

ระบบแนะนำประกัน
INSURANCE RECOMMENDER SYSTEM



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2562

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบแนะนำประกัน

นางสาวชนัญญา ชูศักดิ์ 59010272
นางสาวชุตিকা วีระยุทธวัฒน์ 59010337
ผศ.ดร. รัฐชัย ชาวอุทัย อาจารย์ที่ปรึกษา
ปีการศึกษา 2562

บทคัดย่อ

ระบบแนะนำประกันหรือ Insurex เป็นเว็บแอปพลิเคชันที่ถูกออกแบบและพัฒนาสำหรับการแนะนำแผนประกันที่เหมาะสมกับผู้ใช้โดยอาศัยการประเมินรูปแบบของคุณลักษณะของประกันและผู้ใช้จะซื้อประกันเข้าด้วยกัน ซึ่งกระบวนการส่วนใหญ่จะใช้เทคนิคต่างๆจาก Machine Learning มาคำนวณ วิเคราะห์ และทำนายหาคุณลักษณะต่างๆของประกันแต่ละแผนกับข้อมูลที่ใช้ได้กรอกเข้ามาในระบบ ก่อนจะนำมาหาความใกล้เคียงกับแผนประกันต่างๆ จากนั้นระบบก็จะสร้างรายการแผนประกันที่แนะนำตามความใกล้เคียงกับข้อมูลผู้ใช้งานหนึ่ง พร้อมทั้งแสดงผลผ่านเว็บแอปพลิเคชัน รวมถึงยังมีการคำนวณเบี้ยประกัน ทุนประกัน และรายละเอียดของประกันแนะนำในแต่ละแผนอีกด้วย โดยระบบรองรับการแนะนำประกันทั้งแบบสัญญาหลักและสัญญาเพิ่มเติม เพื่อให้ผู้ใช้ได้ประโยชน์จากการค้นหาแผนประกันที่เหมาะสมได้

Insurance Recommender System

Ms. Chananya Choosak 59010272

Ms. Chutika Weerayuthwattana 59010337

Dr. Rathachai Chawuthai Advisor

Academic Year 2019

Abstract

Finding a suitable insurance policy could be difficult due to the variety of insurance plans and biased agents. In this paper, we demonstrated how we tackled this problem by developing a web application that could recommend insurance plans based on the input data. The recommender system used various models to output characteristics of a life insurance policy. Then, it calculated a similarity between a policy with those attributes and the current policy. Finally, it recommended 5 most similar insurance plans.

Insurance recommender system (Insurex) is a web-based application which is designed and developed for recommending insurance plans for users. The system estimates the patterns similarity between insurance packages' properties and user's preferences. The recommender mechanism mainly employs several techniques from machine learning in order to calculate, analyze, and predict the properties of insurance packages by matching with input user preferences. After finishing the calculation, the system provides a list of recommend insurance plans to users via the web application. It also displays the premium, sum assured, and the detail of each insurance plan. In addition, the system supports both main plan and rider plan in order to recommend appropriate insurance plans to users.

กิตติกรรมประกาศ

ปริญญานิพนธ์ในหัวข้อระบบแนะนำประกันฉบับนี้ สามารถลุล่วงไปได้ด้วยดี เพราะได้รับการสนับสนุนและช่วยเหลือจากบุคคลหลายฝ่าย ซึ่งจะไม่สามารถสำเร็จลุล่วงไปได้ด้วยดี หากปราศจากบุคคลเหล่านี้ อันได้แก่

ผศ.ดร. รัฐชัย ชาวอุทัย อาจารย์ที่ปรึกษาโครงงาน ผู้ให้ความรู้และแนวทางในการศึกษา อันเป็นแนวทางชี้แนะ ช่วยเหลือ และชี้แนะให้ปริญญานิพนธ์มีความสมบูรณ์ คณะผู้จัดทำ ขอขอบพระคุณในความอนุเคราะห์เป็นอย่างสูง

ขอขอบคุณ คุณภักดิ์สินี วัฒนสรวิทย์พงศ์ (ศิษย์เก่าคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง รุ่นที่ 31) และคุณนวลแพร แก้วนุช ผู้ให้ความช่วยเหลือและความรู้ด้านประกันอันเป็นพื้นฐานในการทำโครงงานหัวข้อดังกล่าว

ขอขอบคุณ คุณเพชร มุสิกอินทร์และคุณสุจิรา เสถียรวงศา ผู้ให้ความรู้ด้านประกันเพิ่มเติม และให้ความอนุเคราะห์ในการเก็บข้อมูลต่างๆ อันได้แก่ ข้อมูลแผนประกันและผู้ทำประกันเพื่อการวิจัยและพัฒนาโครงงาน

ขอขอบพระคุณคณาจารย์ภาควิชาวิศวกรรมคอมพิวเตอร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ผู้คอยชี้แนะและให้ความรู้ตลอดช่วงการศึกษาอันเป็นส่วนสำคัญในการทำปริญญานิพนธ์นี้ให้สำเร็จลุล่วงได้ด้วยดี

สุดท้ายนี้ ขอขอบพระคุณบิดามารดาและครอบครัวที่อบรมเลี้ยงดู ให้ความช่วยเหลือ และสนับสนุนจนทำให้ปริญญานิพนธ์ฉบับนี้เสร็จสมบูรณ์

ชัญญญา ชูศักดิ์

ชุดิกา วีระยุทธวัฒน์

สารบัญ

	หน้า
ระบบแนะนำประกัน	I
Insurance Recommender System	II
กิตติกรรมประกาศ.....	III
สารบัญ	IV
สารบัญตาราง	IX
สารบัญรูปภาพ	X
บทที่ 1 บทนำ	1
1.1. ความเป็นมาของปัญหา.....	1
1.2. วัตถุประสงค์ของโครงการ	2
1.3. ขอบเขตของโครงการ	2
1.4. ประโยชน์ที่คาดว่าจะได้รับ	2
1.5. ส่วนประกอบของปริญญาานิพนธ์	2
1.6. การดำเนินการโครงการ	3
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	6
2.1. ประกันชีวิต	6
2.1.1. สัญญาหลัก	6
2.1.2. สัญญาเพิ่มเติม	7
2.2. เว็บไซต์แนะนำประกัน	8
2.2.1. iTax Market	8
2.2.2. Roojai.com.....	9
2.2.3. Easy Compare.....	10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.3. วิทยาศาสตร์ข้อมูล (Data Science)	11
2.4. การเรียนรู้ของเครื่อง (Machine Learning).....	12
2.4.1. Supervised Learning	12
2.4.2. Unsupervised Learning	15
2.4.3. Reinforcement Learning	16
2.5. ระบบแนะนำ (Recommender System)	17
2.5.1. Content-Based Filtering	17
2.5.2. Collaborative Filtering	17
2.6. การเก็บรวบรวมข้อมูล (Data Collection).....	19
2.7. การทำความสะอาดข้อมูล (Data Cleansing).....	19
2.8. การเตรียมพร้อมสำหรับข้อมูล (Data Preprocessing).....	19
2.8.1. Standardizing Data.....	20
2.8.2. Feature Engineering.....	20
2.8.3. Feature Selection	20
2.9. การวิเคราะห์ข้อมูล (Data Analysis).....	21
2.10. เทคนิคการสังเคราะห์เพิ่มตัวอย่างข้างน้อย (Synthetic Minority Over-Sampling Technique)	21
2.11. เทคโนโลยีที่ใช้.....	22
2.11.1. Python.....	22
2.11.2. Anaconda.....	22
2.11.3. Numpy และ Pandas	23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.11.4. Matplotlib และ Seaborn	23
2.11.5. scikit-learn	24
2.11.6. Jupyter	24
2.11.7. React.....	25
2.11.8. Flask.....	25
บทที่ 3 การออกแบบและพัฒนาระบบ	26
3.1. สถาปัตยกรรมของระบบ	26
3.2. ภาพรวมการทำงานและการออกแบบอัลกอริทึม	27
3.2.1. ข้อมูลนำเข้าและข้อมูลส่งออก.....	27
3.2.2. ภาพรวมการทดลอง.....	28
3.2.3. หลักการทำงานของอัลกอริทึม	29
3.2.4. การวัดผลอัลกอริทึม	33
3.2.5. Use Case Diagram	33
3.3. Sequential Diagram.....	34
3.3.1. Sequence Diagram สำหรับการกรอกข้อมูลส่วนตัว.....	34
3.3.2. Sequence Diagram สำหรับการสร้างรายแผนประกันที่แนะนำ	35
3.3.3. Sequence Diagram สำหรับการคำนวณทุนประกันและเบี้ยประกัน	366
3.3.4. Sequence Diagram สำหรับการดูรายแผนประกันที่แนะนำ	38
3.4. Relational Database Diagram	39
3.5. User Interface Design	400
3.5.1. UI สำหรับหน้าหลัก	400

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
3.5.2. UI สำหรับการกรอกข้อมูลส่วนที่ 1.....	411
3.5.3. UI สำหรับการกรอกข้อมูลส่วนที่ 2.....	421
3.5.4. UI สำหรับหน้าแนะนำรายการแผนประกัน	42
บทที่ 4 การทดลองและผลการทดลอง.....	44
4.1. บทนำการทดลอง	44
4.2. ผลการทำความสะอาดข้อมูลและวิเคราะห์ข้อมูลเบื้องต้น.....	44
4.3. ผลการทดลองอัลกอริทึมกับข้อมูลที่จะใช้จริง.....	48
4.3.1. ผลการทดลองอัลกอริทึมที่ใช้ในการทำนายคุณลักษณะของประกันตลอดชีพ.....	48
4.3.2. ผลการทดลองอัลกอริทึมที่ใช้ในการทำนายคุณลักษณะของประกันสะสม ทรัพย์สิน.....	50
4.3.3. ผลการทดลองอัลกอริทึมที่ใช้ในการทำนายการซื้อประกันบ้าน.....	52
4.3.4. ผลการทดลองอัลกอริทึมที่ใช้ในการทำนายการซื้อประกันสัญญาเพิ่มเติม.....	53
4.3. ผลการทดสอบโปรแกรมแนะนำประกัน	55
4.4. ผลการพัฒนาเว็บแอปพลิเคชัน.....	56
4.5.1. การกรอกข้อมูลเบื้องต้นและเป้าหมายการทำประกันเพื่อสร้างรายการแนะนำ ประกัน	56
4.5.2. การดูรายการแผนประกันที่แนะนำ.....	58
4.5. ผลการเปรียบเทียบกับแอปพลิเคชันแนะนำประกันอื่นๆ	59
บทที่ 5 บทสรุปและข้อเสนอแนะ.....	61
5.1. บทสรุป.....	61
5.2. ปัญหา อุปสรรค และแนวทางการแก้ไข.....	62

สารบัญ (ต่อ)

	หน้า
5.3. แนวทางการพัฒนาต่อ	62
บรรณานุกรม	63
ภาคผนวก ก.....	65



สารบัญตาราง

ตาราง	หน้า
1.1 แผนการดำเนินการ โครงการงาน (ภาคเรียนที่ 1).....	3
1.1 แผนการดำเนินการ โครงการงาน (ภาคเรียนที่ 2).....	5
3.1 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการกรอกข้อมูลส่วนตัว.....	34
3.2 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการสร้างรายแผนประกันที่แนะนำ.....	36
3.3 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการคำนวณทุนประกันและเบี้ยประกัน.....	37
3.4 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการดูรายแผนประกันที่แนะนำ.....	38
4.1 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับระยะเวลาชำระเบี้ย (ประกันตลอดชีพ).....	48
4.2 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับเงินปันผล (ประกันตลอดชีพ).....	49
4.3 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับระยะเวลาคุ้มครอง (ประกันตลอดชีพ).....	49
4.4 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับระยะเวลาชำระเบี้ย (ประกันสะสมทรัพย์).....	50
4.5 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับระยะเวลาคุ้มครอง (ประกันสะสมทรัพย์).....	50
4.6 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับเงินคืนทั้งหมด (ประกันสะสมทรัพย์).....	51
4.7 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับความคุ้มครองเมื่อจบสัญญา (ประกันสะสมทรัพย์).....	51
4.8 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับจำนวนครั้งที่ได้เงินคืน (ประกันสะสมทรัพย์).....	52
4.9 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง AHC.....	53
4.10 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง ECIR.....	53
4.11 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองอุบัติเหตุ AI/RCC.....	54
4.12 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยรายวัน HB.....	54
4.13 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HS.....	55
4.14 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HSPG.....	55
4.15 ตารางแสดงผลการทดสอบโปรแกรมด้วย Top-K Precision.....	56
4.16 ตารางแสดงการเปรียบเทียบเว็บแอปพลิเคชัน Insurex กับเว็บแอปพลิเคชันอื่น.....	60

สารบัญรูปภาพ

รูป	หน้า
2.1 เว็บไซต์ iTax Market.....	9
2.2 หน้าแนะนำประกันชีวิตในเว็บไซต์ iTax Market.....	9
2.3 เว็บไซต์ Roojai.com.....	10
2.4 หน้าเสนอแผนประกันภัยรถยนต์ในเว็บไซต์ Roojai.com.....	10
2.5 เว็บไซต์ Easy Compare.....	11
2.6 หน้าแนะนำประกันในเว็บไซต์ Easy Compare.....	11
2.7 ศาสตร์ที่เกี่ยวข้องกับวิทยาศาสตร์ข้อมูล.....	12
2.8 ตัวอย่างแบบจำลอง Decision Tree.....	13
2.9 หลักการทำ Random Forest.....	13
2.10 ตาราง Confusion Matrix.....	14
2.11 ภาพประกอบเพื่ออธิบายการคำนวณค่า Precision และ Recall.....	15
2.12 ประเภทของการเรียนรู้ของเครื่อง.....	16
2.13 กระบวนการทำงานของการเรียนรู้ของเครื่อง.....	16
2.14 หลักการทำงานของระบบแนะนำแบบคัดกรองอิงเนื้อหา.....	17
2.15 หลักการทำงานของระบบแนะนำแบบคัดกรองร่วม.....	18
2.16 Matrix Factorization.....	19
2.17 กราฟแสดงความสัมพันธ์ของคูฟิเจอร์.....	20
2.18 การลดมิติของข้อมูล.....	21
2.19 การสังเคราะห์ข้อมูลใหม่ด้วย SMOTE.....	21
2.20 โลโก้ของ Python.....	22
2.21 โลโก้ของ Anaconda.....	22
2.22 โลโก้ของ Numpy.....	23
2.23 โลโก้ของ Pandas.....	23
2.24 โลโก้ของ Matplotlib.....	23
2.25 โลโก้ของ Seaborn.....	23
2.26 โลโก้ของ scikit-learn.....	24

สารบัญรูปภาพ (ต่อ)

รูป	หน้า
2.27 โลโก้ของ Jupyter	24
2.28 โลโก้ของ React.....	25
2.29 โลโก้ของ Flask	25
3.1 สถาปัตยกรรมของระบบแนะนำประกัน	26
3.2 ตัวอย่างข้อมูลทดลองที่ผ่านการทำความสะอาดแล้ว.....	28
3.3 ภาพรวมการทำงานของอัลกอริทึม	30
3.4 แผนภาพ Flowcharts แสดงการทำงานของระบบแนะนำประกัน.....	31
3.5 แผนภาพ Use Case สำหรับระบบแนะนำประกัน	33
3.6 Sequence Diagram สำหรับการกรอกข้อมูลส่วนตัว.....	34
3.7 Sequence Diagram สำหรับการสร้างรายแผนประกันที่แนะนำ.....	35
3.8 Sequence Diagram สำหรับการคำนวณทุนประกันและเบี้ยประกัน	37
3.9 Sequence Diagram สำหรับการดูรายแผนประกันที่แนะนำ.....	38
3.10 การออกแบบฐานข้อมูลของระบบ.....	40
3.11 หน้า UI สำหรับหน้าหลัก	41
3.12 หน้า UI สำหรับการกรอกข้อมูลส่วนที่ 1	42
3.13 หน้า UI สำหรับการกรอกข้อมูลส่วนที่ 2	43
3.14 หน้า UI สำหรับหน้าแนะนำรายแผนประกัน	43
4.1 กราฟฮิสโทแกรมแสดงเพศและสถานะของผู้เอาประกัน	45
4.2 กราฟฮิสโทแกรมแสดงชั้นอาชีพและประเภทประกันที่ผู้เอาประกันเลือกทำ	45
4.3 กราฟแสดงจำนวนผู้ซื้อประกันตลอดชีพ สะสมทรัพย์ บ้านอายุ	46
4.4 กราฟแสดงช่วงอายุของผู้เอาประกัน	47
4.5 กราฟแสดงช่วงเงินทุนประกันและเบี้ยประกันที่ผู้เอาประกันต้องชำระ	47
4.6 หน้าหลักของเว็บ Insurex	56
4.7 หน้ากรอกข้อมูลส่วนตัว (ส่วนข้อมูลเบื้องต้น).....	57
4.8 หน้ากรอกข้อมูลส่วนตัว (ส่วนเป้าหมายการทำประกัน)	57
4.9 กดปุ่มแนะนำประกันเพื่อเริ่มการแนะนำประกัน.....	58
4.10 หน้าแสดงผลการแนะนำประกัน	59

สารบัญรูปลูกภาพ (ต่อ)

รูป	หน้า
ก.1 User Info Table	65
ก.2 Recommended Plan Table	68
ก.3 Protection Table.....	69
ก.4 Endowment Table	69
ก.5 Annuity Table	70
ก.6 ProPackage Table	71
ก.7 Rider Table	71
ก.8 Cashback for En Table.....	72
ก.9 Critical Illness Table.....	72
ก.10 JobRate Table	73

บทที่ 1

บทนำ

ในบทนี้จะกล่าวถึงเหตุผลที่จัดทำโครงการ ตั้งแต่ความเป็นมาของปัญหา วัตถุประสงค์และขอบเขตของโครงการ ประโยชน์ที่คาดว่าจะได้รับจากโครงการ และแผนการดำเนินงานโครงการ

1.1. ความเป็นมาของปัญหา

ในปัจจุบันนี้ ชีวิตของคนมีความเสี่ยง ซึ่งอาจจะเกิดเหตุการณ์ที่ไม่คาดคิดที่อาจส่งผลกระทบต่อการใช้ชีวิตได้ โดยเฉพาะอย่างยิ่งแล้ว บุคคลที่มีครอบครัว เป็นเสาหลักในการหารายได้ หากเกิดเหตุที่ทำให้ไม่สามารถหารายได้ได้อย่างปกติ ก็ส่งผลกระทบต่อการดำเนินชีวิตของครอบครัวอย่างแน่นอน เพราะครอบครัวนั้นย่อมมีภาระมีค่าใช้จ่ายต่างๆ หากขาดเสาหลักไปก็ประสบปัญหานั้นได้ หากต้องการสร้างประกันและความมั่นคงให้แก่ตนเองหรือคนในครอบครัวยามเกิดเหตุไม่คาดฝันนั้น ประกันชีวิตถือเป็นทางเลือกที่ดีที่จะสร้างหลักประกันความมั่นคงแก่ตนเองหรือครอบครัว เพราะผลประโยชน์ที่ได้จากการทำประกันก็จะสามารถบรรเทาความเสียหายที่เกิดขึ้นจากเหตุไม่คาดฝันได้ อีกทั้งการทำประกันยังช่วยสร้างวินัยทางการเงิน ได้ผลตอบแทนที่มั่นคง และยังลดหย่อนภาษีได้อีกด้วย

แต่ทว่า ถึงแม้ประกันชีวิตจะมีประโยชน์อย่างมากในการสร้างหลักประกันความมั่นคงต่อชีวิต ก็ยังไม่เป็นที่นิยมที่คนในสังคมส่วนใหญ่ เพราะการทำประกันนั้นต้องทำสัญญาผ่านตัวแทนประกันชีวิต ซึ่งเกิดปัญหาต่างๆ เช่น ตัวแทนประกันชีวิตขายประกันโดยไม่คำนึงถึงผู้เอาประกัน ทำให้คุณภาพลักษณะของการทำประกันในทางที่ไม่ดี หรือปัญหาที่เกิดจากทางผู้เอาประกันเอง เช่น ผู้เอาประกันไม่ยอมบอกข้อมูลส่วนตัวที่เป็นจริงแก่ตัวแทนประกัน ทำให้ได้ประกันที่ไม่เหมาะสมกับตนหรือเกิดข้อผิดพลาดเมื่อรับผลประโยชน์ เป็นต้น

ระบบแนะนำประกันจึงถูกพัฒนาขึ้นมาเพื่อแก้ไขปัญหาเหล่านี้ โดยระบบจะรับข้อมูลจากผู้ใช้เพื่อนำมาวิเคราะห์หาแผนประกันที่เหมาะสมแก่ผู้ใช่มากที่สุดและสามารถเลือกแผนประกันที่ต้องการได้ด้วยตนเอง โดยไม่ต้องผ่านตัวแทนประกันชีวิต โดยคณะผู้วิจัยวางแผนจัดทำระบบแนะนำโดยสร้างแบบจำลองจากการเรียนรู้ของเครื่อง (Machine Learning) และจะพัฒนาเว็บไซต์เพื่อใช้ในการแสดงผลรายแผนประกันที่ระบบได้แนะนำให้ผู้ใช้งานได้ง่าย สามารถเลือกแบบประกันได้ตรงความต้องการของตนได้สะดวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2. วัตถุประสงค์ของโครงการ

- 1) เพื่อสร้างแบบจำลองของระบบแนะนำประกัน
- 2) เพื่อพัฒนาเว็บแอปพลิเคชันช่วยแนะนำประกันที่เหมาะสมและตรงความต้องการของผู้ใช้

1.3. ขอบเขตของโครงการ

- 1) ใช้ข้อมูลของผู้เอาประกันจากบริษัทประกันหนึ่งแห่ง
- 2) ใช้ข้อมูลแผนประกันจากบริษัทประกันหนึ่งแห่ง โดยมีประเภทสัญญาต่อไปนี้
สัญญาหลัก ได้แก่ ประกันชีวิต ประกันสะสมทรัพย์ ประกันบำนาญ และประกันคุ้มครอง
ชั่วระยะเวลา
สัญญาเพิ่มเติม ได้แก่ ประกันคุ้มครองโรคร้ายแรง ประกันชดเชยรายวัน ประกันชดเชยค่า
รักษาพยาบาล และประกันอุบัติเหตุ

1.4. ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้แบบจำลองที่สามารถแนะนำประกันได้เหมาะสมและตรงความต้องการของผู้ใช้
- 2) ได้เว็บแอปพลิเคชันที่ช่วยแนะนำประกันให้แก่ผู้ใช้ โดยที่ผู้ใช้สามารถกรอกข้อมูล
ส่วนตัวได้โดยไม่ต้องผ่านการทำประกันผ่านตัวแทนประกันชีวิต
- 3) คณะผู้จัดทำมีความรู้ความเข้าใจด้านประกัน
- 4) คณะผู้จัดทำมีความรู้ความเข้าใจเรื่องการเรียนรู้ของเครื่อง (Machine Learning)
- 5) คณะผู้จัดทำมีความรู้ความเข้าใจในทักษะการนำข้อมูลไปใช้สร้างแบบจำลองเพื่อต่อยอด
ใช้งานได้จริง

1.5. ส่วนประกอบของปฏิญานิพนธ์

- 1) บทที่ 1 บทนำ กล่าวถึงความเป็นมาของปัญหา วัตถุประสงค์ของโครงการ ขอบเขตของ
โครงการ ประโยชน์ที่คาดว่าจะได้รับ ส่วนประกอบของปฏิญานิพนธ์ และการ
ดำเนินการโครงการ(ส่วนภาคเรียนที่ 1)
- 2) บทที่ 2 ทฤษฎีที่เกี่ยวข้อง กล่าวถึงประกันชีวิต การทำวิทยาศาสตร์ข้อมูล การเรียนรู้ของ
เครื่อง และเทคโนโลยีที่ใช้
- 3) บทที่ 3 การออกแบบและพัฒนาระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6. การดำเนินการโครงการ

- 1) ศึกษาความรู้ด้านประกัน การเรียนรู้ของเครื่อง (Machine Learning) รวมไปถึงรวบรวมข้อมูลของแผนประกันจากเว็บไซต์บริษัทประกัน
- 2) ติดต่อกับบริษัทประกันเพื่อขอรับข้อมูลของผู้เอาประกัน
- 3) เก็บข้อมูลของผู้เอาประกัน
- 4) นำข้อมูลทั้งผู้เอาประกันภัยและแผนประกันมาทำความสะอาดข้อมูล (Data Cleaning) และเตรียมการสำหรับสร้างแบบจำลอง (Data Preprocessing)
- 5) ทดลองสอนแบบจำลองโดยใช้อัลกอริทึมต่างๆ (Data Training)
- 6) ออกแบบส่วนติดต่อกับผู้ใช้ (User Interface) ของหน้าเว็บไซต์เพื่อแสดงผลแบบประกัน
- 7) สร้างแบบจำลองแนะนำประกันภัยโดยใช้ข้อมูลผู้เอาประกันและข้อมูลประกัน

ตาราง 1.1 แผนการดำเนินการโครงการ (ภาคเรียนที่ 1)

ลำดับ	งานที่ได้ลงมือทำ	สิงหาคม				กันยายน				ตุลาคม				พฤศจิกายน				
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
1	ศึกษาความรู้ด้านประกัน																	
2	ศึกษาความรู้เกี่ยวกับการเรียนรู้ของเครื่องและการทำระบบแนะนำ																	
3	รวบรวมข้อมูลของแผนประกันจากเว็บไซต์บริษัทประกัน																	
4	ติดต่อกับบริษัทประกันเพื่อขอรับข้อมูลของผู้เอาประกันเพื่อขอรับข้อมูลผู้เอาประกัน																	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับ	งานที่ได้ลงมือทำ	สิงหาคม				กันยายน				ตุลาคม				พฤศจิกายน			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
5	เก็บข้อมูลของผู้เอาประกัน																
6	นำข้อมูลผู้เอาประกันและแผนประกันมาทำความสะอาดข้อมูล (Data Cleaning)																
7	นำข้อมูลมาเตรียมการสำหรับสร้างแบบจำลอง (Data Preprocessing)																
8	ทดลองสอนแบบจำลองโดยใช้อัลกอริทึมต่างๆ (Data Training)																
9	ออกแบบส่วนติดต่อกับผู้ใช้ (User Interface) ของหน้าเว็บไซต์เพื่อแสดงผลแบบประกัน																
10	สร้างแบบจำลองแนะนำประกันภัยโดยใช้ข้อมูลผู้เอาประกันและข้อมูลประกัน																

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 1.2 แผนการดำเนินการโครงการ (ภาคเรียนที่ 2)

ลำดับ	งานที่ได้ลงมือทำ	มกราคม				กุมภาพันธ์				มีนาคม			
		1	2	3	4	1	2	3	4	1	2	3	4
1	ปรับปรุงและพัฒนาแบบจำลองแนะนำประกันและข้อมูลประกัน												
2	เขียนเว็บสำหรับระบบแนะนำประกัน												
3	ทดสอบการทำงานของเว็บและแบบจำลองแนะนำประกัน												
4	ปรับปรุงการทำงานระหว่างเว็บและแบบจำลองแนะนำประกัน												

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องที่คณะผู้จัดทำได้ทำการศึกษาค้นคว้าเพื่อนำมาประกอบการทำโครงการ ดังหัวข้อย่อต่อไปนี้

2.1. ประกันชีวิต

ผลิตภัณฑ์ทางการเงินรูปแบบหนึ่งที่ทำให้ความคุ้มครองแก่ชีวิตเมื่อเกิดเหตุที่ส่งผลกระทบต่อการใช้ชีวิตและการเงินอย่างรุนแรง ได้แก่ เสียชีวิต ทูพลาภาพถาวร และการสูญเสียรายได้ยามชรา ซึ่งหากประสบเหตุเหล่านี้ก็จะได้รับเงินชดเชยตามเงื่อนไขของสัญญาที่ได้ตกลงกับบริษัทเพื่อบรรเทาความเสียหายแก่ตนเองและครอบครัว

ในนัยของกฎหมาย ประกันชีวิตจะมีชื่อเรียกว่า “กรมธรรม์ประกันชีวิต” คือสัญญาต่างตอบแทนของกลุ่มสัญญาระหว่าง “ผู้เอาประกัน” หรือผู้ที่ทำประกันและมีหน้าที่จ่าย “เบี้ยประกัน” ให้กับคู่สัญญาที่ตกลงกันคือ “ผู้รับทำประกัน” หรือบริษัทรับทำประกัน โดยบริษัทรับทำประกันจะต้องเป็นผู้ชดเชยค่าตอบแทนตามความคุ้มครองให้กับผู้เอาประกันหรือ “ผู้รับผลประโยชน์” เมื่อผู้เอาประกันเสียชีวิตหรือมีชีวิตอยู่จนครบสัญญาเป็นจำนวนเท่ากับ “ทุนประกัน” ที่ตกลงกัน ณ เริ่มทำสัญญา หรือหากผู้เอาประกันทุพพลภาพถาวร บริษัทรับทำประกันจะดำเนินการยกเว้นการชำระเบี้ย

ประกันชีวิตนอกจากมีประโยชน์ในด้านการคุ้มครองผู้เอาประกันและผู้รับผลประโยชน์แล้ว บางสัญญาผู้เอาประกันจะได้รับผลตอบแทน เช่น เงินคืนระหว่างสัญญา เงินปันผล เป็นต้น และสามารถนำไปลดหย่อนภาษีได้

ประกันชีวิตสามารถแบ่งตามประเภทสัญญาออกได้ 2 ประเภทใหญ่ ดังนี้

2.1.1. สัญญาหลัก

สัญญาที่ให้ความคุ้มครองแก่ผู้เอาประกัน เมื่อผู้เอาประกันเสียชีวิตหรือมีชีวิตอยู่จนครบสัญญา บริษัทรับประกันจะจ่ายเงินคุ้มครองหรือเงินเอาประกันและเงินตอบแทนเพิ่มเติม(ถ้ามี) ให้กับผู้เอาประกันหรือผู้รับผลประโยชน์ตามที่ผู้เอาประกันได้ระบุไว้เมื่อผู้เอาประกันเสียชีวิตก่อนครบสัญญา สัญญาหลักสามารถแบ่งได้ดังนี้

1) ประกันคุ้มครองตลอดชีพ

ประกันที่ผู้เอาประกันชำระเบี้ยเป็นระยะเวลาหนึ่ง ซึ่งจะได้รับความคุ้มครองระยะยาว (โดยทั่วไปจะหมดสัญญาเมื่อผู้เอาประกันมีอายุครบ 99 ปี) หากผู้เอาประกันมีชีวิตอยู่จนครบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาจะได้รับเงินทุนประกันคืน แต่ถ้าเสียชีวิตก่อนหมดสัญญา จะเป็นผู้รับผลประโยชน์ที่ได้รับเงินคุ้มครองส่วนนั้นแทน

ประกันคุ้มครองตลอดชีพเหมาะสำหรับผู้เอาประกันที่มีครอบครัว ต้องการคุ้มครองถึงคนข้างหลัง เพราะเมื่อตนเสียชีวิต คนข้างหลังจะได้รับเงินคุ้มครองแทนตน (ในฐานะผู้รับผลประโยชน์) ประกันคุ้มครองตลอดชีพบางสัญญาผู้เอาประกันจะได้รับเงินปันผลร่วมด้วย

2) ประกันคุ้มครองชั่วระยะเวลา

คล้ายประกันคุ้มครองตลอดชีพ แต่มีความแตกต่างคือประกันชนิดนี้จะคุ้มครองชั่วระยะเวลา ทำให้เบี้ยประกันมีราคาถูกกว่าชนิดอื่นๆ หากผู้เอาประกันเสียชีวิตก่อนครบกำหนดสัญญา ผู้รับผลประโยชน์จะได้รับเงินคุ้มครองตามจำนวนเงินทุนประกันที่ทำสัญญาไว้ แต่ถ้าผู้เอาประกันมีชีวิตอยู่จนครบสัญญา จะไม่ได้รับเงินทุนประกันคืน ประกันลักษณะนี้มีชื่อเรียกอีกชื่อว่า “เบี้ยจ่ายทิ้ง”

ประกันคุ้มครองชั่วระยะเวลาเหมาะสำหรับผู้เอาประกันที่ต้องการคุ้มครองชั่วระยะเวลาหนึ่ง และไม่ต้องการชำระเบี้ยในราคาสูง

3) ประกันสะสมทรัพย์

ประกันเน้นการออมเงินที่ผู้เอาประกันได้รับความคุ้มครองชั่วระยะเวลาหนึ่ง แต่ในช่วงระยะคุ้มครองจะได้รับเงินคืนระหว่างสัญญา หากผู้เอาประกันมีชีวิตอยู่จนครบสัญญา จะได้รับเงินทุนประกันคืนพร้อมเงินคืนตลอดสัญญา แต่ถ้าผู้เอาประกันเสียชีวิตก่อนครบสัญญา ผู้รับผลประโยชน์จะเป็นผู้รับแทน

เนื่องจากมีเงินคืนระหว่างสัญญา ประกันสะสมทรัพย์จึงมีอัตราชำระเบี้ยต่อทุนประกันสูงกว่าประกันชนิดอื่นๆ ทำให้ทุนประกันที่ทำได้มีไม่สูงหากเทียบกับประกันชนิดอื่นที่มีเบี้ยประกันเท่ากัน

4) ประกันบำนาญ

ประกันที่ผู้เอาประกันต้องทำและชำระเบี้ยก่อนถึงอายุเกษียณ เมื่อเริ่มเกษียณจะได้รับเงินตอบแทนเป็น “เงินบำนาญ” ทุกปีจนถึงอายุครบกำหนดตามสัญญา หากผู้เอาประกันเสียชีวิตในช่วงหลังเกษียณและก่อนอายุครบสัญญา ผู้รับผลประโยชน์จะได้รับเงินคุ้มครองเป็นมูลค่าของเงินบำนาญขาดเซยในส่วนที่ยังไม่ได้รับ

2.1.2. สัญญาเพิ่มเติม

สัญญาที่สามารถแนบเพิ่มเติมเข้ากับสัญญาหลักได้ โดยต้องจ่ายเบี้ยประกันแยกตามทีแต่ละสัญญากำหนดไว้และมีทุนประกันเป็นของตัวเองต่างหาก ซึ่งผู้เอาประกันจะไม่สามารถซื้อสัญญาเพิ่มเติมได้หากไม่ซื้อสัญญาหลัก โดยสัญญาเพิ่มเติมจะมีระยะเวลาคุ้มครองเมื่อสัญญาหลักมีผลบังคับอยู่นั้น สัญญาเพิ่มเติมสามารถแบ่งได้ดังนี้

เอกสารนี้เป็นเอกสารส่งเสริมการขายเพื่อประกอบการพิจารณา ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) ประกันอุบัติเหตุ

ประกันที่ให้ความคุ้มครองแก่ผู้เอาประกันกรณีประสบอุบัติเหตุได้รับบาดเจ็บทางกาย เสียชีวิต หรือทุพพลภาพ บริษัททำประกันจะชำระเงินคุ้มครองให้แก่ผู้เอาประกันหรือผู้รับผลประโยชน์ (กรณีผู้เอาประกันภัยเสียชีวิต)

2) ประกันชดเชยรายวัน

ประกันที่ให้ความคุ้มครองแก่ผู้เอาประกันเมื่อต้องเข้ารับการรักษาในโรงพยาบาล โดยบริษัทประกันจะชำระเงินชดเชยรายวันเป็นค่ารักษาพยาบาลและค่าห้องพักกรณีต้องเข้าพักในโรงพยาบาลระยะเวลาหนึ่ง

3) ประกันคุ้มครองโรคร้ายแรง

ประกันที่ให้ความคุ้มครองแก่ผู้เอาประกันเมื่อเจ็บป่วยด้วยโรคร้ายแรงตามเงื่อนไขที่กำหนดไว้ในกรมธรรม์ บริษัทประกันจะชำระเงินคุ้มครองในอัตราส่วนน้อยกว่าหรือเท่ากับเงินทุนประกัน ประกันคุ้มครองโรคร้ายแรงแบ่งเป็น คุ้มครองโรคร้ายแรงระดับต้นถึงปานกลาง และคุ้มครองโรคร้ายแรงระดับรุนแรง

4) ผลประโยชน์ยกเว้นเบี้ยประกันกรณีทุพพลภาพ

ผู้เอาประกันจะได้รับผลประโยชน์นี้เมื่อทุพพลภาพสิ้นเชิงถาวร ไม่สามารถช่วยเหลือตนเองได้ ผู้เอาประกันจะยังได้รับความคุ้มครองจากสัญญาหลัก โดยที่ไม่ต้องชำระเบี้ยประกัน สัญญาหลักตลอดจนอายุครบสัญญา โดยปกติประกันชนิดนี้จะแนบกับสัญญาหลักโดยไม่ต้องเสียค่าชำระเบี้ยประกันเพิ่มเติม

5) ผลประโยชน์ผู้ชำระเบี้ย

ประกันชนิดนี้จะแนบเพิ่มในสัญญาหลักเมื่อผู้เอาประกันเป็นผู้เยาว์ ต้องให้ผู้ปกครองเป็นผู้ชำระเบี้ย โดยจะให้ความคุ้มครองเมื่อผู้เอาประกันเสียชีวิตหรือทุพพลภาพและผู้ชำระเบี้ยไม่ต้องชำระเบี้ยสัญญาหลักและเพิ่มเติมตลอดจนครบสัญญา หรืออีกกรณีหากผู้ชำระเบี้ยประสบเหตุจนไม่สามารถชำระเบี้ยได้ ผู้เอาประกันจะได้รับเงินคุ้มครองเป็นจำนวนเท่ากับเงินทุนประกัน

2.2. เว็บไซต์แนะนำประกัน

ในหัวข้อนี้จะยกตัวอย่างเว็บไซต์ที่แนะนำประกัน ดังหัวข้อย่อต่อไปนี้

2.2.1. iTax Market

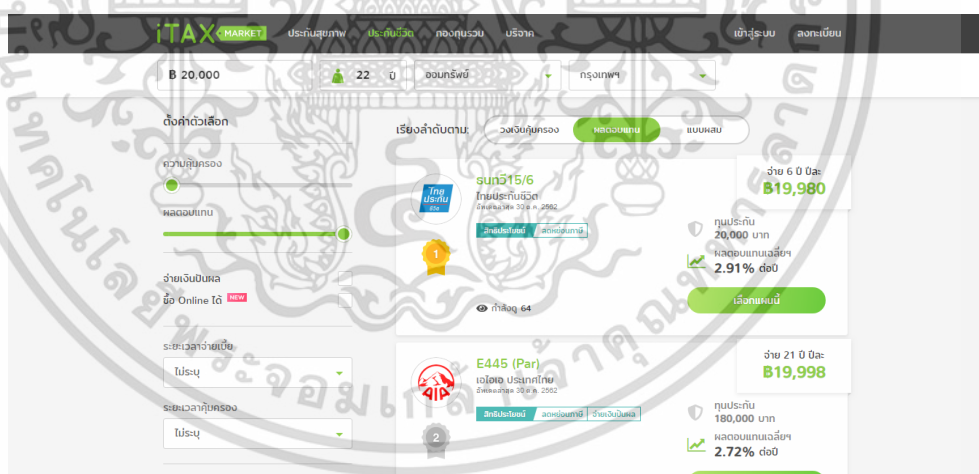
เว็บไซต์ iTax Market เป็นเว็บไซต์แนะนำประกันและกองทุนที่เน้นทำการลดหย่อนภาษีเป็นหลัก สามารถเลือกได้ 3 ชนิดคือ ประกันชีวิต ประกันสุขภาพ และกองทุน ซึ่งใช้งานได้ง่าย เพียงเลือกประเภทประกัน กรอกข้อมูลเพศ อายุ และงบสำหรับชำระเบี้ยประกันชีวิต ดังรูปที่ 2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อคลิกดูประกันตัว Top เว็บไซต์จะแสดงประกันที่ลดหย่อนภาษีและให้ผลประโยชน์แก่ผู้ต้องเสียภาษีมากที่สุด ทั้งยังสามารถเรียงลำดับได้ตามวงเงินคุ้มครอง ผลตอบแทน หรือแบบผสมได้ ดังรูปที่ 2.2



รูปที่ 2.1 เว็บไซต์ iTax Market
(ที่มา <https://www.itax.in.th/market>)



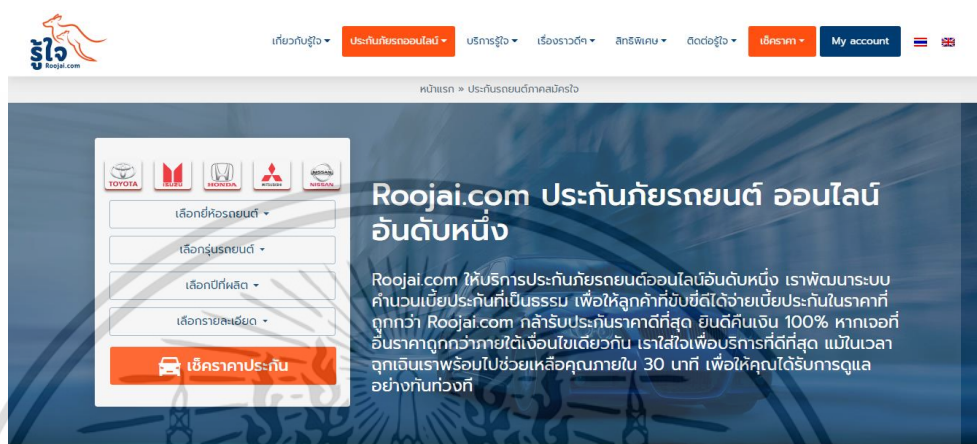
รูปที่ 2.2 หน้าแนะนำประกันชีวิตในเว็บไซต์ iTax Market
(ที่มา <https://www.itax.in.th/market>)

2.2.2. Roojai.com

เว็บไซต์แนะนำประกันภัยสำหรับรถยนต์ โดยจะแนะนำประกันที่ราคาถูกลงและมีความคุ้มค่าต่อผู้ซื้อรถยนต์มากที่สุด ผู้ใช้จะต้องกรอกข้อมูลรถยนต์ ข้อมูลส่วนตัว และพฤติกรรมการใช้รถเบื้องต้น ดังรูปที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อกรอกข้อมูลเรียบร้อยแล้ว ระบบจะคำนวณราคาเบี้ยประกันพร้อมเสนอแผนประกันที่เหมาะสมกับผู้ใช้ตามข้อมูลที่กรอกไปข้างต้น ทั้งยังสามารถปรับเปลี่ยนแผนประกันได้ตามความต้องการ ดังรูปที่ 2.4



รูปที่ 2.3 เว็บไซต์ Roojai.com
(ที่มา <https://www.roojai.com/car-insurance/>)



รูปที่ 2.4 หน้าเสนอแผนประกันภัยรถยนต์ในเว็บไซต์ Roojai.com
(ที่มา <https://www.roojai.com/car-insurance/>)

2.2.3. Easy Compare

เว็บไซต์แนะนำประกันภัยสำหรับรถยนต์ที่เปรียบเทียบประกันรถจากบริษัทชั้นนำในไทย 11 บริษัทที่ผ่านการรับรองจากสำนักงาน คปอ. ผู้ใช้จะต้องกรอกข้อมูลรถยนต์และข้อมูลติดต่อ ดังรูปที่ 2.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อผู้ใช้งานกรอกข้อมูลเรียบร้อยแล้วและคลิกเช็คเบี้ย ระบบจะแนะนำรายการแผนประกันภัยตามข้อมูลรถยนต์พร้อมทุนประกันที่แนะนำ สามารถเลือกเป็นประกันชั้นหนึ่ง ชั้นสอง และชั้นสามได้ หากต้องการประกันที่ดีที่สุดก็สามารถกรอกข้อมูลเพิ่มเติมได้ ดังรูปที่ 2.6

เปรียบเทียบประกันรถ จากบริษัทชั้นนำ

เช็คเบี้ยประกันรถจากบริษัทชั้นนำ
ใบอนุญาตกรมธรรม์ประกันภัยรถยนต์ เลขที่ ๖00007/2560

ปีรถ: 2016 | สี: ส้ม | ยี่ห้อ: Honda | รุ่น: City Sedan 4 Doors S 1.5 2016 | ประเภท: 28 วัน | บริษัท: ไทยพาณิชย์

รูปที่ 2.5 เว็บไซต์ Easy Compare
(ที่มา <https://www.easycompare.co.th/>)

ขอบคุณสำหรับข้อมูลของคุณ

- สามารถใช้งานได้จริงของข้อมูลประกันภัยรถยนต์
- หมายเลขติดต่อฉุกเฉิน 02-206-8555 (ในวันปกติ) หรือ 02-206-8555 (ในวันหยุด)

ต้องการแผนที่ดีที่สุดสำหรับคุณหรือไม่?

ประกันชั้น 1	ประกันชั้น 2+	ประกันชั้น 3+	ประกันชั้น 3
10,000	330,000	5,000	ไม่ระบุ
ประกันภัย (คืนเงิน)	ค่าความเสียหายส่วนแรก	ประเภทการรื้อ	พ.ร.บ.
8,976	ฟรีค่าซ่อมเปลี่ยน	สองผู้	ไม่ระบุ
ฟรีค่าซ่อมเปลี่ยน	อิสระตลอดชีพ		

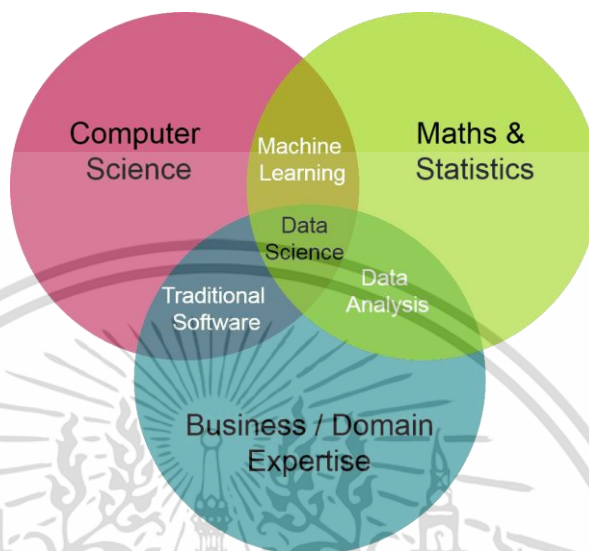
รูปที่ 2.6 หน้าแนะนำประกันในเว็บไซต์ Easy Compare
(ที่มา <https://www.easycompare.co.th/>)

2.3. วิทยาศาสตร์ข้อมูล (Data Science)

การนำข้อมูลหรือองค์ความรู้มาเพื่อใช้ประโยชน์เพื่อค้นหาผลลัพธ์ที่ต้องการรู้ โดยใช้ศาสตร์จากวิทยาการคอมพิวเตอร์ คณิตศาสตร์และสถิติ และความรู้ด้านธุรกิจ ซึ่งกระบวนการทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิทยาศาสตร์ข้อมูลจะครอบคลุมตั้งแต่การเก็บข้อมูล การจัดการข้อมูล การวิเคราะห์ข้อมูล และการนำข้อมูลมาช่วยตัดสินใจ ดังรูปที่ 3.2



รูปที่ 2.7 ศาสตร์ที่เกี่ยวข้องกับวิทยาศาสตร์ข้อมูล

2.4. การเรียนรู้ของเครื่อง (Machine Learning)

การทำให้คอมพิวเตอร์สามารถเรียนรู้ด้วยตนเองได้โดยปราศจากการควบคุมจากโปรแกรมเมอร์ คอมพิวเตอร์จะเรียนรู้จากข้อมูลนำเข้า (Input) ที่ผู้ใส่ส่งเข้าไป แล้ววิเคราะห์ผ่านอัลกอริทึมที่ได้ติดตั้งไว้ จากนั้นถึงส่งออกข้อมูล (Output) ที่ผ่านการวิเคราะห์แล้ว ซึ่งผลลัพธ์ที่ได้ก็จะถูกนำไปใช้ประโยชน์เชิงลึกต่อไป สามารถแบ่งออกเป็น 3 ประเภทดังแสดงในรูปที่ 2.3 ต่อไปนี้

2.4.1. Supervised Learning

การเรียนรู้แบบมีข้อมูลที่กำหนดผลลัพธ์แล้ว คอมพิวเตอร์จะนำข้อมูลจากพีเจอร์มาสอนตนเองจนสามารถคาดเดาผลลัพธ์จากข้อมูลนำเข้าได้เอง การเรียนรู้ลักษณะนี้จะเป็นรูปแบบของ Regression และ Classification

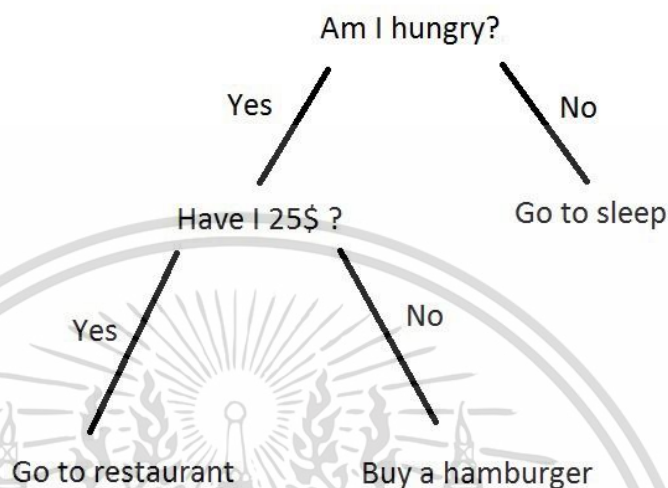
ในส่วนของกรนำเทคนิคและโมเดลที่นำมาใช้ประกอบการทำโครงการ มีทั้งหมด 2 โมเดล ได้แก่

1) Decision Tree

โมเดลการเรียนรู้ของเครื่องประเภทการจำแนก (Classification Model) เป็นแบบจำลองทางคณิตศาสตร์ที่ใช้หาทางเลือกที่ดีที่สุด โดยนำข้อมูลมาสร้างแบบจำลองในรูปแบบของโครงสร้างต้นไม้ โดยมีส่วนประกอบเป็น Node เป็นคุณสมบัติต่างที่ใช้แยกทิศทางของข้อมูล,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

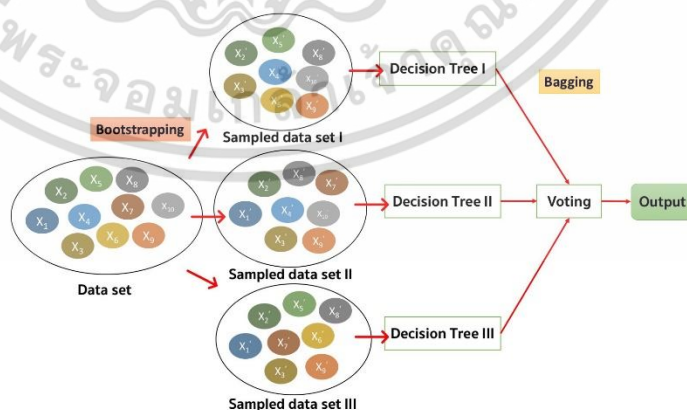
Branch เป็นคุณสมบัติของโหนดที่แตกออกมา และ Leaf เป็นกลุ่มผลการจำแนกข้อมูล ดังแสดงในรูปที่ 2.8



รูปที่ 2.8 ตัวอย่างแบบจำลอง Decision Tree

2) Random Forest

โมเดลการเรียนรู้ของเครื่องประเภทการจำแนก (Classification Model) ซึ่งสร้างจากโมเดล Decision Tree หลายโมเดลมารวมกัน โดยจะให้ Decision Tree ดังกล่าวแยกกันทำนายข้อมูลโดยตัวเอง โดยแต่ละโมเดลจะได้รับข้อมูลที่เป็นข้อมูลย่อยต่างกันภายในชุดข้อมูลเดียวกัน และเลือกผลลัพธ์ที่ถูกจัดจำแนกโดย Decision Tree มากที่สุด ดังแสดงในรูปที่ 2.9



รูปที่ 2.9 หลักการทำ Random Forest

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการสร้างโมเดลเพื่อจำแนกประเภทข้อมูลและได้มีการแบ่งข้อมูลทดสอบแล้วนั้น จะต้องมีการวัดผลประสิทธิภาพข้อมูลก่อนนำโมเดลดังกล่าวไปใช้งานจริง ด้วยวิธีต่างๆดังต่อไปนี้

1) Confusion Matrix

เครื่องมือในการประเมินผลลัพธ์ของการทำนายของโมเดลที่สร้างขึ้น โดยมีพื้นฐานในการเทียบผลลัพธ์จากสิ่งที่โมเดลทำนาย (Predicted) กับผลลัพธ์ที่เกิดขึ้นจริง (Actual) ซึ่งจะเทียบสัดส่วนจากค่าต่อไปนี้ ดังแสดงในรูปที่ 2.10

- True Positive (TP) คือ การทำนายว่าจริง สอดคล้องกับสิ่งที่เกิดขึ้นจริง
- True Negative (TN) คือ การทำนายว่าไม่จริง สอดคล้องกับสิ่งที่เกิดขึ้นไม่จริง
- False Positive (FP) คือ การทำนายว่าจริง แต่สิ่งที่เกิดขึ้นไม่เป็นจริง
- False Negative (FN) คือ การทำนายว่าไม่จริง แต่สิ่งที่เกิดขึ้นเป็นจริง

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

รูปที่ 2.10 ตาราง Confusion Matrix

โดยค่าทั้ง 4 ค่าต่อไปนี้สามารถนำมาคำนวณหาประสิทธิภาพของโมเดลในรูปแบบต่างๆ ได้แก่ Precision, Recall, F-Measure และ Accuracy ซึ่งจะกล่าวในหัวข้อถัดไปตามลำดับ

2) Precision

ผลลัพธ์จากการทำนายของโมเดลที่ถูกต้องและเป็นจริงเทียบกับผลการทำนายว่าจริงทั้งหมด (ทั้งเป็นจริงและไม่เป็นจริง) โดยมีสูตรเป็น

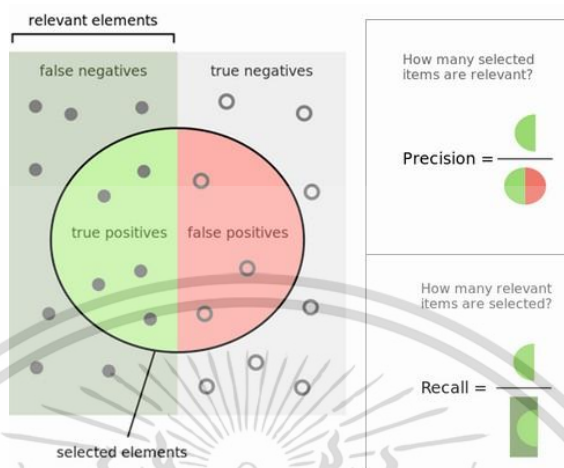
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

3) Recall

ผลลัพธ์การทำนายของโมเดลที่ถูกต้องและเป็นจริงเทียบกับการทำนายกับสิ่งที่เป็นจริง (ทั้งทำนายว่าเป็นจริงและไม่เป็นจริง) โดยมีสูตรเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$



รูปที่ 2.11 ภาพประกอบเพื่ออธิบายการคำนวณค่า Precision และ Recall

4) F-Measure

หรือ F1 Score เป็นค่าเฉลี่ยแบบ Harmonic Mean ระหว่างค่า Precision และ Recall ใช้เป็นเมตริกเดียวในการวัดความสามารถของโมเดล โดยมีสูตรเป็น

$$\text{F-Measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

5) Accuracy

ค่าความถูกต้องจากการทำนายของ โมเดลที่ทำนายได้ตรงกับสิ่งที่เกิดขึ้นจริง โดยมีสูตรเป็น

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

6) Top-K Precision

การวัดผล Precision โดยมีหลักการคือเลือกของขึ้นมา N สิ่งจากรายการแนะนำที่มีของ K สิ่งจากรายการสิ่งของทั้งหมด แล้วนำมาเทียบอัตราส่วนกัน โดยมีสูตรเป็น

$$\text{Top-K Precision} = \frac{N}{K}$$

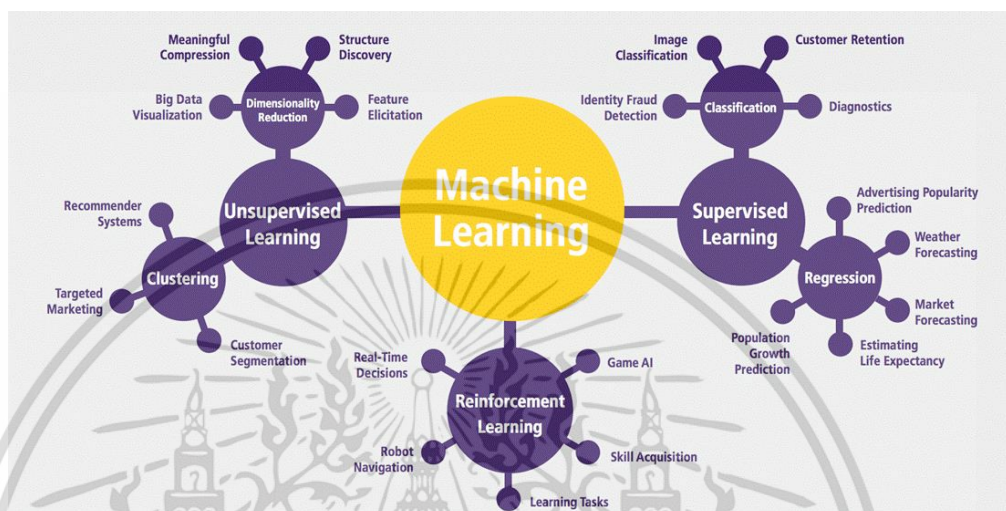
2.4.2. Unsupervised Learning

การเรียนรู้แบบมีข้อมูลที่ไม่กำหนดผลลัพธ์ คอมพิวเตอร์จะไม่สามารถรู้ได้เลยว่าผลลัพธ์ของข้อมูลชนิดนี้เป็นอะไร จึงทำได้เพียงการจัดกลุ่มจากการกระจุกตัวของข้อมูลเท่านั้น เรียกว่า Cluster

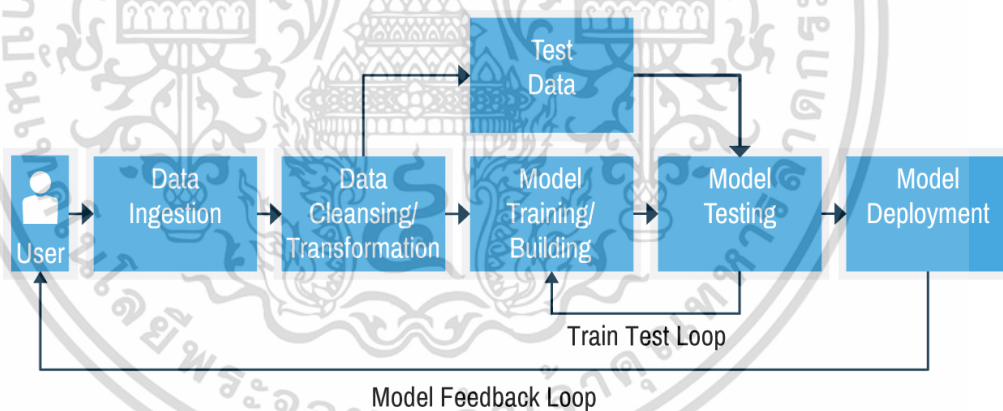
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.3. Reinforcement Learning

การเรียนรู้แบบเสริมกำลัง เป็นการเรียนรู้จากสภาพแวดล้อม กฎเกณฑ์ ลองผิดลองถูก สามารถประยุกต์ผลลัพธ์ได้ตามข้อมูลนำเข้าหรือสภาพแวดล้อมที่เปลี่ยนไปได้



รูปที่ 2.12 ประเภทของการเรียนรู้ของเครื่อง



รูปที่ 2.13 กระบวนการทำงานของการเรียนรู้ของเครื่อง

จากรูปที่ 2.4 แสดงให้เห็นถึงกระบวนการทำงานของการเรียนรู้ของเครื่อง โดยเริ่มจากการรับข้อมูล (Data Ingestion) การทำความสะอาดหรือแปลงข้อมูล (Data Cleansing/Transformation) แบ่งข้อมูลออกเป็นข้อมูลทดสอบ (Test Data) และข้อมูลสำหรับฝึกสอน (Training Data) จากนั้นจึงนำข้อมูลฝึกสอนไปสร้างแบบจำลองเพื่อทดสอบฝึกสอน (Model Training/Building) การทดสอบแบบจำลอง (Model Testing) ซึ่งจะต้องนำข้อมูลทดสอบมาใช้ในการทดสอบแบบจำลอง เมื่อเรียบร้อยแล้วจึงนำแบบจำลองไปใช้งานอื่นต่อไป (Model Deployment)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

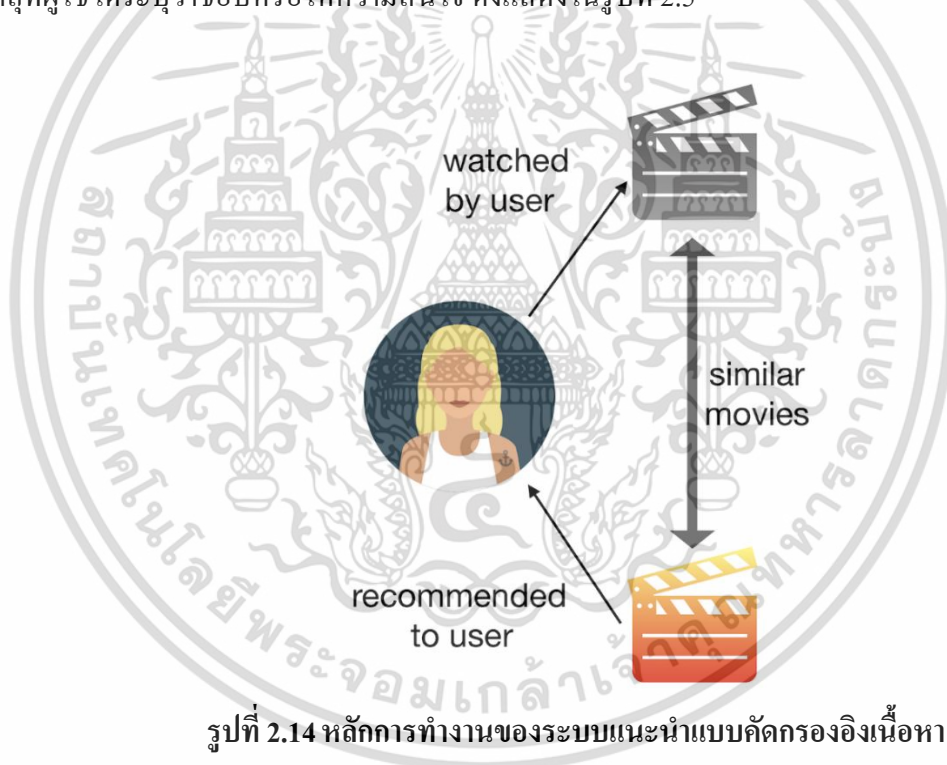
2.5. ระบบแนะนำ (Recommender System)

ระบบที่แนะนำวัตถุให้กับผู้ใช้ที่คาดว่าจะสนใจสิ่งนั้น โดยอ้างอิงจากข้อมูลคะแนนความชอบหรือลักษณะของผู้ใช้ก่อนหน้า ซึ่งเป็นที่นิยมมากในธุรกิจขายสินค้าต่างๆที่จะแนะนำสินค้าให้กับผู้ใช้ที่ไม่เคยเข้าใช้มาก่อนหรือดูสินค้าที่เกี่ยวข้องกับสิ่งที่แนะนำ เป็นต้น

ระบบแนะนำแบ่งเป็น 2 หลักๆ คือ

2.5.1. Content-Based Filtering

การคัดกรองเนื้อหา วิธีการนี้จะใช้ข้อมูลจากผู้ใช้และพฤติกรรมต่างๆของผู้ใช้มาวิเคราะห์หาความคล้ายกับคุณลักษณะของวัตถุหากผู้ใช้เคยมีปฏิสัมพันธ์กับวัตถุดังกล่าวมาก่อน และผู้ใช้ได้ให้คะแนนกับวัตถุนั้น ซึ่งการคัดกรองลักษณะนี้จะแนะนำวัตถุที่มีคุณลักษณะคล้ายกับวัตถุที่ผู้ใช้ได้ระบุว่าชอบหรือให้ความสนใจ ดังแสดงในรูปที่ 2.5



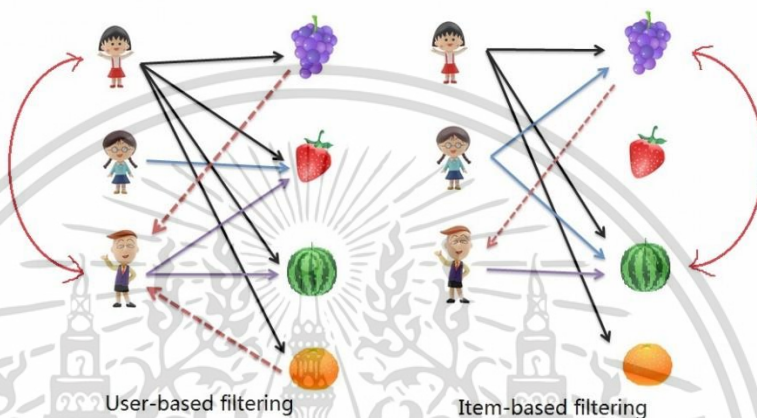
2.5.2. Collaborative Filtering

การคัดกรองร่วม วิธีการนี้จะใช้คะแนนของหลายสิ่งและผู้ใช้หลายคน มาคาดการณ์ว่าผู้ใช้ที่ไม่เคยใช้ของสิ่งนั้นจะสนใจสิ่งนั้นหรือสิ่งอื่นที่คล้ายกันหรือไม่ สามารถจำแนกการคัดกรองได้เป็น 2 แบบ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) คัดกรองโดยหน่วยความจำ (Memory Based)

การคัดกรองวิธีนี้จะใช้ข้อมูลผู้ใช้และวัตถุเพียงอย่างเดียว โดยยึดจากการกระทำของผู้ใช้ที่มีต่อวัตถุเพื่อหาความคล้ายและแนะนำผู้ใช้ใหม่ สามารถแบ่งการคัดกรองออกเป็น คัดกรองโดยผู้ใช้ และคัดกรองโดยวัตถุ ดังรูปที่ 2.6



รูปที่ 2.15 หลักการทำงานของระบบแนะนำแบบคัดกรองร่วม

- คัดกรองโดยผู้ใช้ (User-Based Filtering)
เป็นวิธีการคัดกรองโดยใช้ข้อมูลของวัตถุที่มีความคล้ายคลึงและใกล้เคียงกันกับผู้ใช้ในการแนะนำวัตถุ เช่น ความชอบ กิจกรรมต่างๆ ที่กระทำต่อวัตถุ เป็นต้น
- คัดกรองโดยวัตถุ (Item-Based Filtering)
เป็นวิธีการคัดกรองโดยใช้ความคล้ายคลึงกันของคุณสมบัติระหว่างวัตถุมาแนะนำแก่ผู้ใช้

2) คัดกรองโดยสร้างแบบจำลอง (Model Based)

การคัดกรองวิธีนี้จะใช้ข้อมูลของผู้ใช้และวัตถุสร้างแบบจำลองขึ้นมาใหม่ เรียกว่า User-Item Interactions Matrix แล้วใช้แบบจำลองนี้ในการคาดการณ์ความสนใจของผู้ใช้ที่มีต่อสิ่งๆ หนึ่ง

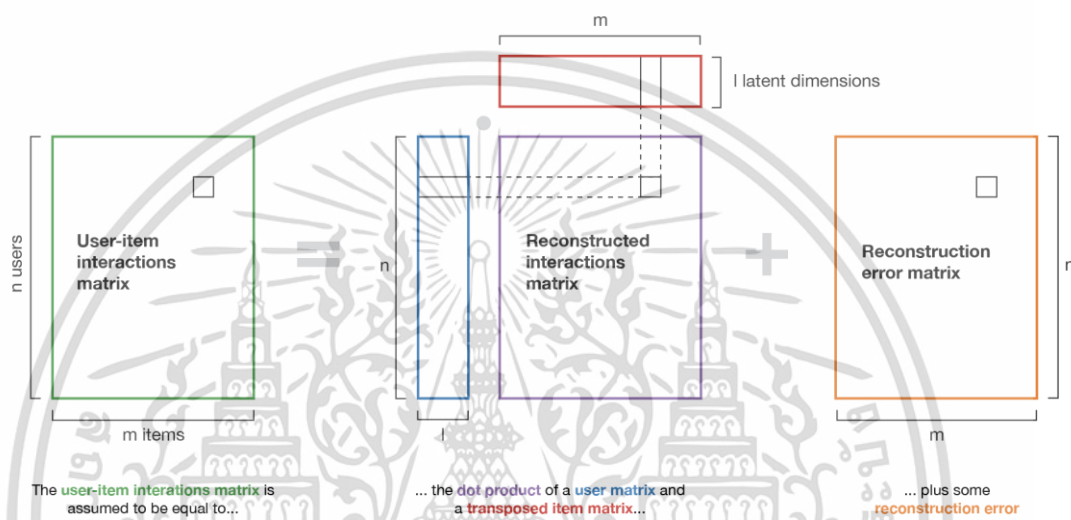
การคัดกรองวิธีนี้จะใช้วิธีการแยกตัวประกอบของเมทริกซ์ (Matrix Factorization) เป็นการนำเอาเมทริกซ์ของผู้ใช้งาน (User-factor Matrix) จำนวน n คนมาทำการคูณแบบสเกลาร์ (Dot Product) กับเวกเตอร์ทรานสโพสของวัตถุ (Item-Factor Matrix) จำนวน m สิ่ง จะได้เมทริกซ์ใหม่เป็น User-Item Interaction Matrix ขนาด $n \times m$ ซึ่งสามารถสร้างได้จากสมการ

$$M = X \cdot Y^T$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ M คือ User-Item Interaction Matrix ขนาด $n \times m$, X คือเมตริกซ์ของผู้ใช้จำนวน n แถว และ Y^T คือเมตริกซ์ทรานสโพสของวัตถุจำนวน m หลัก

แต่ถ้าหากจะนำเมตริกซ์นี้ไปบรรยายผู้ใช้งาน จะต้องเป็นเมตริกซ์ที่มีขนาดไม่ใหญ่มาก จึงต้องทำให้เป็นเมตริกซ์อันดับต่ำ (Low-Rank Matrix) ที่จะบรรยายทั้งผู้ใช้งานและวัตถุที่มีปฏิสัมพันธ์กับผู้ใช้งาน ดังรูปที่ 2.7



รูปที่ 2.16 Matrix Factorization

2.6. การเก็บรวบรวมข้อมูล (Data Collection)

ขั้นตอนเพื่อให้ได้มาซึ่งข้อมูลที่ต้องการนำไปวิเคราะห์ให้ตอบสนองกับความต้องการผลลัพธ์ซึ่งครอบคลุมทั้งการเก็บข้อมูลใหม่ด้วยตนเองหรือรวบรวมข้อมูลที่มีอยู่ก่อนแล้ว ข้อมูลที่เก็บได้แบ่งได้เป็น 3 ประเภทคือ ข้อมูลเชิงปริมาณ และ ข้อมูลเชิงคุณภาพ

2.7. การทำความสะอาดข้อมูล (Data Cleansing)

ขั้นตอนที่คัดกรองข้อมูลดิบโดยเอาส่วนที่ไม่ต้องการออก ข้อมูลที่ไม่ต้องการเหล่านี้มักเป็นข้อมูลที่ไม่ถูกต้อง ข้อมูลที่ไม่สัมพันธ์กัน อาจส่งผลต่อการนำข้อมูลไปใช้ต่อได้ การทำความสะอาดข้อมูลจึงเป็นกระบวนการที่ทำให้ข้อมูลมีประสิทธิภาพมากยิ่งขึ้น

2.8. การเตรียมพร้อมสำหรับข้อมูล (Data Preprocessing)

ขั้นตอนที่ทำให้ข้อมูลอยู่ในรูปแบบที่ถูกต้อง แม่นยำ และอยู่ในสภาพเตรียมพร้อมที่สุดก่อนนำไปใช้งาน สามารถแบ่งขั้นตอนได้หลักๆเป็น 3 ขั้นตอน ดังนี้ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8.1. Standardizing Data

เนื่องจากว่าข้อมูลที่ได้มีขนาดในแต่ละฟีเจอร์ไม่เท่ากัน ถึงแม้ว่าข้อมูลจะสัมพันธ์กัน แต่ยังมีค่าเฉลี่ยและค่าความแปรปรวนต่างกัน จึงต้องปรับปรุงให้ข้อมูลมีค่าความแปรปรวนใกล้เคียงกันและมีค่าเฉลี่ยเป็นกลางมากขึ้น เรียกขั้นตอนนี้ว่า Data Scaling

2.8.2. Feature Engineering

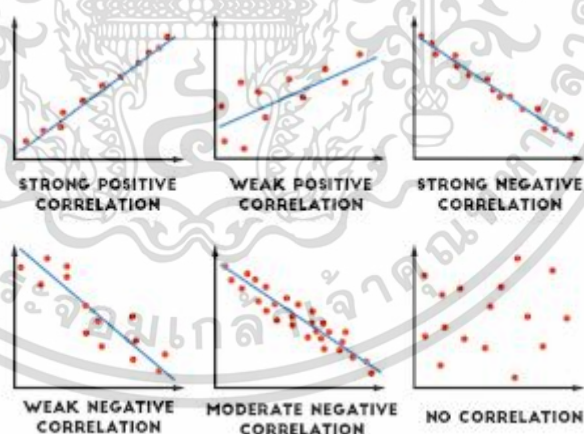
ขั้นตอนนี้เป็นการแปลงข้อมูลประเภทตัวอักษรให้กลายเป็นตัวเลข เพราะ Machine Learning ไม่สามารถรับข้อมูลประเภทตัวอักษรได้ รับเฉพาะข้อมูลที่เป็นตัวเลขเท่านั้น สามารถทำได้หลายวิธีเช่นการทำ Label Encoder, 1-hot Encoding, TF-IDF เป็นต้น

2.8.3. Feature Selection

ขั้นตอนการเลือกฟีเจอร์สำหรับการวิเคราะห์ข้อมูล การเลือกฟีเจอร์ถูกต้องสำคัญต่อการทำโมเดลอย่างมากและทำให้เห็นผลลัพธ์ที่ออกมาชัดเจนยิ่งขึ้น ซึ่งขั้นตอนนี้จะมี 2 ย่อยได้แก่

1) Correlated Feature

วิธีนี้จะใช้หาความสัมพันธ์ของคู่ฟีเจอร์ หากคู่ใดมีค่าเข้า 1 หรือ -1 มากที่สุด หมายถึงคู่ที่มีความสัมพันธ์ต่อกันสูง หากข้อมูลมีการกระจายมากหมายถึงคู่ที่มีความสัมพันธ์น้อยหรือไม่มีเลย ดังรูปที่ 2.8

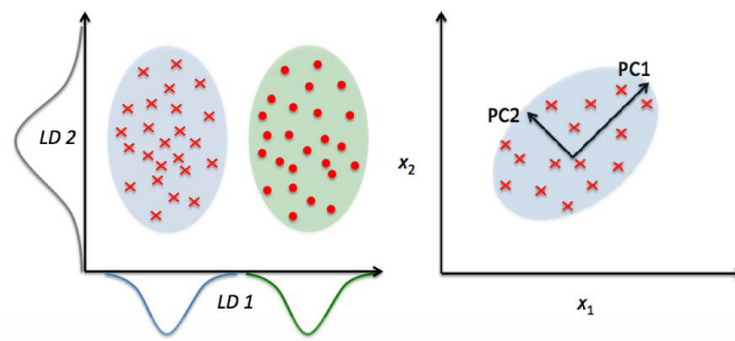


รูปที่ 2.17 กราฟแสดงค่าความสัมพันธ์ของคู่ฟีเจอร์

2) Dimensionality Reduction

ขั้นตอนการลดมิติของข้อมูล โดยจะเลือกแต่ฟีเจอร์ที่จำเป็นและส่งผลกระทบต่อโมเดลเท่านั้น ขั้นตอนนี้เป็นวิธีการของ Unsupervised Learning ดังแสดงในรูปที่ 2.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.18 การลดมิติของข้อมูล

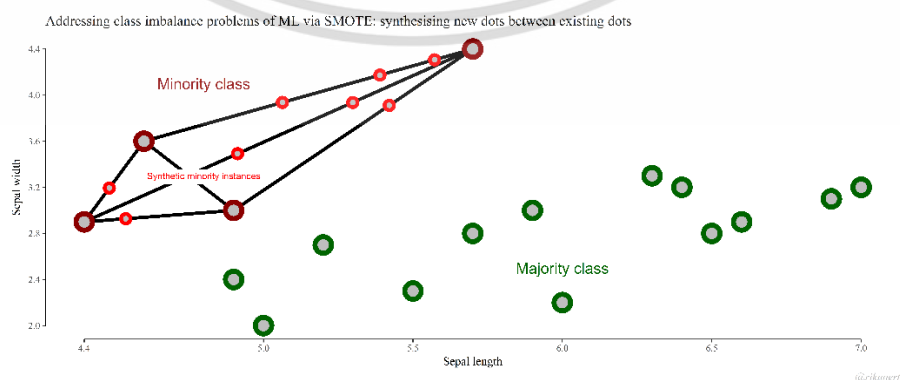
2.9. การวิเคราะห์ข้อมูล (Data Analysis)

ขั้นตอนการนำข้อมูลผ่านการเตรียมพร้อมแล้วมาวิเคราะห์ด้วยวิธีการทางสถิติเพื่อศึกษาพฤติกรรมของข้อมูล โดยสร้างแบบจำลองหรือโมเดลวิเคราะห์ข้อมูลผ่านอัลกอริทึมต่างๆเพื่อให้ได้มาซึ่งผลลัพธ์ที่ต้องการ

อัลกอริทึมต่างๆจากการเรียนรู้ของเครื่องที่เป็นที่นิยมนำมาใช้มีมากมาย เช่น Linear Regression, Logistic Regression, Decision Tree, Random Forest, Naïve Bayes, Support Vector Machine, K Nearest Neighbors, K Mean Clustering เป็นต้น ซึ่งถูกนำไปใช้ประโยชน์ต่างๆ เช่นการคาดเดาผลลัพธ์ (Prediction) หรือทำระบบแนะนำ (Recommender System)

2.10. เทคนิคการสังเคราะห์เพิ่มตัวอย่างข้างน้อย (Synthetic Minority Over-Sampling Technique)

เทคนิคการสังเคราะห์เพิ่มตัวอย่างข้างน้อย (Synthetic Minority Over-Sampling หรือ SMOTE) เป็นเทคนิคสำหรับจัดการปัญหาข้อมูลที่ไม่มีความสมดุลระหว่างกลุ่ม ด้วยการเพิ่มจำนวนข้อมูลนั้น จะทำให้การสังเคราะห์ข้อมูลขึ้นมาใหม่ ณ จุดที่อยู่ระหว่างข้อมูลเดิม ดังรูป 2.14



รูปที่ 2.19 การสังเคราะห์ข้อมูลใหม่ด้วย SMOTE

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.11. เทคโนโลยีที่ใช้

2.11.1. Python



รูปที่ 2.20 โลโก้ของ Python

เป็นภาษาโปรแกรมระดับสูงที่ไม่ยึดติดกับแพลตฟอร์ม และเป็น OpenSource ที่เปิดให้คนมาพัฒนาภาษาให้มีประสิทธิภาพสูงขึ้น และใช้งานได้ครอบคลุมทุกลักษณะงาน โดยภาษาถูกพัฒนาไว้อย่างรัดกุมเพื่อให้เข้าใจง่าย มีโครงสร้างไม่ซับซ้อน มีลักษณะการประมวลผลที่ละบรรทัดเหมือนภาษา C แต่มีโครงสร้างทางภาษาน้อยกว่า

ภาษาไพธอนมี Interpreter ที่ใช้งานได้ในตัว ซึ่งยังสนับสนุน Interactive Shell ให้สามารถเขียนโปรแกรมบน Command Line Interface ได้

2.11.2. Anaconda



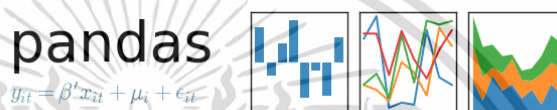
รูปที่ 2.21 โลโก้ของ Anaconda

เป็นตัวแจกจ่ายของภาษาไพธอนที่รวมหลายโมดูลเข้าด้วยกัน เช่น โมดูลสำหรับการทำวิทยาศาสตร์ข้อมูล เช่น Data Analysis, Data Visualization, Machine Learning, Deep Learning หรือโมดูลสำหรับวิเคราะห์ภาพอย่าง OpenCV เป็นต้น

2.11.3. Numpy และ Pandas



รูปที่ 2.22 โลโก้ของ Numpy



รูปที่ 2.23 โลโก้ของ Pandas

Numpy เป็นส่วนเสริมของภาษาไพธอนที่มีความสามารถในการคำนวณคณิตศาสตร์ รวมไปถึงการจัดการข้อมูลอาร์เรย์ เวกเตอร์และเมตริกซ์ อีกทั้งยังสามารถคำนวณพีชคณิตเชิงเส้นได้อีกด้วย ซึ่งครอบคลุมการทำงานได้ใกล้เคียงกับซอฟต์แวร์สำหรับคำนวณคณิตศาสตร์อย่าง Matlab ได้เป็นอย่างดี

Pandas เป็นส่วนเสริมของภาษาไพธอนสำหรับจัดการ โครงสร้างข้อมูล มักใช้ในการจัดเตรียมข้อมูล ซึ่งข้อมูลใน Pandas สามารถแบ่งออกได้เป็น 2 ประเภทคือ Series และ Data Frame

2.11.4. Matplotlib และ Seaborn



รูปที่ 2.24 โลโก้ของ Matplotlib



Seaborn

รูปที่ 2.25 โลโก้ของ Seaborn

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Matplotlib เป็นส่วนเสริมของภาษาไพธอนสำหรับการสร้างกราฟจากข้อมูลอาร์เรย์ 2 มิติในรูปแบบแกน X และแกน Y สามารถสร้างกราฟได้หลายรูปแบบ เช่น กราฟเส้น ฮิสโตแกรม กราฟสแกตเตอร์ เป็นต้น

Seaborn เป็นส่วนเสริมของภาษาไพธอน โดยมีพื้นฐานจาก Matplotlib แต่มีจุดเด่นคือสามารถสร้างกราฟได้หลากหลายและมีผลลัพธ์ที่เข้าใจง่ายกว่า Matplotlib

2.11.5. scikit-learn



รูปที่ 2.26 โลโก้ของ scikit-learn

โมดูลของภาษาไพธอนสำหรับใช้งานด้าน Machine Learning และ Data Mining โดยเฉพาะ ประกอบด้วยอัลกอริธึมสำหรับสร้างแบบจำลองการวิเคราะห์ข้อมูลมากมายทั้งแบบ Supervised Learning และ Unsupervised Learning และครอบคลุมทุกกระบวนการทำ Machine Learning ทุกขั้นตอน ด้วยโครงสร้างที่ง่ายต่อการใช้งานแต่มีประสิทธิภาพสูง ทำให้ scikit-learn เป็นที่นิยมในการทำวิทยาศาสตร์ข้อมูล

2.11.6. Jupyter

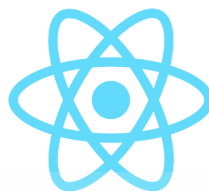


รูปที่ 2.27 โลโก้ของ Jupyter

เครื่องมือสำหรับเขียนโปรแกรมภาษาไพธอน มีจุดเด่นคือสามารถแบ่งส่วนการรันโปรแกรมและแสดงผลการทำงานได้ทันที อีกทั้งยังสามารถเขียนข้อความในรูปแบบ Markdown ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.11.7. React



รูปที่ 2.28 โลโก้ของ React

Javascript Library ที่ถูกพัฒนาโดย Facebook มีพื้นฐานการทำงานบนแนวคิดแบบ MVC (Model View Controller) และรองรับการเขียนด้วย JSX เน้นการพัฒนาเว็บไซต์ส่วนหน้าบ้าน (Front-End) โดยการเขียนเว็บจะเขียนเป็น Component, State และ Properties เป็นหลัก

2.11.8. Flask



รูปที่ 2.29 โลโก้ของ Flask

Microframework ของภาษาไพธอนที่ใช้งานด้านการสร้างเว็บแอปพลิเคชัน มีจุดเด่นคือใช้งานได้ง่ายต่อการจัดการระบบเว็บ (Back-End) และสามารถรันเซิร์ฟเวอร์ได้ในตัว สามารถรองรับส่วนขยายได้ มีประสิทธิภาพสูง

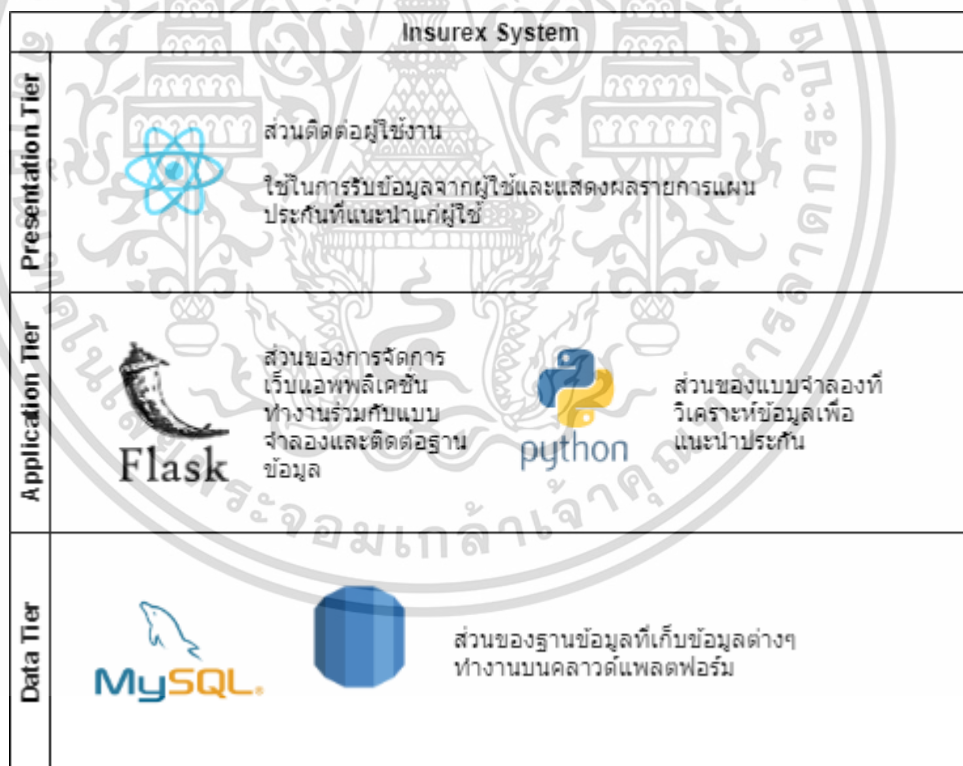
บทที่ 3

การออกแบบและพัฒนาระบบ

ในบทนี้จะกล่าวถึงการออกแบบและพัฒนาระบบแนะนำประกัน ประกอบด้วย สถาปัตยกรรมระบบ การออกแบบอัลกอริทึมของแบบจำลอง แผนภาพ Use Case แผนภาพ Sequence แผนภาพความสัมพันธ์ของฐานข้อมูล User Interface ของระบบ และข้อมูลที่ใช้ในการทดลอง

3.1. สถาปัตยกรรมของระบบ

การพัฒนาระบบจะแบ่งเป็น 2 ส่วนหลักๆ คือ ส่วนของแบบจำลองแนะนำประกันและส่วนของเว็บไซต์ ซึ่งมีหลักการทำงานดังนี้



รูปที่ 3.1 สถาปัตยกรรมของระบบแนะนำประกัน

จากรูปที่ 3.1 แสดงให้เห็นถึงหลักการทำงาน โดยภาพรวมของระบบแนะนำประกัน ระบบจะมี ส่วนติดต่อกับผู้ใช้งาน (Front-End) เพื่อรับข้อมูลจากผู้ใช้มา จากนั้นจึงส่งข้อมูลที่ได้รับไปวิเคราะห์ ด้วยแบบจำลอง Machine Learning ที่อยู่ในส่วนของการจัดการเว็บไซต์ (Back-End) เพื่อไปติดต่อ เอกสารนี้เป็นเอกสารที่สงวนเวลาสำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาดเห็นไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กับฐานข้อมูลประกันชีวิตแล้วส่งรายการแผนประกันที่ระบบแนะนำกลับไปแสดงผลที่ส่วนติดต่อผู้ใช้งาน ทุกครั้งที่ผู้ใช้งานคลิกเลือกดูแผนประกัน ส่วนติดต่อผู้ใช้งานจะส่งข้อมูลการคลิกไปที่ระบบจัดการเว็บไซต์เพื่อปรับปรุงแบบจำลองต่อไป

ในส่วนของการติดต่อผู้ใช้งานจะใช้ React ส่วนจัดการเว็บแอปพลิเคชันและฐานข้อมูลประกันใช้ Flask ซึ่งเป็นไลบรารีสำหรับสร้างเว็บแอปพลิเคชันของภาษา Python ร่วมกับ Amazon Relational Database Service (RDS) ซึ่งเป็นระบบฐานข้อมูลบนคลาวด์ที่ทำงานบนแพลตฟอร์ม Amazon Web Service (AWS) โดยฐานข้อมูลที่เลือกใช้จะใช้เป็น MySQL

3.2. ภาพรวมการทำงานและการออกแบบอัลกอริทึม

ในส่วนนี้จะกล่าวถึงการเลือกใช้ชุดข้อมูลในการทดลองทั้งการนำเข้าและส่งออก การออกแบบอัลกอริทึม และการวัดผลการทำงานของอัลกอริทึม ดังหัวข้อย่อยต่อไปนี้

3.2.1. ข้อมูลนำเข้าและข้อมูลส่งออก

ในการเลือกข้อมูลนำเข้ามาใช้ในการทดลองนั้น จำเป็นต้องเลือกคุณลักษณะที่มีผลต่อการซื้อแผนประกันชนิดต่างๆ ให้มากที่สุด เช่น อายุจะส่งผลต่อการเลือกแผนประกันที่มีระยะเวลาคุ้มครองหรือระยะเวลาการชำระเบี้ย หรือ งบเบี้ยประกันส่งผลต่อการเลือกแผนประกันที่มีระยะเวลาการชำระเบี้ยที่แปรผกผันกับจำนวนเบี้ยที่ต้องชำระ หรือพฤติกรรมการสูบบุหรี่ก็จะนำมาวิเคราะห์แผนประกันที่เกี่ยวข้องกับสุขภาพหรือโรคร้ายแรง เป็นต้น

ข้อมูลนำเข้า ประกอบด้วยข้อมูลพื้นฐานของผู้ใช้งานคือ อายุ เพศ สถานะ อาชีพ งบเบี้ยประกันที่ชำระได้ พฤติกรรมการสูบบุหรี่ เป้าหมายของการทำประกัน โดยข้อมูลดังกล่าวจะนำไปผ่านกระบวนการทำความสะอาดข้อมูลด้วยวิธี Normalization และ One-hot Encoding ดังรูปที่ 3.2

Age	SumAssured	TotalPremium	MainPremium	OtherPremium	HospitalPremium	IncomePremium	AccidentPremium	CriticalIllness	PayorPremium	Gender	Female	Gender	male	Marital	Stati	Marital	Stati	Marital	Stati	Marital	Stati	JobRate_0	JobRate_1	JobRate_2	JobRate_3
0.3570631	-0.3567582	-0.4918962	0.91249032	0.20731796	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.1032727	-0.5467881	-0.3273893	-1.6854051	-0.4267421	1.42498901	0.39581669	2.32674147	0.80183736	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.2301679	3.0637799	0.31133169	0.91249032	0.20731796	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.3570631	-0.0717134	1.11077459	0.94201186	-0.4267421	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	
1.18188189	-0.5467881	0.45657643	-1.8034912	-0.4267421	0.89964788	1.24818846	-0.4378252	2.81513987	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.4205107	0.21333146	0.06881301	-1.3311466	-0.4267421	0.61677189	4.41414076	-0.4378252	-0.1209263	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.1032727	1.16348094	-0.2917788	0.91249032	0.20731796	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.2936155	-0.5467881	-0.2544778	-1.7149266	-0.4267421	1.38457815	0.27404929	2.32674147	1.05390018	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.2936155	1.16348094	-0.2462925	0.91249032	0.20731796	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.6108535	-0.0717134	-0.1775514	0.94201186	-0.4267421	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.7377487	-0.5467881	-0.4364229	-1.5673189	-0.4267421	1.10170216	0.51798409	2.91914861	0.80183736	-0.2069354	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	
-0.0236225	0.1373195	0.30678973	0.91249032	0.20731796	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
2.07014828	-0.5467881	0.47034466	-1.6558835	-0.4267421	2.87977981	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
-1.6732601	-0.4707762	-0.0578969	0.94201186	-0.4267421	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
-0.9118889	0.30834641	-0.4043646	-0.7997589	-0.4267421	1.0612913	0.88288628	-0.4378252	-0.1209263	-0.2069354	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	
1.75291028	-0.5467881	-0.3790813	-1.1835389	-0.4267421	2.23320612	-0.4565551	-0.4378252	-0.5403643	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	
-0.0236225	0.02330157	-0.3338397	-0.9768881	-0.4267421	0.73800446	1.00465367	1.33939625	0.21462413	-0.2069354	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.2936155	-0.5467881	-0.6904233	0.8292804	-0.4267421	-0.6763755	-0.4565551	-0.4378252	4.49289195	-0.2069354	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	
0.86464389	-0.4517732	-0.3528794	0.88296878	0.841378	-0.6763755	-0.4565551	-0.4378252	-0.5403643	-0.2069354	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	
-0.2774129	-0.5467881	-0.7640577	0.2039734	-0.4267421	-0.6763755	-0.4565551	-0.4378252	1.5668258	-0.2069354	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	

รูปที่ 3.2 ตัวอย่างข้อมูลทดลองที่ผ่านการทำความสะอาดแล้ว

ข้อมูลส่งออก ประกอบด้วยรายแผนประกันชีวิตที่เหมาะสมกับผู้ใช้โดยอ้างอิงจากข้อมูลที่ผู้ใช้กรอกมาและคาดว่าผู้ใช้งานสนใจ จำนวน 5 แผน

3.2.2. ภาพรวมการทดลอง

ในการทดลองวิเคราะห์ข้อมูลเพื่อวิเคราะห์หารายการแผนประกันที่แนะนำ มีกระบวนการดังต่อไปนี้

- 1) Data Collection - คณะผู้จัดทำได้เดินทางไปเก็บข้อมูลด้วยตัวเองที่บริษัทประกันหนึ่งแห่ง โดยทำการกรอกข้อมูลจากเล่มกรมธรรม์จำนวน 8 เล่ม เก็บรวบรวมเป็นไฟล์ Excel
- 2) Data Cleansing - ทำความสะอาดข้อมูลโดยการกรองเอาข้อมูลที่ไม่ถูกต้องหรือข้อมูลบางส่วนขาดหายออก
- 3) Data Pre-Processing – ทำการคัดเลือกคุณลักษณะที่คาดว่าจะมีความเกี่ยวข้องกับผลการเลือกซื้อแผนประกัน เปลี่ยนสภาพข้อมูลตัวอักษรเป็นตัวเลขเพื่อการนำไปวิเคราะห์ที่มีประสิทธิภาพและหลากหลายมากขึ้น ปรับสเกลข้อมูลด้วยวิธี Normalization ปรับข้อมูลประเภท Categorical ให้เป็นคอลัมน์แยกด้วยวิธี One-Hot Encoding รวมไปถึงการสังเคราะห์ตัวอย่างข้อมูลเพิ่มเติมด้วยวิธี SMOTE ในกรณีคลาสที่มีน้อยเกินไป จนได้ข้อมูลพร้อมสำหรับการใช้งานดังข้อที่ 3.2.1
- 4) Data Exploration - ทำการสำรวจชุดข้อมูลที่จะใช้ในการทดลอง โดยการสร้างกราฟรูปแบบ Histogram, Scatter Plot และ Heatmap เพื่อสังเกตการณ์กระจายตัวของข้อมูลในแต่ละคอลัมน์ และหาค่าความสัมพันธ์ระหว่างคู่คอลัมน์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) Data Analysis – ขั้นตอนการใช้โมเดลหรือแบบจำลองในการวิเคราะห์ทำนายผลของข้อมูลแบ่งออกเป็นขั้นตอนย่อย 2 ขั้นตอนดังนี้

- ขั้นตอนการวิเคราะห์ข้อมูลของฟีเจอร์ของประกัน โดยโมเดลที่เลือกใช้จะเลือกใช้นั้นเป็นโมเดลที่ใช้ทำนายข้อมูลประเภท Classification ได้แก่ Decision Tree, Random Forest, K-Nearest Neighbor, Support Vector Machine, Logistic Regression และ Naïve Bayes เนื่องจากข้อมูลนำเข้าดังในข้อที่ 3.2.1 เป็นข้อมูลประเภท Categorical จึงเลือกใช้เทคนิคเหล่านี้เพราะโมเดลดังกล่าวสามารถทำงานได้ดีกับชุดข้อมูลประเภท Categorical จากนั้นจึงคัดเลือกโมเดลที่ให้ผลความแม่นยำสูงที่สุดในแต่ละฟีเจอร์ของประกันแต่ละชนิดมาเป็นข้อมูลนำเข้าของประกันแต่ละชนิด
- ขั้นตอนการหาค่าความแม่นยำระหว่างประกันที่โมเดลแนะนำกับประกันที่ผู้ใช้ซื้อจริง หลังจากได้โมเดลส่งออกเป็นการทำนายผลของฟีเจอร์ของประกันแล้วจะนำฟีเจอร์ส่วนนี้มาเป็นข้อมูลนำเข้าประกอบกันเป็นโมเดลของแผนประกันตามชนิด (ประกันแต่ละชนิดจะมีฟีเจอร์ไม่เหมือนกัน) โดยใช้กระบวนการหาความคล้ายด้วยวิธียุคลิด เพื่อค้นหาความคล้ายระหว่างฟีเจอร์ที่ทำนายโดยโมเดลเทียบกับฟีเจอร์ของแต่ละแผนประกัน ซึ่งจะให้ผลเป็นแผนประกันที่โมเดลวิเคราะห์มาแล้วว่าเหมาะสมกับข้อมูลที่มี

6) Model Evaluation – ขั้นตอนการวัดผลข้อมูลนี้ จะกล่าวโดยละเอียดในหัวข้อ 3.2.4

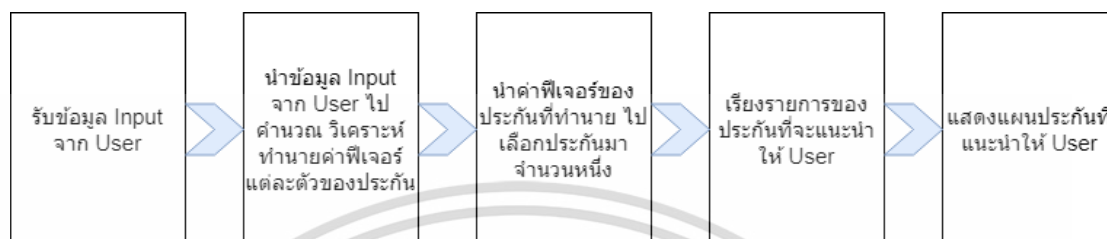
3.2.3. หลักการทำงานของอัลกอริธึม

การออกแบบอัลกอริธึมสำหรับการวิเคราะห์ข้อมูลประกันและผู้ใช้งานสำหรับแนะนำประกันนั้น จะแยกแบบจำลองตามประเภทของประกัน เพราะแต่ละประเภทประกันจะมีฟีเจอร์ไม่เหมือนกัน ซึ่งจะแยกเมื่อรับข้อมูลเป้าหมายความสำคัญของการทำประกันของผู้ใช้งาน จากนั้นจะนำข้อมูลของผู้ใช้งานไปวิเคราะห์ตามเป้าหมายที่ผู้ใช้ได้เลือกในตอนต้น เช่น หากเลือกให้ความสำคัญกับการคุ้มครองตลอดชีพ จะนำข้อมูลไปวิเคราะห์ในสายอัลกอริธึมของโมเดลคุ้มครองตลอดชีพ เป็นต้น

โดยจะมีอัลกอริธึมของโมเดลที่ใช้มาวิเคราะห์ทำนายแต่ละฟีเจอร์ของประกันซึ่งได้กล่าวในหัวข้อที่ 3.2.2 ส่วน Data Analysis เมื่อโปรแกรมทำนายได้ครบทุกฟีเจอร์ที่จำเป็นแล้ว จะนำผลลัพธ์ที่เป็นการทำนายฟีเจอร์ที่ดีที่สุดไปประกอบกันเป็นโมเดลหลักของแต่ละชนิดประกัน เพื่อคำนวณและหาแผนประกันที่ใกล้เคียงที่สุด โดยเทียบผลการทำนายฟีเจอร์กับฟีเจอร์จริงจาก

ข้อมูลในแต่ละแผนประกันด้วยวิธีการหาความคล้ายแบบยุคลิด สุดท้ายระบบจะแสดงข้อมูลเป็นเอกสารนี้เป็นข้อสรุปที่ส่งวนเวียนให้กับผู้ใช้ในเพื่อทำการเลือกให้ท่านนั้น ข้อมูลนี้เห็นได้ชัดว่าไม่ใช่ว่าการมีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

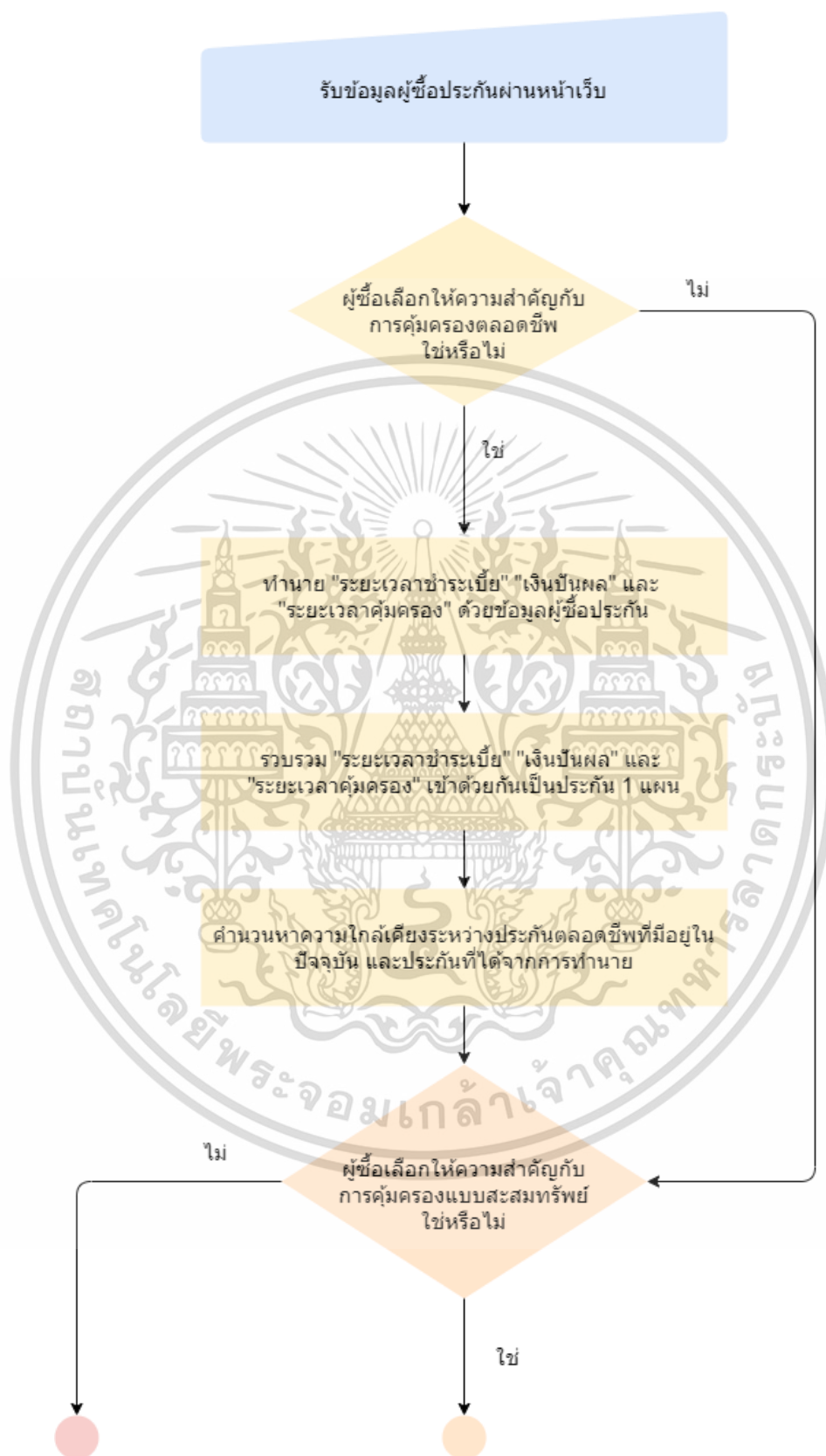
รายแผนประกันจำนวน 5 แผน (เรียงตามความใกล้เคียงของผู้ใช้มากที่สุด) ที่คาดว่าผู้ใช้งานจะสนใจในแต่ละประเภท ก่อนจะนำไปรวมกันกับประเภทอื่น (ในกรณีที่เลือกความสำคัญมากกว่า 1 ช่อง) เพื่อจัดเรียงเป็นแผนประกันที่แนะนำจริง โดยมีภาพรวมการทำงานโดยคร่าว ดังรูปที่ 3.3



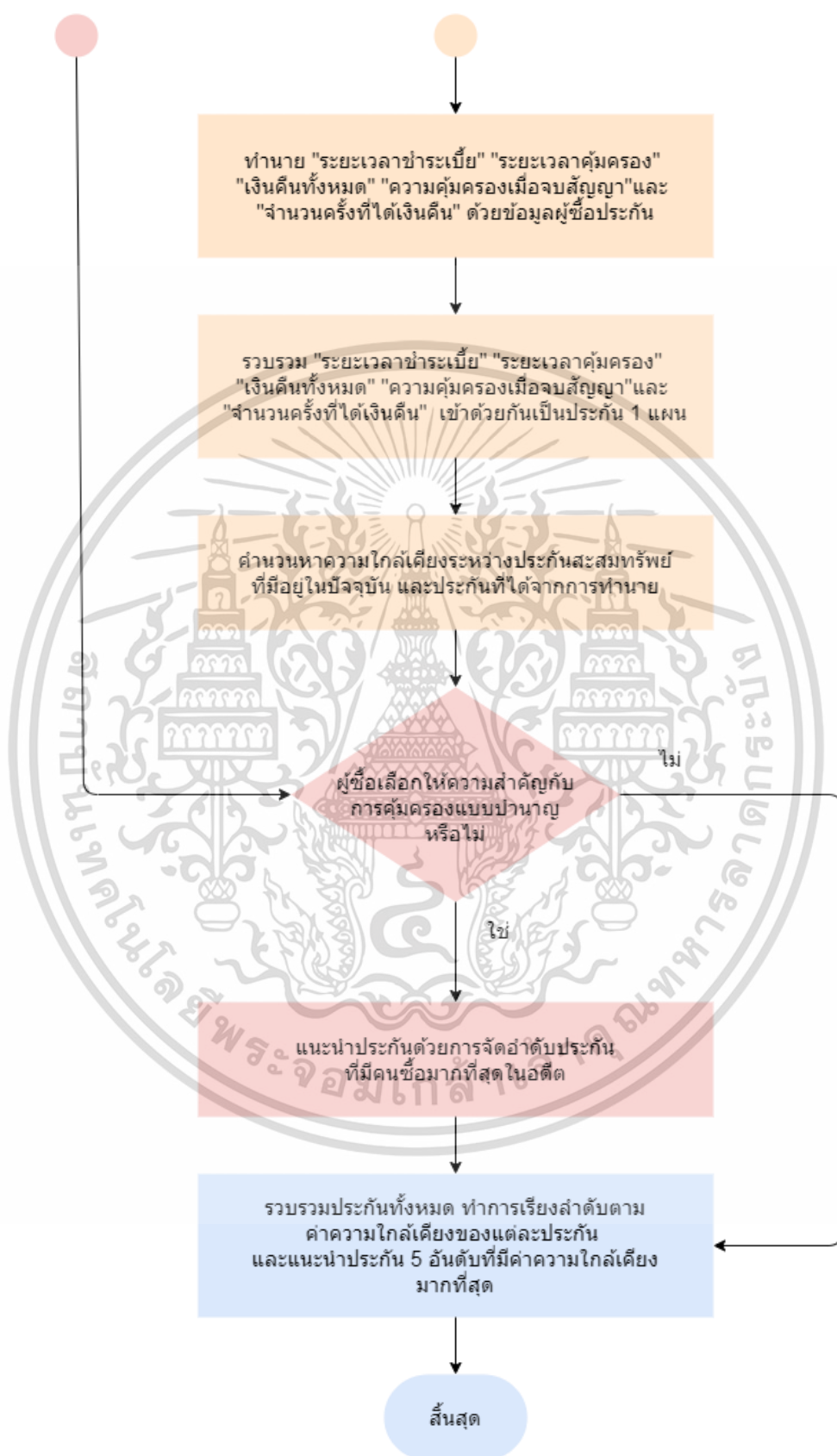
รูปที่ 3.3 ภาพรวมการทำงานของอัลกอริทึม

ซึ่งสามารถแสดงลำดับขั้นตอนการทำงานทั้งหมดของระบบได้ดังรูปที่ 3.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.4 แผนภาพ Flowcharts แสดงการทำงานของระบบแนะนำประกัน

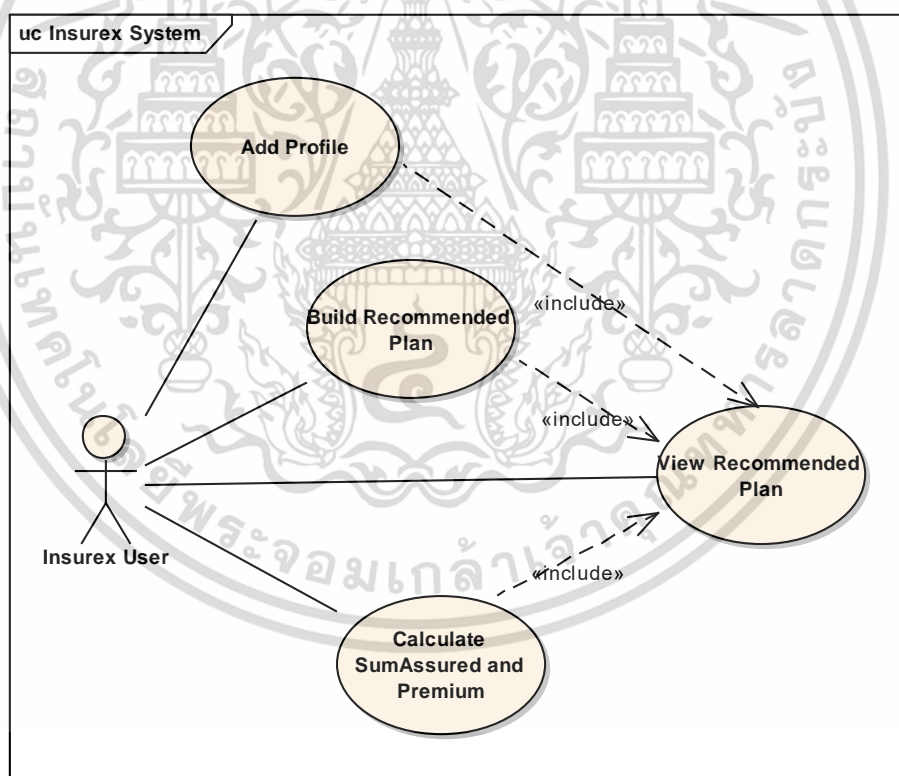
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.4. การวัดผลอัลกอริทึม

ในการวัดผลอัลกอริทึมจะแบ่งเป็น 2 ส่วนคือ

- 1) ส่วนการทำนายฟีเจอร์ - ใช้การวัดผลด้วย Confusion Matrix และ Classification Report ซึ่งประกอบด้วย Precision, Recall, F-Measure และ Accuracy โยในขั้นตอนนี้จะหาโมเดลที่ให้ผลการทำนายที่แม่นยำที่สุดในแต่ละฟีเจอร์
- 2) ส่วนการแนะนำแผนประกัน - ได้เลือกใช้การวัดผลวิธี Top-K Precision เนื่องจากเป็นการสร้างระบบแนะนำ มีจำนวนแผนประกันที่แนะนำชัดเจน จึงใช้แผนประกันจากที่ผู้ใช้เลือกซื้อจริงเป็นตัวประเมินผลการแนะนำรายการแผนประกัน หากแผนประกันที่มีการซื้อจริงตรงกับแผนประกันในรายการแนะนำ จะถือว่าโมเดลสามารถแนะนำแผนประกันได้แม่นยำ

3.2.5. Use Case Diagram



รูปที่ 3.5 แผนภาพ Use Case สำหรับระบบแนะนำประกัน

จากแผนภาพ Use Case ดังแสดงในรูปที่ 3.5 ผู้ใช้งานจะมีบทบาทเป็น Actor ซึ่งสามารถใช้งานระบบได้ตั้งแต่กรอกข้อมูลส่วนตัวตามที่ระบุไว้ในข้อ 3.2.1 เพื่อสร้างรายแผนประกันที่แนะนำจากข้อมูลที่กรอกไปข้างต้น ดูรายการแผนประกันและกดเลือกแผนประกันที่สนใจเพื่อดูข้อมูลเพิ่มเติม เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

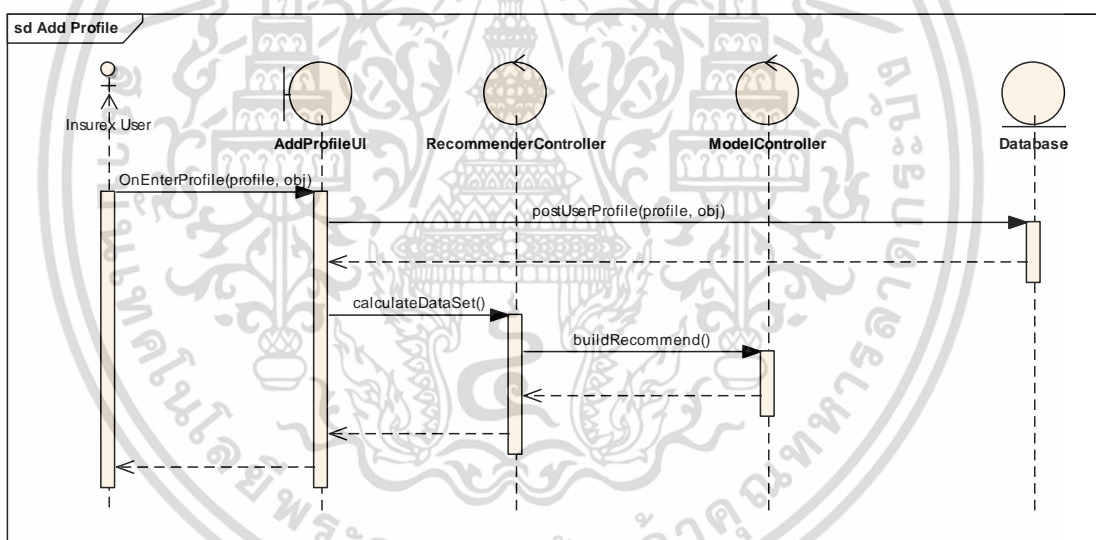
รวมไปถึงแก้ไขข้อมูลที่กรอกไปแล้วได้ รวมไปถึงจะมีการคำนวณเบี้ยประกันที่ต้องชำระและทุนประกันด้วย

3.3. Sequential Diagram

ในส่วนนี้จะกล่าวถึงการออกแบบแผนภาพ Sequence โดยจะอธิบายการทำงานของเว็บแอปพลิเคชันในแต่ละฟังก์ชัน ดังหัวข้อต่อไปนี้

3.3.1. Sequence Diagram สำหรับการกรอกข้อมูลส่วนตัว

ผู้ใช้งานต้องการกรอกข้อมูลส่วนตัวพร้อมทั้งเลือกเป้าหมายในการทำประกันตามที่ปรากฏบนหน้าจอ เมื่อกรอกเสร็จแล้วระบบจะทำการนำข้อมูลที่ได้รับไปเก็บในฐานข้อมูลและประมวลผลร่วมกับคอนโทรลเลอร์ของระบบแนะนำและแบบจำลอง ดังแสดงในรูปที่ 3.6 โดยมีตัวแปรและฟังก์ชันดังแสดงคำอธิบายในตารางที่ 3.1



รูปที่ 3.6 Sequence Diagram สำหรับการกรอกข้อมูลส่วนตัว

ตาราง 3.1 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการกรอกข้อมูลส่วนตัว

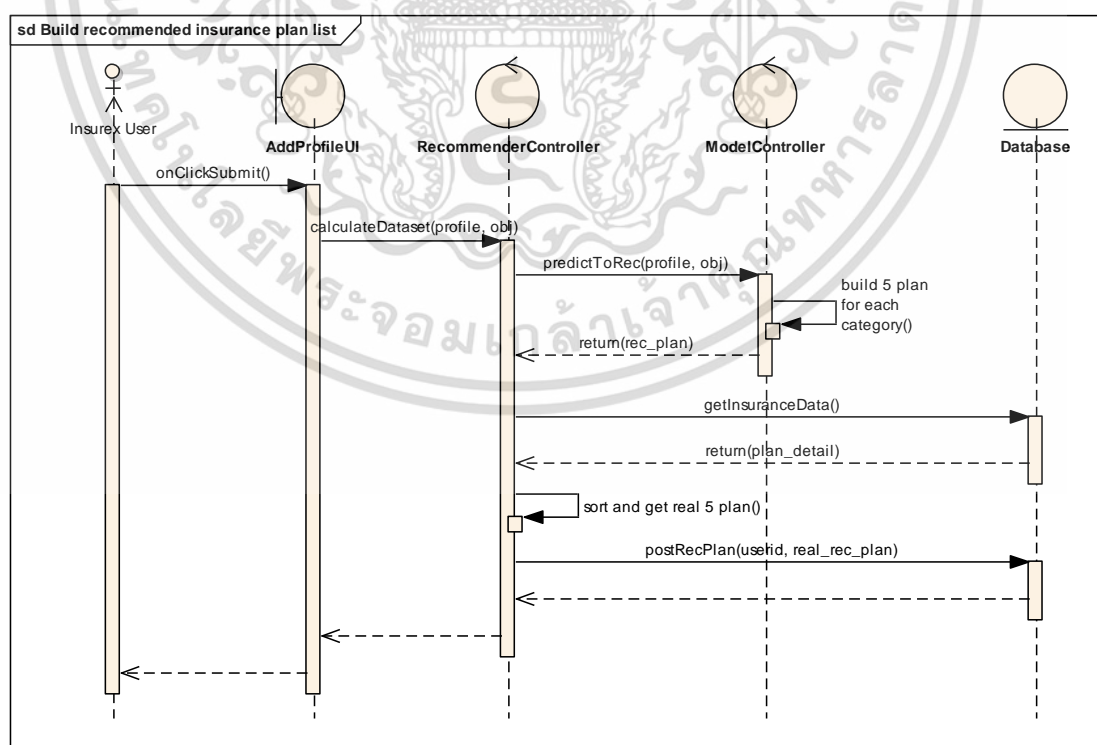
ตัวแปร	คำอธิบาย
profile (dict)	ข้อมูลส่วนตัวของผู้ใช้งาน ประกอบด้วย เพศ อายุ สถานะ พฤติกรรมการสูบบุหรี่ งบเบี้ยประกันที่ชำระได้
obj (dict)	ข้อมูลเป้าหมายในการทำประกัน ประกอบด้วย เป้าหมาย น้ำหนักความสำคัญในการคุ้มครอง (คุ้มครองตลอดชีพ สะสม ทรัพย์สิน บำนาญ และอื่นๆ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานภายในเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชัน	คำอธิบาย
onClickAddProfile	กดปุ่มเริ่มการกรอกข้อมูล
onEnterProfile	กรอกข้อมูลส่วนตัวและข้อมูลเป้าหมายในการทำประกัน
onClickSubmit	คลิกเพื่อทำการสร้างรายการแผนประกันแนะนำ
postUserProfile	ส่งข้อมูลของผู้ใช้ไปเก็บในฐานข้อมูล
calculateDataSet	คำนวณข้อมูลที่ได้อ้างอยู่ในรูปของ Dataset เพื่อส่งไปทำนายกับโมเดล
buildRecommend	สร้างแผนประกันจากแบบจำลอง

3.3.2. Sequence Diagram สำหรับการสร้างรายแผนประกันที่แนะนำ

หลังจากที่ผู้ใช้กรอกข้อมูลเสร็จและกดปุ่มแนะนำประกันแล้ว ระบบจะทำการนำข้อมูลที่ได้รับไปประมวลผลร่วมกับแบบประกันที่มีอยู่ในฐานข้อมูลและข้อมูลผู้ใช้ที่มีคุณลักษณะคล้ายกับผู้ใช้ปัจจุบันเพื่อสร้างรายการแผนประกันจำนวน 5 แผน แล้วจึงส่งข้อมูลการแนะนำดังกล่าวกลับไปเก็บในฐานข้อมูล ดังแสดงในรูปที่ 3.7 โดยมีตัวแปรและฟังก์ชันดังแสดงคำอธิบายในตารางที่ 3.2



รูปที่ 3.7 Sequence Diagram สำหรับการสร้างรายแผนประกันที่แนะนำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

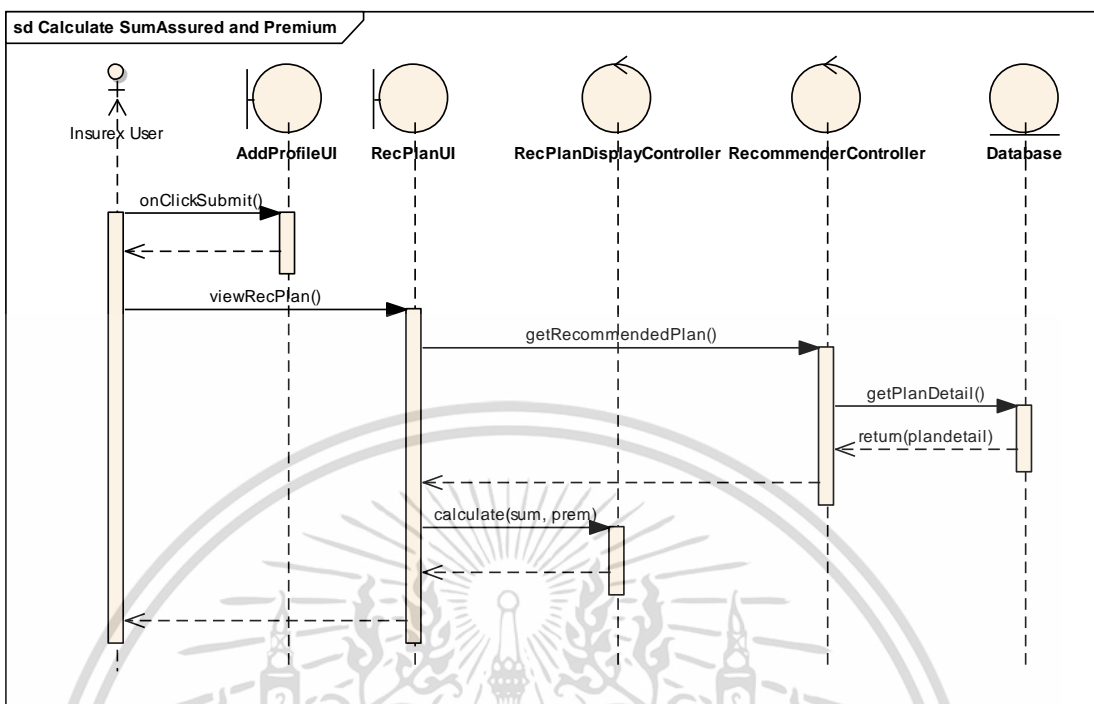
ตาราง 3.2 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการสร้างรายแผนประกันที่แนะนำ

ตัวแปร	คำอธิบาย
profile (dict)	ข้อมูลส่วนตัวของผู้ใช้งาน ประกอบด้วย เพศ อายุ สถานะ พฤติกรรมการสูบบุหรี่ งบประมาณประกันที่ชำระได้
obj (dict)	ข้อมูลเป้าหมายในการทำประกัน ประกอบด้วย เป้าหมาย น้ำหนักความสำคัญในการคุ้มครอง (คุ้มครองตลอดชีพ สะสมทรัพย์ บำนาญ และอื่นๆ)
rec_plan (dict)	รายชื่อแผนประกัน โดยมีค่า value เป็นคะแนนค่าความใกล้เคียง ซึ่งจะถูกคำนวณมาจากชนิดของแผนประกันแต่ละสาย
plan_detail (dict)	ข้อมูลของแผนประกันแต่ละแผน จะถูกดึงจากฐานข้อมูล
userid (int)	รหัสตัวแทนของ user ในการกรอกข้อมูลแต่ละครั้ง
real_rec_plan (dict)	รายการแผนประกันจริงที่ผ่านการเรียงคะแนนเรียบร้อยแล้ว ประกอบด้วยชื่อแผนประกันและรายละเอียดของแต่ละแผน
ฟังก์ชัน	คำอธิบาย
onClickSubmit	คลิกเพื่อทำการสร้างรายการแผนประกันแนะนำ
return	ส่งค่ากลับคืน
calculateDataSet	คำนวณข้อมูลที่ได้ให้อยู่ในรูปของ Dataset เพื่อส่งไปทำนายกับโมเดล
predictToRec	ทำนายแผนประกันจาก Dataset ที่ได้รับมาด้วย โมเดลของแต่ละชนิดแผนประกัน
getInsuranceData	รับข้อมูลรายละเอียดแผนประกันที่เก็บในฐานข้อมูล
postRecPlan	ส่งข้อมูลผู้ใช้และรายการแผนประกันที่แนะนำไปเก็บในฐานข้อมูล

3.3.3. Sequence Diagram สำหรับการคำนวณทุนประกันและเบี้ยประกัน

เมื่อระบบทำการสร้างแผนประกันที่แนะนำและได้รายละเอียดของแต่ละแผนแล้ว ก็ จะทำการคำนวณทุนประกันและเบี้ยประกันจากงบประมาณประกันที่ผู้ใช้ได้กรอกมาและอัตราเบี้ย ประกันต่อทุนประกันสำหรับทุนประกัน และกับอัตราส่วนความสำคัญในแต่ละเป้าหมายสำหรับ เบี้ยประกัน ดังรูปที่ 3.8 และตารางที่ 3.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 Sequence Diagram สำหรับการคำนวณทุนประกันและเบี้ยประกัน

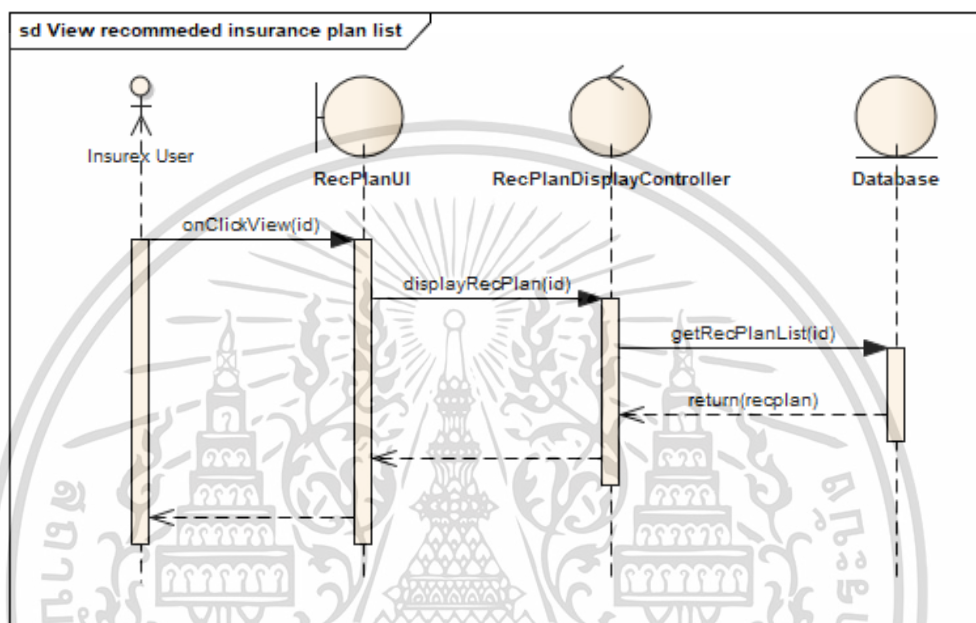
ตาราง 3.3 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการคำนวณทุนประกันและเบี้ยประกัน

ตัวแปร	คำอธิบาย
plan_detail (dict)	ข้อมูลของแผนประกันแต่ละแผน จะถูกดึงจากฐานข้อมูล
sum (int)	เงินทุนประกัน (Sum Assured)
prem (int)	เบี้ยประกัน (Premium)
ฟังก์ชัน	คำอธิบาย
onClickSubmit	คลิกเพื่อทำการสร้างรายการแผนประกันแนะนำ
viewRecPlan	ดูแผนประกันแนะนำ
getRecommendedPlan	ดึงข้อมูลแผนประกันที่แนะนำ
getPlanDetail	ดึงข้อมูลแผนประกัน
return	ส่งค่ากลับคืน
calculate	คำนวณทุนประกันและเบี้ยประกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.4. Sequence Diagram สำหรับการดูรายการแผนประกันที่แนะนำ

ระบบจะใช้ id เมื่อผู้ใช้ทำการกรอกข้อมูลและสร้างรายการแผนประกันที่แนะนำแล้ว ไปดึงข้อมูลจากฐานข้อมูลเพื่อแสดงรายการแผนประกันที่แนะนำจำนวน 5 แผน ดังแสดงในรูปที่ 3.9 โดยมีตัวแปรและฟังก์ชันดังแสดงคำอธิบายในตารางที่ 3.4



รูปที่ 3.9 Sequence Diagram สำหรับการดูรายการแผนประกันที่แนะนำ

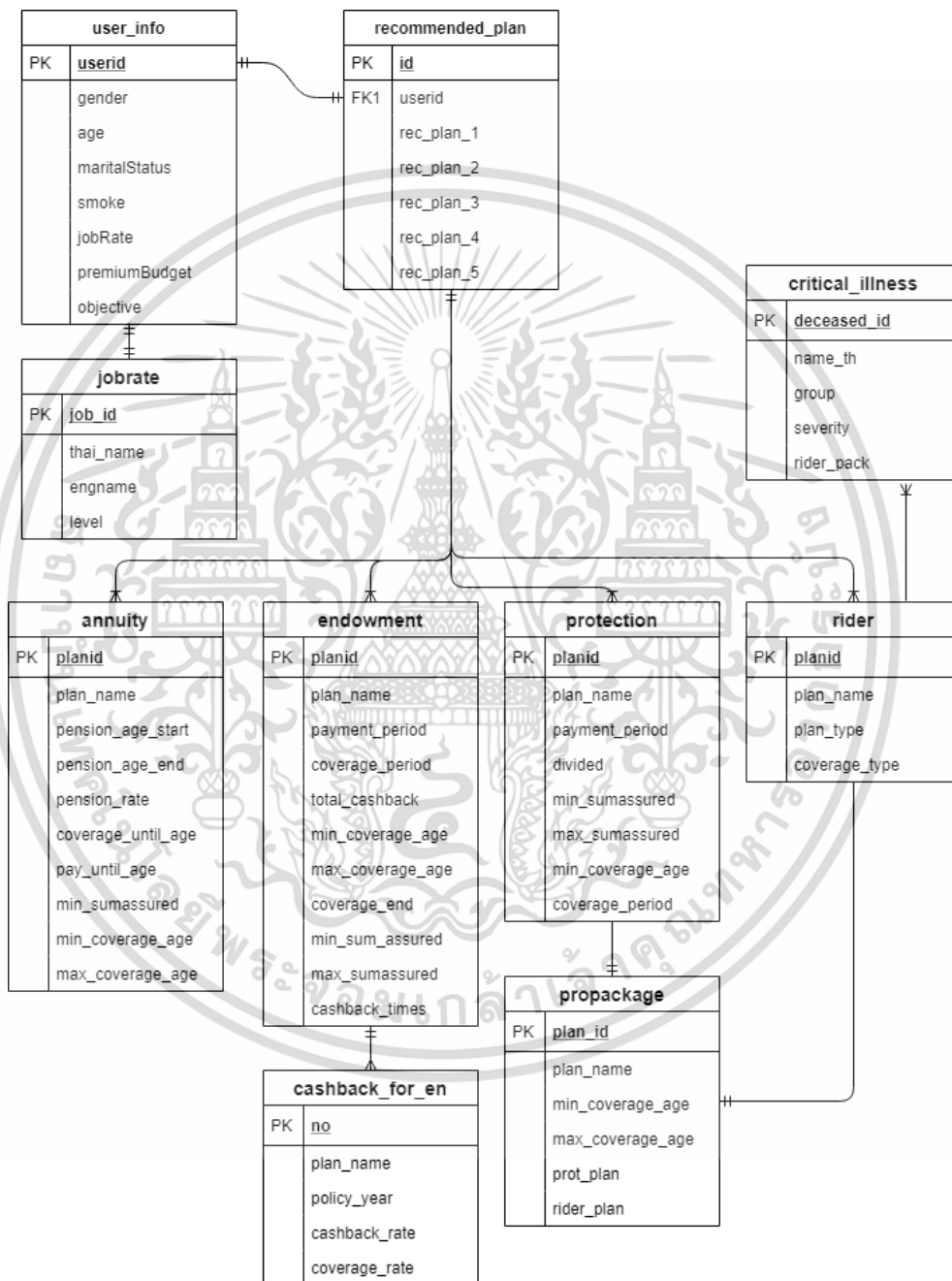
ตาราง 3.4 คำอธิบายตัวแปรและฟังก์ชันที่ใช้ในแผนภาพการดูรายการแผนประกันที่แนะนำ

ตัวแปร	คำอธิบาย
id (int)	รหัสตัวแทนของผู้ใช้งาน
recplan (list)	รายการแผนประกันที่แนะนำแก่ผู้ใช้
ฟังก์ชัน	คำอธิบาย
onClickView	คลิกเพื่อดูรายการแผนประกันที่แนะนำ
displayRecPlan	แสดงผลรายการแผนประกันที่แนะนำ
getRecPlan	ดึงข้อมูลรายการแผนประกันที่แนะนำด้วย id ของผู้ใช้
return	ส่งค่ากลับคืน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4. Relational Database Diagram

เนื่องจากข้อมูลทั้งของประกันและของผู้เอาประกัน(ผู้ใช้) มีลักษณะโครงสร้างที่ชัดเจน จึงเลือกใช้การเก็บในฐานข้อมูลแบบ SQL ซึ่งได้ออกแบบฐานข้อมูลของระบบดังแสดงในรูปที่ 3.10



รูปที่ 3.10 การออกแบบฐานข้อมูลของระบบ

จากรูปที่ 3.10 ในฐานข้อมูลของระบบแนะนำ มีความสัมพันธ์กันระหว่างตาราง โดยแต่ละตารางมีข้อมูลดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1) User_info เป็นตารางเก็บข้อมูลของผู้เอาประกันเมื่อกรอกข้อมูล
- 2) JobRate เป็นตารางเก็บข้อมูลชั้นอาชีพ
- 3) Protection เป็นตารางเก็บข้อมูลของประกันประเภทคุ้มครองตลอดชีพ
- 4) Endowment เป็นตารางเก็บข้อมูลของประกันประเภทสะสมทรัพย์
- 5) cashBackForEn เป็นตารางเก็บข้อมูลเงินคืนและอัตราความคุ้มครองตลอดสัญญาของประกันประเภทสะสมทรัพย์
- 6) Annuity เป็นตารางเก็บข้อมูลของประกันประเภทบำนาญ
- 7) Rider เป็นตารางเก็บข้อมูลของประกันสัญญาเพิ่มเติม
- 8) Critical Illness เป็นตารางเก็บข้อมูลของโรคร้ายแรง
- 9) ProPackage เป็นตารางเก็บข้อมูลแพ็คเกจประกันประเภทคุ้มครองตลอดชีพและสัญญาเพิ่มเติม
- 10) Recammended_plan เป็นตารางที่เก็บข้อมูลของแผนประกันที่แนะนำโดยอิงจากรหัสตัวแทนของผู้ใช้ในการกรอกครั้งนั้นๆ

3.5. User Interface Design

เป็นส่วนติดต่อผู้ใช้ เพื่อให้ผู้ใช้สามารถกรอกข้อมูล ดูรายละเอียดแผนประกัน และเลือกแผนประกันที่สนใจเพื่อดูรายละเอียด

3.5.1. UI สำหรับหน้าหลัก

เป็นหน้าเริ่มต้นเมื่อเข้าเว็บ เป็นหน้าที่แสดงข้อมูลเบื้องต้นของระบบ พร้อมปุ่มเพื่อนำไปสู่การแนะนำแผนประกัน ดังแสดงในรูปที่ 3.11

หน้าแรก ทำไมต้องเป็นเรา? บริการประกันที่ร่วมมือ ติดต่อเรา

Insurex

ประกันชีวิต ประกันสุขภาพ ประกันอุบัติเหตุ
อยากทำประกัน แต่ไม่รู้อะไรเกี่ยวกับประกันแบบไหน เท่าไหร่?

ให้เราช่วยสิ

ทำไมต้องเป็นเรา?

ง่ายตาย
เพียงกรอกข้อมูลส่วนตัวเล็กน้อย
เราก็สามารถแนะนำประกัน
ให้คุณได้ทันที

รูปที่ 3.11 หน้า UI สำหรับหน้าหลัก

3.5.2. UI สำหรับการกรอกข้อมูลส่วนที่ 1

เป็นหน้าสำหรับการกรอกข้อมูลส่วนตัวในส่วนที่ 1 ดังแสดงในรูปที่ 3.12

Insurex

กรอกข้อมูลส่วนตัว

เพศ
 ชาย หญิง

อายุ
0

สถานะ
 โสด สมรส หย่าร้าง หย่าอยู่

เคยสูบบุหรี่หรือไม่
 เคย ไม่เคย

อาชีพ
เลือกอาชีพ

จบชั้นประกัน (บาท/ปี)
บาท/ปี

ต่อไป

รูปที่ 3.12 หน้า UI สำหรับการกรอกข้อมูลส่วนที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

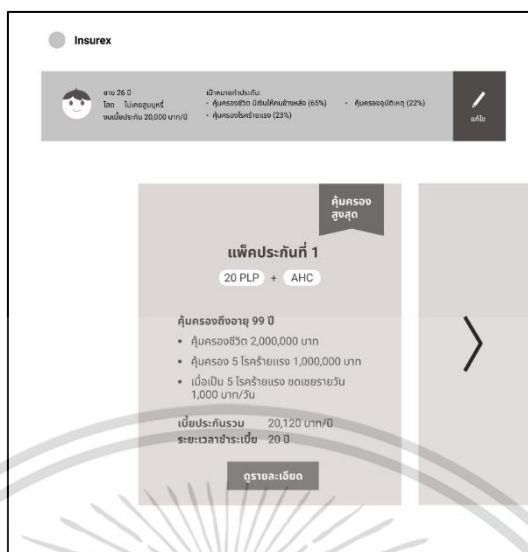
3.5.3. UI สำหรับการกรอกข้อมูลส่วนตัว 2

เป็นหน้าสำหรับกรอกข้อมูลส่วนตัวในส่วนที่ 2 ในหน้านี้จะเป็นการกำหนดเป้าหมายในการทำประกันและการให้ความสำคัญกับความเสี่ยงต่างๆ ดังแสดงในรูปที่ 3.13

รูปที่ 3.13 หน้า UI สำหรับการกรอกข้อมูลส่วนตัว 2

3.5.4. UI สำหรับหน้าแนะนำรายการแผนประกัน

เมื่อกรอกข้อมูลส่วนตัว เป้าหมาย และให้ความสำคัญกับความเสี่ยงแล้ว ระบบจะแสดงรายการแผนประกันที่แนะนำ พร้อมรายละเอียดของแผนประกันนั้นและข้อมูลที่ใช้ได้ให้ไว้ในรูปที่ 3.12 และ 3.13 อีกทั้งยังคำนวณทุนประกันและเบี้ยประกันของแต่ละแผนได้ โดยจะแสดงเป็นรายการทั้งหมด 5 รายการด้วยกัน ดังรูปที่ 3.14



รูปที่ 3.14 หน้า UI สำหรับหน้าแนะนำรายแผนประกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลองและผลการทดลอง

ในส่วนนี้จะกล่าวถึงการทดลองและผลการทดลองทั้งด้านการวิเคราะห์ข้อมูลและการพัฒนาเว็บแอปพลิเคชัน ตั้งแต่ผลการวิเคราะห์ข้อมูลเบื้องต้น ผลการทดลองอัลกอริธึมกับข้อมูลที่จะใช้จริง ผลการทดสอบโปรแกรมแนะนำประกัน ผลการพัฒนาแอปพลิเคชัน ผลการทดสอบโดยผู้เชี่ยวชาญ และผลการเปรียบเทียบกับแอปพลิเคชันแนะนำประกันอื่นๆ

4.1. บทนำการทดลอง

หลังจากออกแบบระบบภายในแล้ว ได้ทำการทดลองและวัดผลดังนี้

- 1) ภาพรวมการทดลอง
- 2) ผลการทำความสะอาดข้อมูลและวิเคราะห์ข้อมูลเบื้องต้น
- 3) ผลการทดลองอัลกอริธึมกับข้อมูลที่จะใช้จริง เพื่อทำนายแต่ละคุณลักษณะของแผนประกันโดยอัลกอริธึมที่ได้เลือกใช้มีต่อไปนี้
 - อัลกอริธึมหลัก ได้แก่ Logistic, Decision Tree, Random Forest, Support Vector Machine, Naïve Bayes และ K Nearest Neighbors สำหรับการจัดกลุ่มแผนประกันตามลักษณะของผู้ใช้งานและลักษณะของแผนประกัน
- 4) ผลการทดสอบโปรแกรมแนะนำประกัน จะใช้การวัดผลแบบ Top K Precision ซึ่งหมายถึงถ้ามีรายการแผนประกัน n แผน แล้วแผนประกันที่ถูกต้อง หรือแผนประกันที่ผู้ใช้กดเลือกเป็นหนึ่งใน K ประกัน ถือว่าโปรแกรมสามารถทำนายได้ดี
- 5) ผลการพัฒนาเว็บแอปพลิเคชัน
- 6) ผลการเปรียบเทียบกับแอปพลิเคชันแนะนำประกันอื่นๆ พร้อมทั้งข้อได้เปรียบและข้อเสียเปรียบ

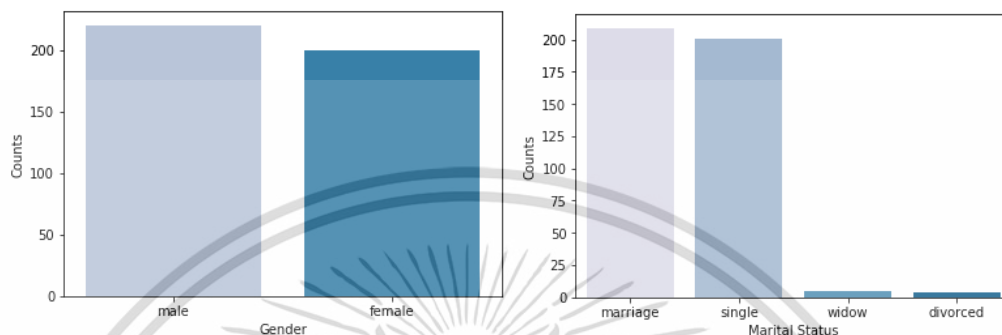
4.2. ผลการทำความสะอาดข้อมูลและวิเคราะห์ข้อมูลเบื้องต้น

ในหัวข้อนี้จะกล่าวถึงการทำความสะอาดข้อมูลและการวิเคราะห์ข้อมูลเบื้องต้น

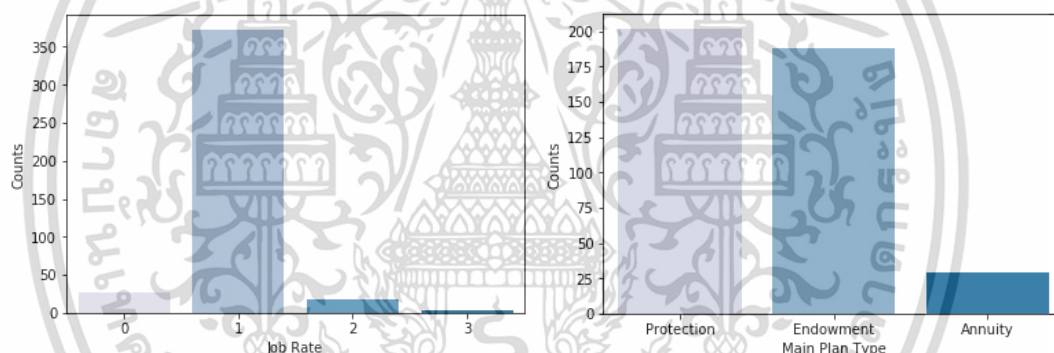
เริ่มจากการทำความสะอาดข้อมูลนั้น ได้ทำตั้งแต่การแจกแจงข้อมูล ลบข้อมูลที่มีข้อมูลบางฟิลด์สูญหายและไม่สามารถแก้ไขหรือต่อเติมได้ และลบข้อมูลผู้ซื้อประกันที่ไม่ได้ขายในปัจจุบัน หรือไม่มีข้อมูลของประกัน ไปจนถึงการทำ Standardization ให้ข้อมูลอยู่ในรูปแบบเดียวกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในส่วนของการวิเคราะห์ข้อมูลเบื้องต้นนี้จะเป็นการสำรวจข้อมูลผู้เอาประกันที่เก็บรวบรวมได้มาสำรวจและวิเคราะห์เบื้องต้นก่อนนำไปทดลองกับอัลกอริทึมในรูปแบบกราฟการกระจายและกราฟฮิสโทแกรมดังต่อไปนี้

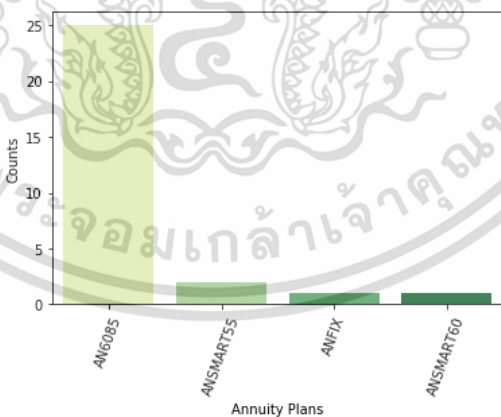
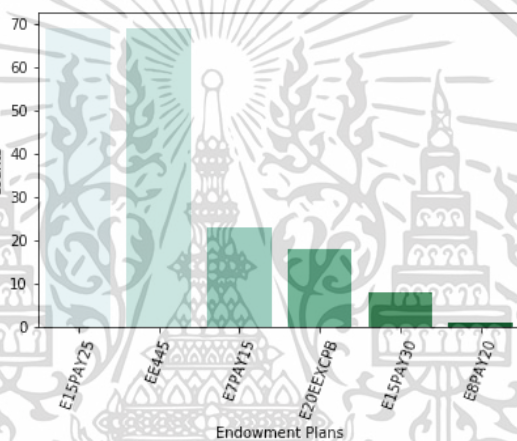
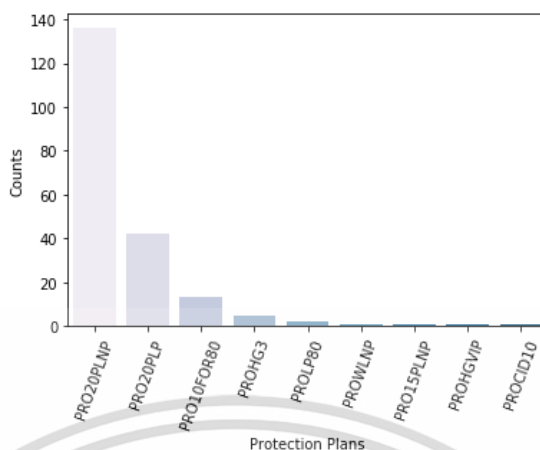


รูปที่ 4.1 กราฟฮิสโทแกรมแสดงเพศและสถานะของผู้เอาประกัน



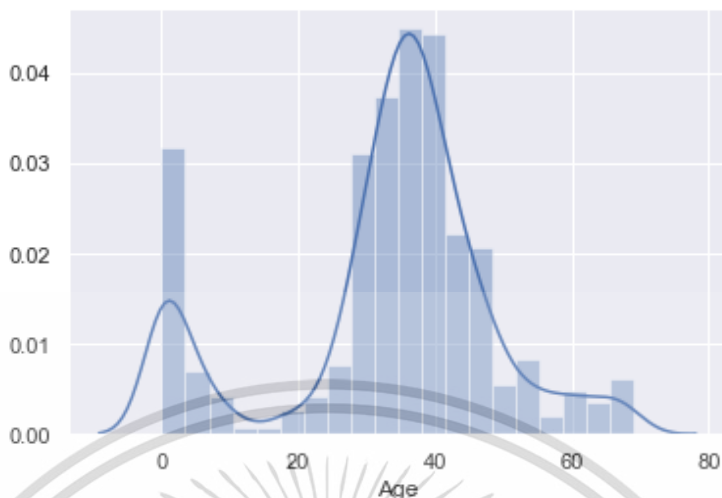
รูปที่ 4.2 กราฟฮิสโทแกรมแสดงชั้นอาชีพและประเภทประกันที่ผู้เอาประกันเลือกทำ

จากรูปที่ 4.1 และ 4.2 พบว่าผู้เอาประกันเพศชายและหญิงทำประกันเท่าๆกัน และมักมีสถานะโสดหรือสมรสแล้ว ส่วนใหญ่ผู้เอาประกันจะมีชั้นอาชีพระดับ 1 (กลุ่มอาชีพความเสี่ยงต่ำ เช่น พนักงานบริษัท เป็นต้น) ส่วนของประกันที่ส่วนใหญ่เลือกทำจะเป็นคุ้มครองตลอดชีพ สะสมทรัพย์ และบำนาญตามลำดับ



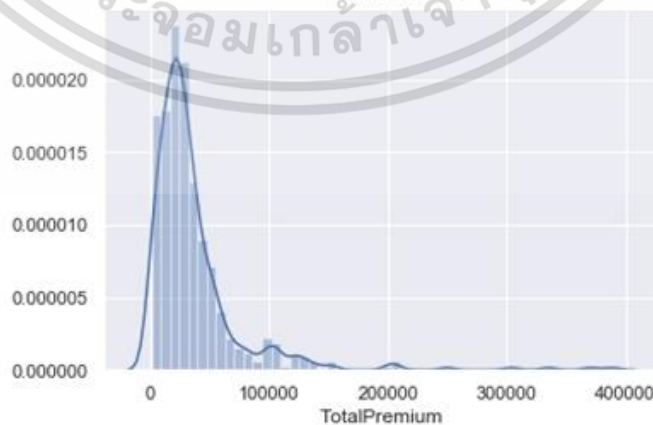
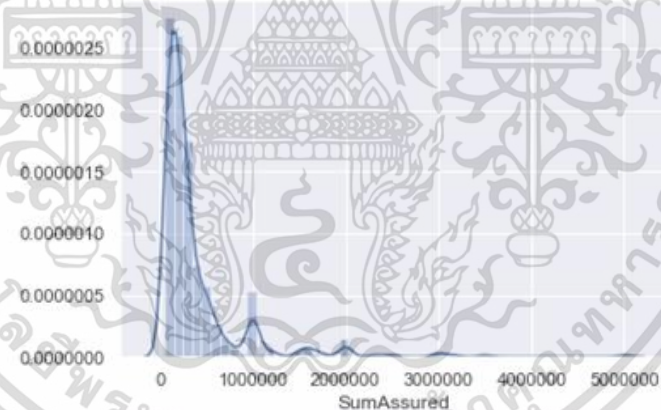
รูปที่ 4.3 กราฟแสดงจำนวนผู้ซื้อประกันตลอดชีพ สะสมทรัพย์ และบำนาญ

จากรูป 4.3 จะเห็นได้ว่าข้อมูลที่ได้นั้น ไม่มีค้อยสมมูลนัก โดยในส่วนของประกันตลอดชีพ แผนประกันที่ได้รับความนิยมคือ PRO20PLNP ในขณะที่ประกันแบบสะสมทรัพย์ แผนที่มีจำนวนผู้ซื้อมากที่สุดคือ EE445 และ AN6085 ในส่วนของประกันบำนาญ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรวิจัยเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 กราฟแสดงช่วงอายุของผู้เอาประกัน

กราฟการกระจายจากรูปที่ 4.4 แสดงให้เห็นถึงช่วงอายุของผู้เอาประกัน พบว่าผู้เอาประกันส่วนใหญ่มักทำประกันช่วงอายุ 25-50 ปี



รูปที่ 4.5 กราฟแสดงช่วงเงินทุนประกันและเบี้ยประกันที่ผู้เอาประกันต้องชำระ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กราฟการกระจายจากรูปที่ 4.5 แสดงให้เห็นถึงเงินทุนประกันที่ผู้เอาประกันเลือกเมื่อทำประกันและเบี้ยประกันที่ผู้เอาประกันต้องชำระ โดยเบี้ยประกันนี้รวมทั้งเบี้ยประกันของสัญญาหลักและสัญญาเพิ่มเติมโดยคำนวณจากเงินทุนประกัน อายุ เพศของผู้เอาประกันตามเงื่อนไขของแต่ละแผนประกันได้กำหนดไว้

4.3. ผลการทดลองอัลกอริธึมกับข้อมูลที่จะใช้จริง

คณะผู้จัดทำได้ทำการประเมินผลการทดลองอัลกอริธึมกับข้อมูลที่จะใช้จริง โดยแจกแจงตามฟีเจอร์ของแผนประกันและผลการทำนายการซื้อประกันสัญญาเพิ่มเติมได้ดังนี้

4.3.1. ผลการทดลองอัลกอริธึมที่ใช้ในการทำนายคุณลักษณะของประกันตลอดชีพ

คุณลักษณะของประกันตลอดชีพที่เลือกมาทำนายคือ ระยะเวลาชำระเบี้ย (ปี) เงินปันผล (มี หรือ ไม่มี) ระยะเวลาคุ้มครอง (ปี) โดยมีผลการทดลองดังตาราง 4.1 – 4.3

ตาราง 4.1 ตารางแสดงผลการทดลองอัลกอริธึมสำหรับระยะเวลาชำระเบี้ย (ประกันตลอดชีพ)

ระยะเวลาชำระเบี้ย				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	95.4	95.5	95.4	95.4
Random Forest	98.5	98.5	98.5	98.5
K-Nearest-Neighbors	90.8	91.1	90.8	90.9
Support Vector Machine	92.3	92.4	92.3	92.3
Logistic Regression	86.2	86.9	86.2	86.3
Naïve Bayes	49.2	76.7	49.2	44.3

จากตาราง 4.1 จะเห็นได้ว่าอัลกอริธึมที่เหมาะสมที่สุดคืออัลกอริธึม Random Forest

ตาราง 4.2 ตารางแสดงผลการทดลองอัลกอริธึมสำหรับเงินปันผล (ประกันตลอดชีพ)

เงินปันผล				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	92.3	92.2	92.3	92.2
Random Forest	98.5	98.5	98.5	98.5
K-Nearest-Neighbors	89.2	90.7	89.2	88.6
Support Vector Machine	87.7	88.4	87.7	87.1
Logistic Regression	86.2	86.4	86.2	86.2
Naïve Bayes	83.1	86.5	83.1	81

จากตาราง 4.2 จะเห็นได้ว่าอัลกอริธึมที่เหมาะสมที่สุดในการใช้ทำนายเงินปันผลของประกันแบบตลอดชีพคืออัลกอริธึม Random Forest

ตาราง 4.3 ตารางแสดงผลการทดลองอัลกอริธึมสำหรับระยะเวลาคุ้มครอง (ประกันตลอดชีพ)

ระยะเวลาคุ้มครอง				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	96.9	97.2	96.9	96.9
Random Forest	98.5	98.5	98.5	98.4
K-Nearest-Neighbors	90.8	91.1	90.8	90.9
Support Vector Machine	92.2	92.4	92.3	92.3
Logistic Regression	86.1	86.9	86.2	86.3
Naïve Bayes	49.2	79.7	49.2	44.3

จากตาราง 4.3 จะเห็นได้ว่าอัลกอริธึมที่เหมาะสมที่สุดในการใช้ทำนายระยะเวลาคุ้มครองของประกันแบบตลอดชีพคืออัลกอริธึม Random Forest

4.3.2. ผลการทดลองอัลกอริธึมที่ใช้ในการทำนายคุณลักษณะของประกันสะสมทรัพย์

คุณลักษณะของประกันตลอดชีพที่เลือกมาทำนายคือ ระยะเวลาชำระเบี้ย (ปี) ระยะเวลาคุ้มครอง (ปี) เงินคืนทั้งหมด(%) ความคุ้มครองเมื่อจบสัญญา(%) จำนวนเงินคืนทั้งหมด (ครั้ง) โดยมีผลการทดลองดังตาราง 4.4 – 4.8

ตาราง 4.4 ตารางแสดงผลการทดลองอัลกอริธึมสำหรับระยะเวลาชำระเบี้ย (ประกันสะสมทรัพย์)

ระยะเวลาชำระเบี้ย				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	74	74.4	74	74
Random Forest	81.1	82	81.1	81.1
K-Nearest-Neighbors	66.7	67.6	66.7	66.4
Support Vector Machine	63.7	64.3	63.7	63.8
Logistic Regression	63.7	63.8	63.7	62.7
Naïve Bayes	37.6	49.2	37.6	31

จากตาราง 4.4 จะเห็นได้ว่าอัลกอริธึมที่เหมาะสมที่สุดในการใช้ทำนายระยะเวลาชำระเบี้ยของประกันแบบสะสมทรัพย์คืออัลกอริธึม Random Forest

ตาราง 4.5 ตารางแสดงผลการทดลองอัลกอริธึมสำหรับระยะเวลาคุ้มครอง (ประกันสะสมทรัพย์)

ระยะเวลาคุ้มครอง				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	76.8	80.6	76.8	77.5
Random Forest	75.4	79.2	75.3	76.5
K-Nearest-Neighbors	65.2	68.4	65.2	65.9
Support Vector Machine	63.7	67.6	63.8	64.1
Logistic Regression	58	59.2	58	58
Naïve Bayes	53.6	58.3	53.6	50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตาราง 4.5 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการใช้ทำนายระยะเวลาคุ้มครองของประกันแบบสะสมทรัพย์คืออัลกอริทึม Decision Tree

ตาราง 4.6 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับเงินคืนทั้งหมด (ประกันสะสมทรัพย์)

เงินคืนทั้งหมด				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	74	78	74	74.6
Random Forest	73.9	77.8	74	74.1
K-Nearest-Neighbors	63.8	67.6	64	64.5
Support Vector Machine	63.8	67.6	63.8	64.1
Logistic Regression	58	59.2	58	58
Naïve Bayes	53.6	58.2	53.6	50

จากตาราง 4.6 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการใช้ทำนายเงินคืนทั้งหมดของประกันแบบสะสมทรัพย์คืออัลกอริทึม Decision Tree

ตาราง 4.7 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับความคุ้มครองเมื่อจบสัญญา (ประกันสะสมทรัพย์)

ความคุ้มครองเมื่อจบสัญญา				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	68.1	74.3	68.1	70.1
Random Forest	73.9	77.3	73.9	75.1
K-Nearest-Neighbors	66.7	69	66.7	66.7
Support Vector Machine	65.2	66.2	65.2	65
Logistic Regression	66.7	68	66.7	66.9
Naïve Bayes	43.4	54.9	43.5	39.7

จากตาราง 4.7 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการใช้ทำนายความคุ้มครองเมื่อจบสัญญาของประกันแบบสะสมทรัพย์คืออัลกอริทึม Random Forest

ตาราง 4.8 ตารางแสดงผลการทดลองอัลกอริทึมสำหรับจำนวนครั้งที่ได้เงินคืน (ประกันสะสมทรัพย์)

จำนวนครั้งที่ได้เงินคืน				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	78.2	82	78.2	77.1
Random Forest	77	80.1	76.8	78
K-Nearest-Neighbors	65.2	68.4	65.2	66
Support Vector Machine	64	67.6	63.8	64.1
Logistic Regression	58	59.2	58	57.9
Naïve Bayes	53.6	58.2	53.6	50

จากตาราง 4.8 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการใช้ทำนายจำนวนครั้งที่ได้เงินคืนของประกันแบบสะสมทรัพย์คืออัลกอริทึม Random Forest

4.3.3. ผลการทดลองอัลกอริทึมที่ใช้ในการทำนายการซื้อประกันบ้านอายุ

เนื่องจากว่าข้อมูลของผู้ทำประกันบ้านอายุมีน้อยเกินไป ทำให้ไม่สามารถนำมาทดลองกับแบบจำลองได้ จึงทำการกำหนดการคำนวณแบบพิเศษโดยพิจารณาจากแผนประกันที่มีคนซื้อมากที่สุด ดังนี้

- กำหนดคะแนนความใกล้เคียงประกันบ้านอายุจากค่าความสำคัญของประกันบ้านอายุเมื่อกรอกในหน้ากรอกข้อมูล จากนั้นนำไปหารด้วย 10 แล้วจึงคูณด้วย 40
- กรณีผู้ใช้เลือกประกันบ้านอายุอย่างเดียว ให้นำเสนอแผนประกันบ้านอายุตามความนิยมชื่อของข้อมูลดังนี้
 - 1) แผนประกัน AN6085
 - 2) แผนประกัน ANSMART55
 - 3) แผนประกัน ANSMART60
 - 4) แผนประกัน ANFIX
 - 5) แผนประกัน ANSMART65
- กรณีมีการเลือกประกันหลักอื่นร่วมด้วย คิดเป็นหากคะแนนความใกล้เคียงบ้านอายุที่ได้มีค่าน้อยกว่า 5 ให้นำเสนอแผนประกันบ้านอายุที่มีความนิยมมากที่สุดเพียงแผนเดียว หากคะแนนความใกล้เคียงบ้านอายุไม่น้อยกว่า 5 ให้นำเสนอ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แผนประกันที่มีความนิยมสูงสุด 2 อันดับ โดยกำหนดให้ทั้งสองแผนมีคะแนนความใกล้เคียงเท่ากัน

4.3.4. ผลการทดลองอัลกอริธึมที่ใช้ในการทำนายการซื้อประกันสัญญาเพิ่มเติม

สัญญาเพิ่มเติมที่ได้เลือกมาทำนายแบ่งเป็น สัญญาเพิ่มคุ้มครองโรคร้ายแรง ได้แก่ แผน AHC และ ECIR, สัญญาเพิ่มเติมคุ้มครองอุบัติเหตุ ได้แก่ แผน AI/RCC, สัญญาเพิ่มเติมชดเชยรายวัน ได้แก่ แผน HB และสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล ได้แก่ แผน HS และ HSPG โดยมีผลการทดลองดังตารางที่ 4.9-4.14

ตาราง 4.9 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง AHC

การเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง AHC				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	88.1	87.7	88.1	87.9
Random Forest	89.3	90.1	89.3	89.6
K-Nearest-Neighbors	88.1	89.3	88.1	88.5
Support Vector Machine	85.7	87.1	85.7	86.1
Logistic Regression	91.7	91.8	91.7	91.7
Naïve Bayes	39.3	83.5	39.3	37.1

จากตารางที่ 4.9 จะเห็นได้ว่าอัลกอริธึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง AHC คืออัลกอริธึม Logistic Regression

ตาราง 4.10 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง ECIR

การเลือกซื้อสัญญาเพิ่มเติมคุ้มครองโรคร้ายแรง ECIR				
อัลกอริธึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	95.2	95.5	95.2	95.3
Random Forest	97.6	97.8	97.6	97.6
K-Nearest-Neighbors	90.5	90.9	90.5	90.6
Support Vector Machine	96.4	96.9	96.4	96.5
Logistic Regression	95.2	95.5	95.2	95.3
Naïve Bayes	85.7	90.8	85.7	86.5

เอกสารนี้เป็นเอกสารสงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้ทำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.10 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมคุ้มครอง โรคร้ายแรง ECIR คืออัลกอริทึม Random Forest

ตาราง 4.11 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมคุ้มครองอุบัติเหตุ AI/RCC

การเลือกซื้อสัญญาเพิ่มเติมคุ้มครองอุบัติเหตุ AI/RCC				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	95.2	95.8	95.2	95.4
Random Forest	95.2	95.8	95.2	95.4
K-Nearest-Neighbors	96.4	96.7	96.4	96.5
Support Vector Machine	96.4	96.7	96.4	96.5
Logistic Regression	92.9	92.9	92.9	92.9
Naïve Bayes	25	90.6	25	27

จากตารางที่ 4.11 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมคุ้มครองอุบัติเหตุ AI/RCC คืออัลกอริทึม Support Vector Machine

ตาราง 4.12 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยรายวัน HB

การเลือกซื้อสัญญาเพิ่มเติมชดเชยรายวัน HB				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	90.5	93.8	90.5	91.2
Random Forest	90.5	93.8	90.5	91.2
K-Nearest-Neighbors	83.3	86.3	83.3	84.3
Support Vector Machine	85.7	87.4	85.7	86.3
Logistic Regression	83.3	87.4	83.3	84.6
Naïve Bayes	76.2	86.5	76.2	78.8

จากตารางที่ 4.12 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมชดเชยรายวัน HB คืออัลกอริทึม Random Forest

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.13 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HS

การเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HS				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	92.9	95.1	92.9	93.3
Random Forest	89.3	92.5	89.3	90.1
K-Nearest-Neighbors	85.7	88.8	85.7	86.7
Support Vector Machine	86.9	88.3	86.9	87.4
Logistic Regression	81	82.1	81	81.5
Naïve Bayes	21.4	63.4	21.4	20

จากตารางที่ 4.12 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HS คืออัลกอริทึม Decision Tree

ตาราง 4.14 ตารางแสดงผลการทดลองการเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HSPG

การเลือกซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HSPG				
อัลกอริทึม	Accuracy	Precision	Recall	F1 Score
Decision Tree	94	93.9	94	93.9
Random Forest	94	94.4	94	93.4
K-Nearest-Neighbors	89.3	88.4	89.3	88.6
Support Vector Machine	90.5	89.9	90.4	89.1
Logistic Regression	88.1	86.7	88.1	87.1
Naïve Bayes	79.8	91.6	79.8	82.6

จากตารางที่ 4.12 จะเห็นได้ว่าอัลกอริทึมที่เหมาะสมที่สุดในการทำนายการซื้อสัญญาเพิ่มเติมชดเชยค่ารักษาพยาบาล HSPG คืออัลกอริทึม Random Forest

4.3. ผลการทดสอบโปรแกรมแนะนำประกัน

เมื่อได้คุณลักษณะของประกันจากการทำนายของอัลกอริทึมในข้อ 4.4 แล้ว นำไปคำนวณเพื่อหาแผนประกันที่มีคุณลักษณะใกล้เคียงกับที่สุด จากนั้นทำการวัดผลด้วย Top K Position ได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.15 แสดงผลการทดสอบโปรแกรมด้วย Top-K Position

	ประกันตลอดชีพ	ประกันสะสมทรัพย์
จำนวนการแนะนำที่ผ่าน Top-K (K = 3)	95.4 %	75 %

4.4. ผลการพัฒนาเว็บแอปพลิเคชัน

ในส่วนนี้จะเป็นการทดสอบการเข้าใช้งานเว็บแอปพลิเคชันตามการใช้งานหลัก โดยเว็บแอปพลิเคชันดังกล่าวจะใช้ชื่อเว็บคือ Insurex ดังนี้

4.5.1. การกรอกข้อมูลเบื้องต้นและเป้าหมายการทำประกันเพื่อสร้างรายการแนะนำประกัน

ผู้ใช้งานจะเข้ามาสู่หน้าหลักของเว็บก่อน เมื่อต้องการใช้งานระบบแนะนำประกัน ให้กดปุ่ม 'ให้เราช่วยสิ' เพื่อเข้าสู่ขั้นตอนการกรอกข้อมูล ดังรูปที่ 4.6



รูปที่ 4.6 หน้าหลักของเว็บ Insurex

เมื่อเข้าสู่หน้ากรอกข้อมูลแล้ว ให้ทำการกรอกข้อมูลตามที่แสดงบนหน้าเว็บ โดยแบ่งเป็นสองส่วนคือ ข้อมูลพื้นฐาน ประกอบด้วย เพศ อายุ สถานะ ประวัติการสูบบุหรี่ งบประมาณประกัน ในส่วนนี้ต้องกรอกให้ครบถ้วน และส่วนเป้าหมายในการทำประกัน ส่วนนี้สามารถเลือกได้ตามความต้องการ (แต่ต้องเลือกอย่างน้อยหนึ่งข้อ) เมื่อกดเลือกเป้าหมายนั้น จะต้องกรอกค่าความสำคัญของเป้าหมายที่เลือกด้วย ดังรูปที่ 4.7 และ 4.8

เพศ

ชาย

หญิง

อายุ

23

สถานะ

โสด

สมรส

หย่าร้าง

หม้าย

เคยสูบบุหรี่หรือไม่?

เคย

ไม่เคย

งบเฉลี่ยประกัน (บาท/ปี)

1500

รูปที่ 4.7 หน้ากรอกข้อมูลส่วนตัว (ส่วนข้อมูลเบื้องต้น)

เป้าหมายในการทำประกัน

คุ้มครองชีวิต มีเงินให้คนข้างหลัง

คุ้มครองชีวิต มีเงินคืน

คุ้มครองชีวิต มีเงินใช้หลังเกษียณ

คุ้มครองโรคร้ายแรง

คุ้มครองอุบัติเหตุ

มีเงินชดเชยรายได้ระหว่างรักษาตัว

คุ้มครองเมื่อเจ็บป่วย มีเงินค่ารักษาพยาบาล

"คุ้มครองชีวิต มีเงินคืน"

จาก 1-10, คุณจะให้ความสำคัญด้านนี้เท่าไร?

10

"คุ้มครองอุบัติเหตุ"

จาก 1-10, คุณจะให้ความสำคัญด้านนี้เท่าไร?

7

รูปที่ 4.8 หน้ากรอกข้อมูลส่วนตัว (ส่วนเป้าหมายการทำประกัน)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.2. การดูรายการแผนประกันที่แนะนำ

เมื่อกรอกข้อมูลเรียบร้อยแล้ว สามารถกดปุ่ม ‘แนะนำประกัน’ เพื่อดูรายการแผนประกันที่ระบบแนะนำได้เลย ดังรูปที่ 4.9

"คัดกรองชีวิต มีเงินคืน"
จาก 1-10, คุณจะให้ความสำคัญด้านนี้เท่าไร?
10

"คัดกรองอุบัติเหตุ"
จาก 1-10, คุณจะให้ความสำคัญด้านนี้เท่าไร?
7

กลับ แนะนำประกัน

รูปที่ 4.9 กดปุ่มแนะนำประกันเพื่อเริ่มการแนะนำประกัน

เมื่อกดปุ่มแนะนำประกันแล้ว ระบบจะเข้าสู่กระบวนการนำข้อมูลที่ได้ไปคำนวณวิเคราะห์กับ โมเดลในระบบ แล้วจึงแสดงผลเป็นแผนประกันที่ระบบแนะนำตามข้อมูลที่ใช้กรอก และเป้าหมายที่ต้องการพร้อมทั้งแสดงรายละเอียดของแต่ละแผนประกัน จำนวน 5 แผน อีกทั้งยังคำนวณและแสดงผลทุนประกันที่สามารถทำได้และเบี้ยประกันที่ต้องชำระในแต่ละแผน ดังรูปที่ 4.10

ประกันแนะนำสำหรับคุณ



E5PAY10 AI/RCC

- กุณประกันอุบัติเหตุ 99000 บาท ชดเชยเมื่อเสียชีวิตหรือสูญเสียอวัยวะจากอุบัติเหตุ 100%
- ค่าคุ้มครองชีวิตด้วยกุณประกัน: 2000 บาท
- ค่าคุ้มครองชีวิตพร้อมรับเงินคืน 10 ปี
- รับเงินคืนสูงสุด 500%
- ค่าคุ้มครองสูงสุดสูงสุด 500%
- เบี้ยประกันรวม 1500 บาท/ปี
- ระยะเวลาชำระเบี้ย 5 ปี



E20EEXCPB AI/RCC

- กุณประกันอุบัติเหตุ 99000 บาท ชดเชยเมื่อเสียชีวิตหรือสูญเสียอวัยวะจากอุบัติเหตุ 100%

รูปที่ 4.10 หน้าแสดงผลการแนะนำประกัน

4.5. ผลการเปรียบเทียบกับแอปพลิเคชันแนะนำประกันอื่นๆ

ในส่วนนี้จะเป็นการเปรียบเทียบคุณลักษณะระหว่างเว็บแอปพลิเคชัน Insurex กับเว็บแอปพลิเคชันแนะนำประกันอื่นๆ ตามที่ได้อ้างอิงในหัวข้อที่ 2.2 แต่เนื่องจากเว็บไซต์ Roojai.com และ EasyCompare.com เป็นเว็บไซต์แนะนำประกันประเภทรถยนต์ จึงไม่นำสองเว็บไซต์นี้มาเปรียบเทียบ ดังตารางที่ 4.10

ตาราง 4.16 ตารางแสดงการเปรียบเทียบเว็บแอปพลิเคชัน Insurex กับเว็บแอปพลิเคชันอื่น

	Insurex	iTax
สามารถแนะนำแผนประกัน	สามารถทำได้	สามารถทำได้
จุดประสงค์ของการแนะนำประกัน	เน้นเหมาะสมกับลักษณะผู้ใช้	เน้นเพื่อลดหย่อนภาษี
สามารถเลือกชนิดประกันได้	สามารถทำได้	สามารถทำได้
จำนวนการเลือกเป้าหมายการทำประกัน	เลือกได้มากกว่า 1 แบบ	เลือกได้แบบเดียว
นำเสนอประกันเสริมร่วมกับประกันหลัก	แนะนำประกันเสริมร่วมกับประกันหลัก	แนะนำได้อย่างใดอย่างหนึ่ง
ประกันเสริมที่แนะนำ	โรคร้ายแรง อุบัติเหตุ ชดเชยรายวัน ค่ารักษาพยาบาล	โรคร้ายแรงและรักษาพยาบาลเท่านั้น
แสดงข้อมูลรายละเอียดแผนประกันในหน้าแนะนำ	มีการแสดงข้อมูลรายละเอียด	ไม่ได้จัดทำฟีเจอร์นี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

บทสรุปและข้อเสนอแนะ

5.1. บทสรุป

แรงจูงใจที่ทำให้เกิดโครงการนี้ขึ้น เกิดจากการที่คณะผู้จัดทำได้สังเกตเห็นถึงปัญหาที่เกิดจากการทำประกัน คือ ผู้เอาประกันเมื่อทำประกันแล้วไม่ได้รับผลประโยชน์ที่เพียงพอ ทั้งจากการให้ข้อมูลส่วนตัวที่ไม่เป็นความจริงแก่ตัวแทน หรือเกิดจากตัวแทนที่ขายประกันโดยไม่คำนึงถึงความเหมาะสมและความต้องการของผู้เอาประกัน ดังนั้นปัญญาประดิษฐ์จึงได้ทำการศึกษาประกันและพัฒนาาระบบแนะนำประกันในชื่อ Insurex ซึ่งเป็นเว็บแอปพลิเคชันที่สามารถแนะนำประกันชีวิตที่มีการใช้อัลกอริทึมจาก Machine Learning และ Recommender System มาช่วยในการประมวลผลทำนายและแนะนำประกันชีวิตที่เหมาะสมและตรงความต้องการของผู้ที่สนใจจะทำประกัน ผู้เอาประกัน หรือผู้ใช้งานเว็บแอปพลิเคชันทั่วไปที่มีความสนใจในเรื่องประกัน เพียงกรอกข้อมูลส่วนตัวคือ เพศ อายุ สถานะ พฤติกรรมการสูบบุหรี่ งบประมาณที่สามารถชำระได้ต่อปี และระบุเป้าหมายพร้อมทั้งให้คะแนนความสำคัญของเป้าหมายที่เลือก โดยประกันที่แนะนำดังกล่าวครอบคลุมทั้งประกันประเภทคุ้มครองตลอดชีพ ประเภทสะสมทรัพย์ และประเภทบำนาญ เมื่อกรอกข้อมูลดังกล่าวเสร็จแล้ว ระบบจะนำข้อมูลไปคำนวณพร้อมทั้งวิเคราะห์ในแบบจำลองที่ทางคณะผู้จัดทำได้ทำการทดลองแล้วว่าสามารถทำนายแผนประกันที่ใกล้เคียงกับลักษณะของผู้ใช้จากข้อมูลที่ให้มา จากนั้นระบบจะนำเสนอผลลัพธ์ผ่านเว็บแอปพลิเคชันเป็นรายการแผนประกันแนะนำจำนวนทั้งสิ้น 5 แผนที่ใกล้เคียงกับผู้ใช้งานพร้อมทั้งแสดงรายละเอียด จำนวนทุนประกัน และเบี้ยประกันของชุดประกันแต่ละแผน ตามเป้าหมายที่ต้องการที่ได้เลือกไว้ และมั่นใจว่าผู้ใช้งานจะได้รับผลประโยชน์สูงสุดจากประกันที่ถูกแนะนำเหล่านั้น เว็บแอปพลิเคชัน Insurex นี้จะช่วยอำนวยความสะดวกแก่ผู้ใช้งานที่สนใจในด้านประกัน แก้ไขปัญหาต่างๆระหว่างการทำประกันกับตัวแทน ก็สามารถเลือกทำประกันที่มีความเหมาะสมกับตนเองได้

5.2. ปัญหา อุปสรรค และแนวทางการแก้ไข

- 1) มีข้อจำกัดด้านการเก็บข้อมูล เนื่องจากข้อมูลผู้เอาประกันเป็นความลับทางธุรกิจ จึงได้เพียงข้อมูลเป็นส่วนน้อยที่ทางบริษัทเปิดเผยเพื่อให้ทำวิจัย
- 2) ข้อมูลที่ได้มาจากบริษัทเพียงบริษัทเดียว ทำให้ข้อมูลไม่หลากหลาย ควรรับข้อมูลจากหลายๆบริษัท
- 3) ข้อมูลของประกันมีลักษณะไม่เหมือนกันตามแต่ละประเภท จึงต้องทำความสะอาดข้อมูล จัดเก็บข้อมูลแยกตามประเภทของประกัน รวมไปถึงการฝึกสอนข้อมูลก็ต้องใช้อัลกอริทึมแตกต่างกันตามประเภท ทำให้ต้องใช้เวลาไปกับการปรับปรุงข้อมูลเป็นอย่างมาก เพื่อให้ได้ข้อมูลที่เหมาะสมกับอัลกอริทึมแต่ละชนิด
- 4) ข้อมูลที่นำมาทดลองมีจำนวนน้อย ทำให้มี Bias ของข้อมูลสูง ส่งผลให้ผลลัพธ์การฝึกสอนแบบจำลองมักจะเกิดการ Overfitted ต้องหาวิธีการคำนวณอื่น จึงจำเป็นต้องใช้ข้อมูลเพิ่มอีกมาก เพื่อพัฒนาความแม่นยำ
- 5) การวัดผลจริงกับผู้ใช้ในวงกว้างเป็นเรื่องท้าทาย ทำให้ต้องอาศัยความร่วมมือกับบริษัทประกันในอนาคต

5.3. แนวทางการพัฒนาต่อ

- 1) พัฒนาต่อยอดให้รองรับแผนประกันจากบริษัทอื่น เพื่อการทำตัวเลือกเปรียบเทียบความคุ้มค่าของแผนประกันชนิดเดียวกัน รวมไปถึงการอัปเดตแผนประกันใหม่ๆ ให้สอดคล้องกับปัจจุบัน
- 2) เพิ่มเดิมการให้ความรู้ด้านประกันโดยย่อในหน้าเว็บหลัก เพื่อการตัดสินใจที่ดีขึ้นของผู้ใช้
- 3) พัฒนาระบบเข้ากับ Search Engine Optimization เพื่อให้เว็บแอปพลิเคชันสามารถเข้าถึงได้จากการค้นหาผ่านระบบ Search Engine
- 4) พัฒนาเพิ่มเติมให้มีตัวเลือกการติดต่อกับบริษัทหรือตัวแทนกรณีสนใจซื้อแผนประกันที่เว็บได้แนะนำมาให้
- 5) พัฒนาให้หน้าแสดงรายการแนะนำประกัน สามารถเลือกตัวเลือกการเรียงลำดับได้ เช่น เรียงลำดับตามผลประโยชน์ที่ได้รับ ตามความคุ้มค่า ตามจำนวนเบี้ยประกัน เป็นต้น

บรรณานุกรม

แอดมินเพจ ออมให้เงินโต แบบเข้าใจง่ายๆ (นามแฝง). 2562. **ออมให้เงินโต**. กรุงเทพฯ : พราว.

Francesco Ricci, Lior Rokart, Bracha Shapira, Paul B. Kantor. 2010. **Recommender Systems Handbook**. New York : Springer.

Assad Abbas, Kashif Bilal, Limin Zhang, Samee U. Khan. 2014. "A cloud based health insurance plan recommendation system." **Future Generation Computer Systems**. 43(10) : 99-109.

David Zubriczky. 2016. "Recommender Systems meet Finance: A literature review." **2nd International Workshop on Personalization and Recommender Systems in Financial Services**. Bari : ResearchGate.

Akshay Hinduja and Manju Pandey. 2017. "Multicriteria Recommender System for Life Insurance Plans based on Utility Theory." **Indian Journal of Science and Technology**. 10(14) : 1-8.

Baptiste Rocca. 2019. **Introduction to recommender systems**. [Online].

Available : <https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada>

Karan Sivarat. 2019. [บันทึกความเข้าใจ] ว่าด้วยเรื่อง Cross Origin Resource Sharing-CORS. [Online].

Available : <https://siamchamnankit.co.th/บันทึกความเข้าใจ-ว่าด้วยเรื่อง-cors-cross-origin-resource-sharing-7751b0c5d4ff>

Miguel Grinberg. 2013. **Designing a RESTful API with Python and Flask**. [Online].

Available : <https://blog.miguelgrinberg.com/post/designing-a-restful-api-with-python-and-flask>

Miguel Grinburg. 2020. **How To Create a React + Flask Project**. [Online]

Available : <https://blog.miguelgrinberg.com/post/how-to-create-a-react--flask-project>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Mr.P L (นามแฝง). 2019. **Data Preprocessing** นั้นสำคัญอย่างไร? แล้วจะทำเมื่อไหร่?. [Online].

Available : <https://medium.com/mmp-li/data-preprocessing-นั้นสำคัญอย่างไร-แล้วจะทำเมื่อไหร่-69306f880e68>

Shuyu Luo. 2018. **Introduction to Recommender System**. [Online].

Available : <https://towardsdatascience.com/intro-to-recommender-system-collaborative-filtering-64a238194a26>

Will Koehrsen. 2018. **A Complete Machine Learning Project Walk-Through in Python: Part One**. [Online].

Available : <https://towardsdatascience.com/a-complete-machine-learning-walk-through-in-python-part-one-c62152f39420>

Zike Qin. 2018. **Full-Stack Beginner's Guide (Flask+React+SQLAlchemy)**. [Online].

Available: <http://zike.io/posts/Full-Stack-Beginner-s-Guide-Flask-React-SQLAlchemy/>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

ตัวอย่างข้อมูลที่เก็บในฐานข้อมูล

ในส่วนนี้จะกล่าวถึงข้อมูลที่ใช้ในการทำโครงการที่เก็บในฐานข้อมูล ดังตัวอย่างต่อไปนี้

User Info

ใช้สำหรับเก็บข้อมูลที่รับจากหน้ากรอกข้อมูล ประกอบด้วยคอลัมน์ต่อไปนี้ดังรูปที่ ก.1

userid คือ รหัสตัวแทนของผู้ใช้ในการกรอกข้อมูลแต่ละครั้ง

gender คือ เพศ แบ่งเป็น ชาย (male) และ หญิง (female)

age คือ อายุ

maritalStatus คือ สถานะ แบ่งเป็น โสด (single), สมรส (marriage), หย่าร้าง (divorced) และ หม้าย (widowed)

smoke คือ พฤติกรรมการสูบบุหรี่ แบ่งเป็น เคย (yes) และ ไม่เคย (no)

jobRate คือ ชั้นอาชีพ ในที่นี้ให้มีค่าตั้งต้นเป็นระดับที่ 1

premiumBudget คือ งบประมาณที่สามารถชำระได้ต่อปี

{isWholeLife, isSaving, isAnnuity, isCriticalIllness, isAccident, isIncome, isHospital} คือ การเลือกเป้าหมายการรับประกัน เป็นบูลีน หากถูกเลือกจะมีค่าเป็น 1 ไม่ถูกเลือกมีค่าเป็น 0

{wholeLifeScore, savingScore, annuityScore, criticalIllnessScore, accidentScore, incomeScore, hospitalScore} คือ คะแนนความสำคัญของแต่ละเป้าหมาย หากมีการเลือกเป้าหมาย คะแนนจะมีค่าตั้งแต่ 1-10 หากไม่มีการถูกเลือกจะมีค่าเป็น 0

userid	gender	age	maritalStatus	smoke	jobRate	premiumBudget	isWholeLife	isSaving	isAnnuity	isCriticalIllness	isAccident	isIncome	isHospital
30	male	10	single	no	1	1000	1	0	0	0	0	0	0
31	female	60	single	yes	1	45000	1	1	0	0	0	0	0
32	female	60	single	yes	1	45000	1	1	0	0	0	0	0
33	female	13	single	no	1	1000	1	0	0	0	0	0	0
34	male	34	marriage	yes	1	20000	1	0	0	0	1	0	0

รูปที่ ก.1 User Info Table

Recommended Plan

ใช้สำหรับเก็บข้อมูลประวัติการแนะนำแผนประกัน โดยผูกกับ userid ของตาราง User Info ประกอบด้วยคอลัมน์ต่อไปนี้ ดังรูปที่ ก.2

id คือ รหัสตัวแทนของการแนะนำแผนประกันแต่ละครั้ง

userid คือ รหัสตัวแทนของผู้ใช้ในการกรอกแต่ละครั้ง เป็น foreign key ที่เชื่อมกับคอลัมน์

userid จากตาราง User Info

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

rec_plan_1-5 คือ แผนประกันที่ระบบแนะนำ เป็นข้อมูลในรูปแบบ JSON ดังตัวอย่างต่อไปนี้

```
{
  "coverageDetail": [
    "คุ้มครองชีวิตพร้อมรับเงินคืน 10 ปี",
    "รับเงินคืนสูงสุด 500%",
    "คุ้มครองสูงสุดสูงสุด 500%"
  ],
  "keyString": "key-E5PAY10",
  "other": "",
  "plan": [
    "E5PAY10"
  ],
  "premiumPaymentPeriod": "5 ปี",
  "title": "แพ็คเกจประกันที่ 1",
  "totalPremium": "1500 บาท"
},
{
  "coverageDetail": [
    "คุ้มครองชีวิตพร้อมรับเงินคืน 20 ปี",
    "รับเงินคืนสูงสุด 330%",
    "คุ้มครองสูงสุดสูงสุด 300%"
  ],
  "keyString": "key-E20EEXCPB",
  "other": "",
  "plan": [
    "E20EEXCPB"
  ],
  "premiumPaymentPeriod": "20 ปี",
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

"title": "แพ็คเกจประกันที่ 2",
"totalPremium": "1500 บาท"
},
{
"coverageDetail": [
"คุ้มครองชีวิตพร้อมรับเงินคืน 15 ปี",
"รับเงินคืนสูงสุด 203%",
"คุ้มครองสูงสุดสูงสุด 200%"
],
"keyString": "key-E7PAY15",
"other": "",
"plan": [
"E7PAY15"
],
"premiumPaymentPeriod": "7 ปี",
"title": "แพ็คเกจประกันที่ 3",
"totalPremium": "1500 บาท"
},
{
"coverageDetail": [
"คุ้มครองชีวิตพร้อมรับเงินคืน 20 ปี",
"รับเงินคืนสูงสุด 188%",
"คุ้มครองสูงสุดสูงสุด 154%"
],
"keyString": "key-E8PAY20",
"other": "",
"plan": [
"E8PAY20"
],
"premiumPaymentPeriod": "8 ปี",

```

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

"totalPremium": "1500 บาท"
},
{
"coverageDetail": [
"คุ้มครองชีวิตพร้อมรับเงินคืน 42 ปี",
"รับเงินคืนสูงสุด 445%",
"คุ้มครองสูงสุดสูงสุด 200%"
],
"keyString": "key-E445",
"other": "",
"plan": [
"E445"
],
"premiumPaymentPeriod": "21 ปี",
"title": "แพ็คเกจประกันที่ 5",
"totalPremium": "1500 บาท"
}

```

id	userid	rec_plan_1	rec_plan_2	rec_plan_3	rec_plan_4	rec_plan_5
1	18	{plan: "ANSMART55",...	{plan: "PROWLNP",...	{plan: "PROLP80",...	{plan: "PROLP70",...	{plan: "ANSMART60",...
2	21	{plan: "ANSMART55",...	{plan: "ANSMART6",...	{plan: "PROWLNP",...	{plan: "PROLP80",...	{plan: "PROLP70",...
3	22	{plan: "PROWLNP",...	{plan: "PROLP80",...	{plan: "PROLP70",...	{plan: "E5PAY10",...	{plan: "E7PAY15", "o...
4	23	{plan: "PROWLNP",...	{plan: "PROLP80",...	{plan: "PROLP70",...	{plan: "E5PAY10",...	{plan: "E7PAY15", "o...
5	24	{plan: "PROWLNP",...	{plan: "PROLP80",...	{plan: "PROLP70",...	{plan: "E5PAY10",...	{plan: "E7PAY15", "o...

รูปที่ ก.2 Recommended Plan Table

Protection

ใช้สำหรับเก็บข้อมูลแผนประกันประเภทคุ้มครองตลอดชีพ ดังรูปที่ ก.3 ประกอบด้วยคอลัมน์ต่อไปนี้

planid คือ รหัสตัวแทนแผนประกัน

plan_name คือ ชื่อแผนประกัน

payment_period คือ ระยะเวลาชำระเบี้ย

diviend คือ มีเงินปันผล ถ้าไม่มีเงินปันผลจะแทนด้วย 0 ถ้ามีจะแทนด้วย 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

min_sumassured คือ จำนวนเงินทุนประกันขั้นต่ำ หากไม่มีการกำหนดขั้นต่ำจะแทนด้วย -1
max_sumassured คือ จำนวนเงินทุนประกันสูงสุด หากไม่มีการกำหนดขั้นสูงสุดจะแทนด้วย -

1

min_coverage_age คือ อายุขั้นต่ำที่สามารถทำประกันได้

max_coverage_age คือ อายุสูงสุดที่สามารถทำประกันได้

coverage_period คือ คุ้มครองถึงอายุที่กำหนด

planid	plan_name	payment_period	dividend	min_sumassured	max_sumassured	min_coverage_age	max_coverage_age	coverage_period
101	PROLP70	70	0	500000	-1	18	60	70
102	PROLP80	80	0	500000	-1	18	60	80
103	PROCID10	10	0	200000	4999999	0	65	99
104	PROCID20	20	0	200000	4999999	0	65	99
105	PROCID10P	10	0	5000000	15000000	0	65	99

รูปที่ ก.3 Protection table

Endowment

ใช้สำหรับเก็บข้อมูลแผนประกันประเภทสะสมทรัพย์ ดังรูปที่ ก.4 ประกอบด้วยคอลัมน์
ต่อไปนี้

planid คือ รหัสตัวแทนแผนประกัน

plan_name คือ ชื่อแผนประกัน

payment_period คือ ระยะเวลาชำระเบี้ย

coverage_period คือ ระยะเวลาการคุ้มครอง

total_cashback คือ ได้รับเงินคืนรวมทั้งสิ้น

min_coverage_age คือ อายุขั้นต่ำที่สามารถทำประกันได้

max_coverage_age คือ อายุสูงสุดที่สามารถทำประกันได้

coverage_end คือ ความคุ้มครองสูงสุด

min_sumassured คือ จำนวนเงินทุนประกันขั้นต่ำ หากไม่มีการกำหนดขั้นต่ำจะแทนด้วย -1

max_sumassured คือ จำนวนเงินทุนประกันสูงสุด หากไม่มีการกำหนดขั้นสูงสุดจะแทนด้วย -

1

cashback_times คือ จำนวนครั้งในการได้รับเงินคืน

planid	plan_name	payment_period	coverage_period	total_cashback	min_coverage_age	max_coverage_age	coverage_end	min_sum_assured	max_sum_assured	cashback_time
201	ESPAY10	5	10	5	16	60	5	20000	1000000	9
202	E7PAY15	7	15	2.03	0	70	2	100000	-1	8
203	E8PAY20	8	20	1.88	0	70	1.54	100000	-1	10
204	E15PAY25	15	25	1.44	0	70	1.35	100000	-1	24
205	E15PAY30	15	30	1.78	0	70	1.5	100000	-1	29
206	E20EEXCPB	20	20	3.3	0	70	3	100000	-1	5
207	E445	21	42	4.45	0	57	2	100000	-1	28

รูปที่ ก.4 Endowment Table

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Annuity

ใช้สำหรับเก็บข้อมูลแผนประกันประเภทบำนาญ ดังรูปที่ ก.5 ประกอบด้วยคอลัมน์ต่อไปนี้

planid คือ รหัสตัวแทนแผนประกัน

plan_name คือ ชื่อแผนประกัน

pension_age_start คือ อายุเริ่มต้นรับเงินบำนาญ

pension_age_end คือ อายุสิ้นสุดการรับเงินบำนาญ

pension_rate คือ อัตราเงินบำนาญที่ได้รับต่อปีเทียบกับเงินทุนประกัน

coverage_until_age คือ คຸ້ມครองจนถึงอายุ

pay_until_age คือ ชำระเบี้ยประกันถึงอายุ

min_sumassured คือ จำนวนเงินทุนประกันขั้นต่ำ หากไม่มีการกำหนดขั้นต่ำจะแทนด้วย -1

min_coverage_age คือ อายุขั้นต่ำที่สามารถทำประกันได้

max_coverage_age คือ อายุสูงสุดที่สามารถทำประกันได้

planid	plan_name	pension_age_start	pension_age_end	pension_rate	coverage_until_age	pay_until_age	min_sumassured	min_coverage_age	max_coverage_age
301	AN6085	60	85	0.1	60	60	100000	30	55
302	ANFIX	60	85	0.0625	60	60	200000	20	55
303	ANSMART55	55	90	0.15	55	55	100000	20	50
304	ANSMART60	60	90	0.15	60	60	100000	20	50
305	ANSMART65	65	90	0.15	65	65	100000	20	50

รูปที่ ก.5 Annuity Table

ProPackage

ใช้สำหรับเก็บข้อมูลแพ็คเกจประกันคุ้มครองตลอดชีพพร้อมสัญญาเพิ่มเติม ดังรูปที่ ก.6 ประกอบด้วยคอลัมน์ต่อไปนี้

planid คือ รหัสตัวแทนแผนประกัน

plan_name คือ ชื่อแผนประกัน

min_coverage_age คือ อายุขั้นต่ำที่สามารถทำประกันได้

max_coverage_age คือ อายุสูงสุดที่สามารถทำประกันได้

{PROWLNP, PRO10FOR80, PRO20PLP, PROLC} คือ ประกันคุ้มครองตลอดชีพที่ถูกจัดเป็นชุดแผนประกัน หากในชุดนั้นมีประกันจะมีข้อมูลเป็นเงินทุนประกันที่แน่นอน ไม่มีเป็น -1

{AHC, ADB/RCC, ADD/RCC, WP, ECIR, H&S, HB} คือ สัญญาเพิ่มเติมที่ถูกจัดเป็นชุดแผนประกัน หากในชุดนั้นมีประกันจะมีข้อมูลเป็นเงินทุนประกันที่แน่นอน ไม่มีเป็น -1

plan_id	plan_name	min_coverage_age	max_coverage_age	AHC	ADB/RCC	ADD/RCC	PROWLNP	PRO10FOR80	PRO20PLP	PROLC	WP	ECIR	H&S	HB
501	PROHG1	16	55	100000	100000	-1	75000	-1	-1	-1	0	-1	1000	400
502	PROHG2	16	55	300000	300000	-1	75000	-1	-1	-1	0	-1	2200	600
503	PROHG3	16	55	400000	400000	-1	75000	-1	-1	-1	0	-1	2800	800
504	PROHGVIP	16	55	700000	1000000	-1	75000	-1	-1	-1	0	-1	4000	1000
505	PROHGS1	16	55	100000	100000	-1	100000	-1	-1	-1	0	-1	1000	400

รูปที่ ก.6 ProPackage Table

Rider

ใช้สำหรับเก็บข้อมูลแผนประกันประเภทสัญญาเพิ่มเติม ดังรูปที่ ก.7 ประกอบด้วยคอลัมน์ต่อไปนี้

planid คือ รหัสตัวแทนแผนประกัน

plan_name คือ ชื่อแผนประกัน

plan_type คือ ประเภทของสัญญาเพิ่มเติม

coverage_type คือ ประเภทของความคุ้มครอง ได้แก่ คุ้มครองโรคร้ายแรง (4), ชดเชยรายวัน และค่ารักษาพยาบาล (5) และ คุ้มครองอุบัติเหตุ (6)

planid	plan_name	plan_type	coverage_type
401	AHC	1	4
402	CIP	1	4
403	CIPG	1	4
404	CIC	1	4
405	CIT	2	4

รูปที่ ก.7 Rider Table

Cashback for Endowment

ใช้สำหรับเก็บข้อมูลเงินคืนของประกันประเภทสะสมทรัพย์ ดังรูปที่ ก.8 ประกอบด้วยคอลัมน์ต่อไปนี้

no คือ หมายเลขแถว

plan_name คือ ชื่อแผนประกัน

policy_year คือ ปีกรมธรรม์ที่

cashback_rate คือ อัตราเงินคืนระหว่างสัญญาต่อหนึ่งปีกรมธรรม์

coverage_rate คือ อัตราความคุ้มครองในแต่ละปี

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

no	plan_name	policy_year	cashback_rate	coverage_rate
1	ESPAY10	1	0	1
2	ESPAY10	2	0.04	2
3	ESPAY10	3	0.04	3
4	ESPAY10	4	0.04	4
5	ESPAY10	5	0.04	5
6	ESPAY10	6	0.04	5
7	ESPAY10	7	0.04	5
8	ESPAY10	8	0.04	5
9	ESPAY10	9	0.04	5
10	ESPAY10	10	4.68	5

รูปที่ ก.8 Cashback for En Table

Critical Illness

ใช้สำหรับเก็บข้อมูลโรคร้ายแรง ดังรูปที่ ก.9 ประกอบด้วยคอลัมน์ต่อไปนี้

deceased_id คือ รหัสโรคร้ายแรง

name_th คือ ชื่อโรคร้ายแรง

group คือ กลุ่มของโรคร้ายแรง

severity คือ ระดับความรุนแรง

{AHC, CIP, CIPG, ECIR, PBCI, CIC, CIT} คือ สัญญาเพิ่มเติมที่ครอบคลุมโรคดังกล่าว หากครอบคลุมจะแทนด้วย 1 หากไม่ครอบคลุมแทนด้วย 0

deceased_id	name_th	group	severity	AHC	CIP	CIPG	ECIR	PBCI	CIC	CIT
111	โรคมะเร็งระยะไม่ลุกลาม	1	1	0	0	1	0	0	1	1
112	การผ่าตัดเนื้องอกต่อมไธสมงออก	1	1	0	0	1	0	0	1	1
121	โรคหลอดเลือดหัวใจตีบที่รักษาด้วยการสวนเลือดหัวใจ	2	1	0	0	1	0	0	1	1
122	การผ่าตัดลอกเยื่อหุ้มหัวใจ	2	1	0	0	1	0	0	1	1
123	การผ่าตัดโรคเส้นหัวใจด้วยการสวนหลอดเลือด	2	1	0	0	1	0	0	1	1
124	การรักษาโรคหลอดเลือดแดงใหญ่เออร์ตาโดยวิธีใส่	2	1	0	0	1	0	0	1	1
125	การใส่เครื่องกรองลิ่มเลือดในหลอดเลือดดำใหญ่	2	1	0	0	1	0	0	1	1

รูปที่ ก.9 Critical Illness Table

JobRate

ใช้สำหรับเก็บข้อมูลอาชีพในแต่ละชั้น ดังรูปที่ ก.10 ประกอบด้วยคอลัมน์ต่อไปนี้

job_id คือ รหัสอาชีพ

thainame คือ ชื่ออาชีพภาษาไทย

engname คือ ชื่ออาชีพภาษาอังกฤษ

level คือ ระดับชั้นอาชีพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

job_id	thainame	engname	level
101	ผู้บริหารระดับสูง	CEO	1
102	เจ้าของกิจการขนาดใหญ่	Large Business Owner	1
103	แพทย์	Docter	1
104	พยาบาล	Nurse	1
105	นุรุษพยาบาล	Male Nurse	1

รูปที่ ก.10 JobRate Table



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้