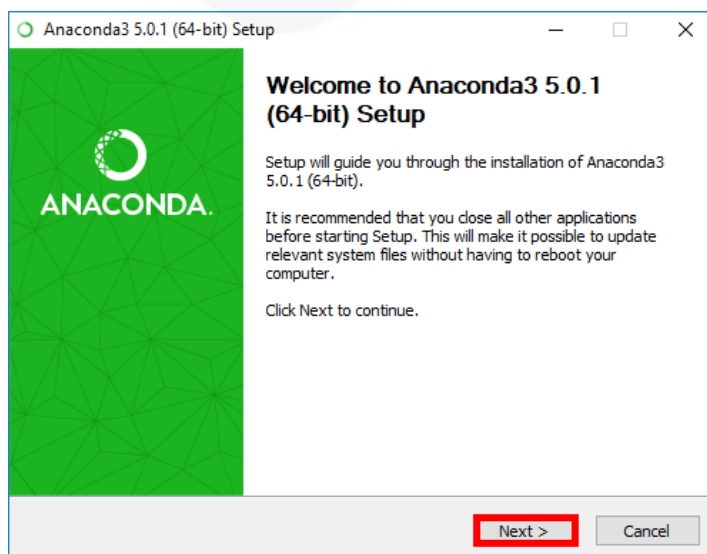
1. เข้าเว็บไซต์ <https://www.anaconda.com/download/> เพื่อติดตั้ง Anaconda ดังรูป เลือกดาวน์โหลดเวอร์ชันสำหรับระบบปฏิบัติการ Windows



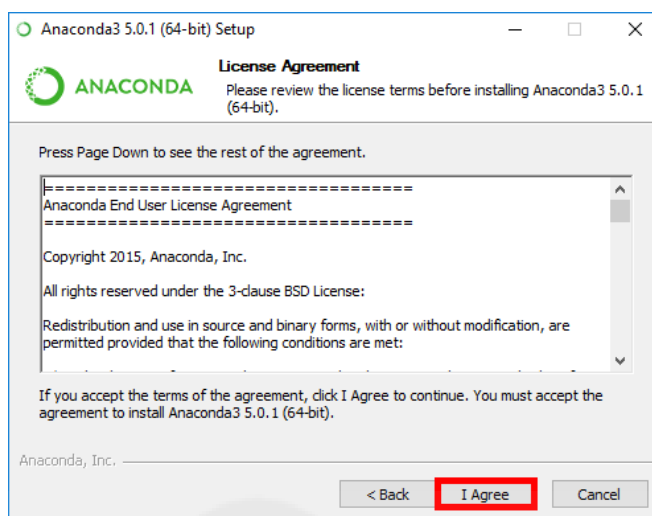
2. คลิกที่ปุ่ม Download (Python 3.6 version)



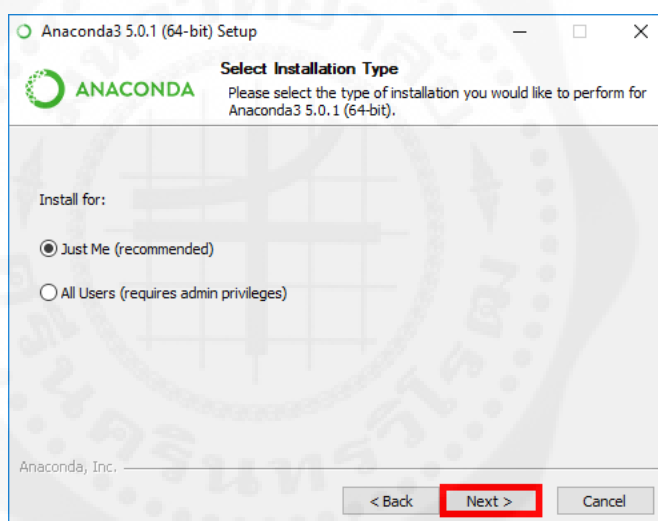
3. เมื่อเปิดตัวติดตั้งขึ้นมาแล้ว จะแสดงหน้าจอ Welcome to Anaconda3 5.0.1 (64-bit) Setup ให้กดปุ่ม Next >



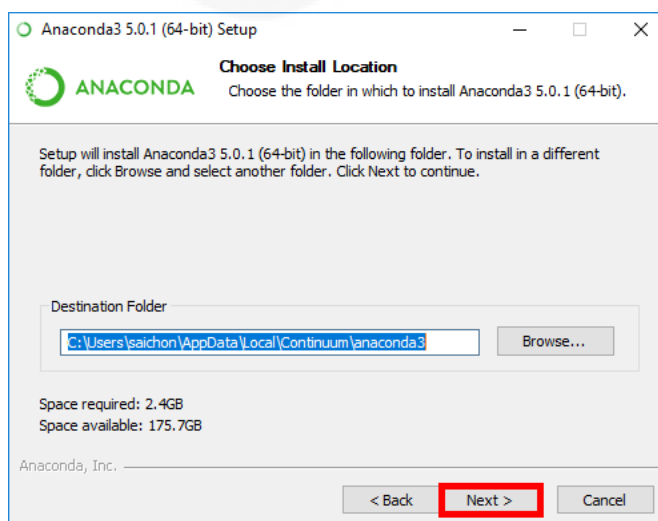
4. หน้าจอจะแสดง License Agreement กดที่ปุ่ม I Agree



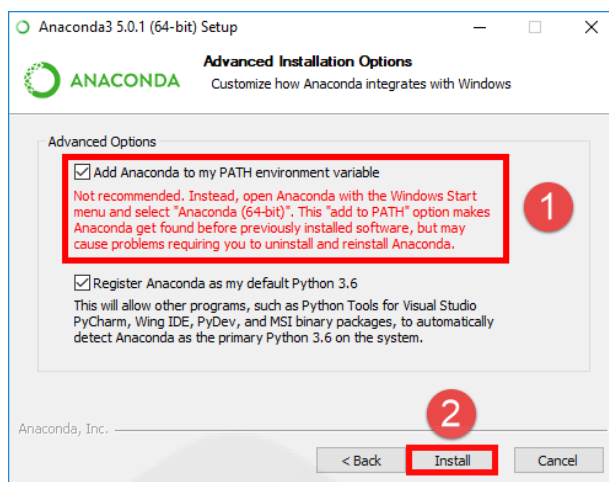
5. หน้าจอจะแสดง Select Installation Type เลือก Just Me (recommended) แล้วกดปุ่ม Next >



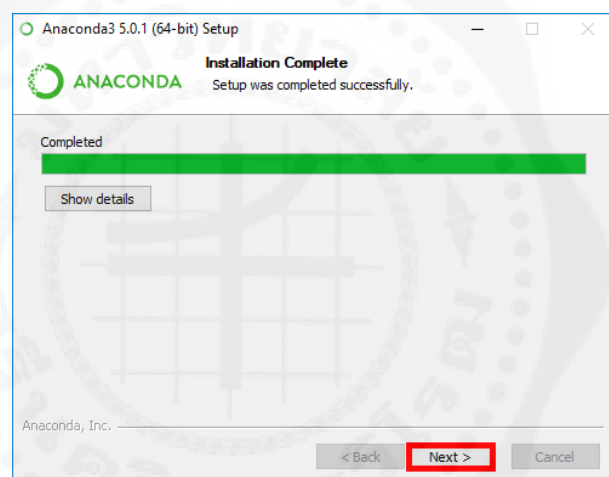
6. หน้าจอจะแสดง Choose Install Location คลิกที่ปุ่ม Next >



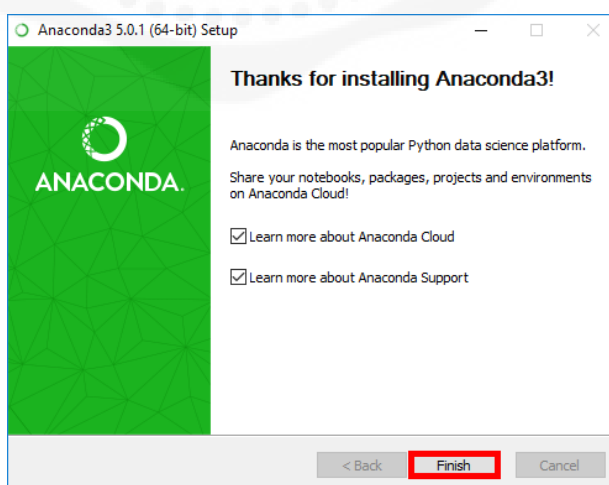
7. หน้าจอจะแสดง Advanced Installation Options เลือก Add Anaconda to my PATH environment variable แล้วกดปุ่ม Install



8. รอก่อนว่าการติดตั้งจะเสร็จสมบูรณ์ เมื่อติดตั้งเสร็จสมบูรณ์แล้วให้คลิกปุ่ม Next >



9. กดที่ปุ่ม Finish จะเป็นการเสร็จสิ้นขั้นตอนในการติดตั้ง Anaconda



ภาคผนวก ค

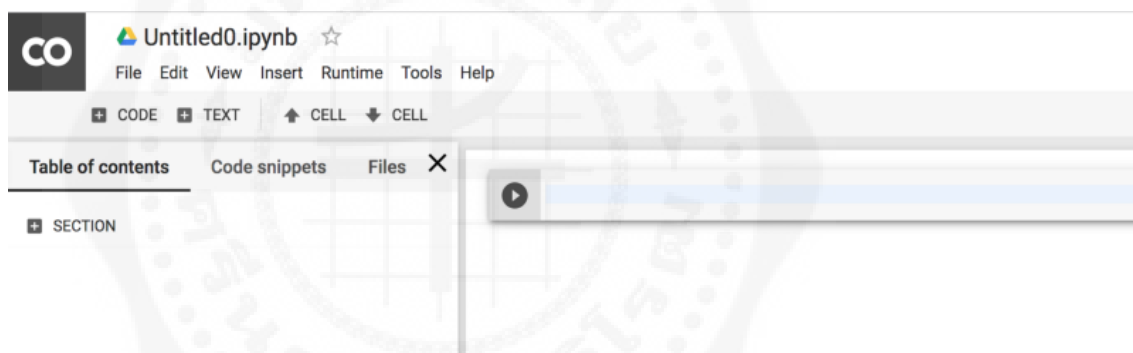
สอนการทำงานบน Google Colaboratory

วิธีใช้เบื้องต้นของ Google Colaboratory

1. การสร้างไฟล์ใหม่ ให้เข้าไปที่เว็บ <http://colab.research.google.com/> จากนั้นกดปุ่ม “NEW PYTHON 3 NOTEBOOK” เพื่อสร้าง Notebook ใหม่

NEW PYTHON 3 NOTEBOOK ▼

2. เริ่มแรกจะได้เป็นหน้าเปล่าๆ และให้ช่องสำหรับรันสคริปต์เพียงช่องเดียว โดยสามารถที่จะใส่คำสั่ง Bash ร่วมกันกับโค้ดภาษา Python ได้เลย เพียงแค่ใส่เครื่องหมาย “!” เพื่อแยกระหว่างคำสั่ง Bash กับคำสั่งภาษา Python



3. สามารถแบ่งโค้ดออกเป็นส่วนๆ โดยกดปุ่ม “+ Section” ด้านซ้ายมือสุด เมื่อเรากดย่อ โค้ดที่ต่อจาก Section นั้นก็จะถูกย่อตาม



4. การเพิ่มช่องสำหรับใส่สคริปต์ให้กดปุ่ม “CODE”

+ CODE

5. หากต้องการใส่ข้อความอธิบายใน Notebook ของเรากดปุ่ม “TEXT” และ

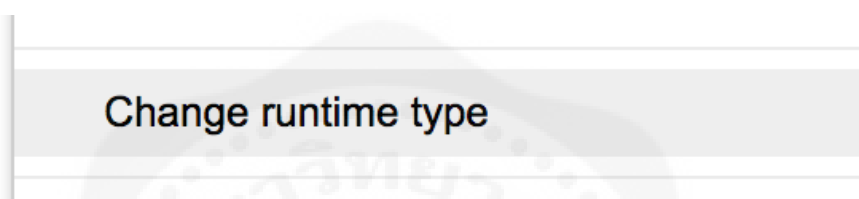


6. หากต้องการย้ายตำแหน่งวัตถุต่าง ๆ บนหน้าจอให้เลือก Cell นั้น ๆ ก่อนแล้วค่อยกดปุ่มดังรูป

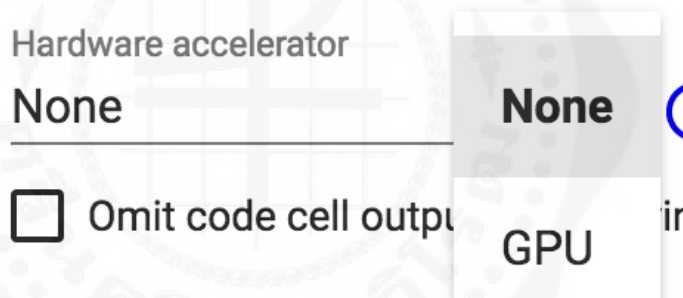


การเลือกใช้ GPU หรือ CPU และเวอร์ชัน Python

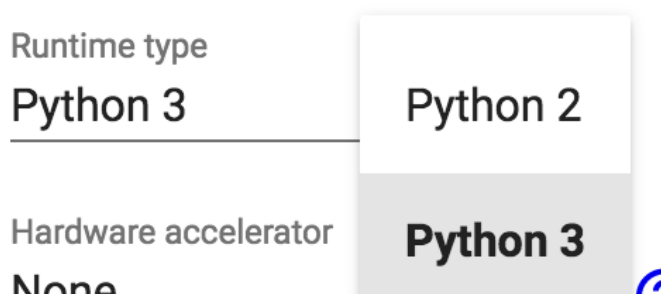
1. เลือกเมนู “Runtime” ตรงแถบเมนูด้านบนซ้ายบน เลือก “Change runtime type”



2. กดเลือก “GPU” เพื่อเปิดใช้งาน



3. เลือกเวอร์ชัน Python ที่จะให้ Compile โดย Runtime Type มีให้เลือกอยู่ 2 เวอร์ชัน คือ Python 2.7.14 และ Python 3.6.3



การใช้ Google Colaboratory ร่วมกับ Google Drive

1. ให้พิมพ์คำสั่งดังต่อไปนี้

```
from google.colab import drive
drive.mount('/content/drive')
```

2. ระหว่างรันสคริปต์จะปรากฏช่องว่างให้ใส่ Authen code ที่ได้จากการกดลิงก์ด้านบน เพื่อเชื่อมต่อกับ Google Drive ของเรา

Go to this URL in a browser: <https://accounts.google.com/>

Enter your authorization code:

3. หากต้องการเปลี่ยน root path ของไดเรกทอรี เพื่อลดความยาวของการเรียก Path ไฟล์ โดยจะได้ไม่ต้องเขียน Path ที่ยาว ให้ใช้คำสั่งดังต่อไปนี้

```
import os
os.chdir("drive/My Drive/(ชื่อโฟลเดอร์ใน Google Drive ที่ต้องการ)")
```

4. เมื่อเปลี่ยน Path เข้ามาในโฟลเดอร์ของไดร์ฟเราแล้ว เวลาแก้ไขไฟล์เช่น ลบหรือสร้างใหม่ มันก็จะอัปเดตไปยัง Google Drive ให้ทันที