

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์ สำหรับการเสนอสินเชื่อส่วนบุคคล

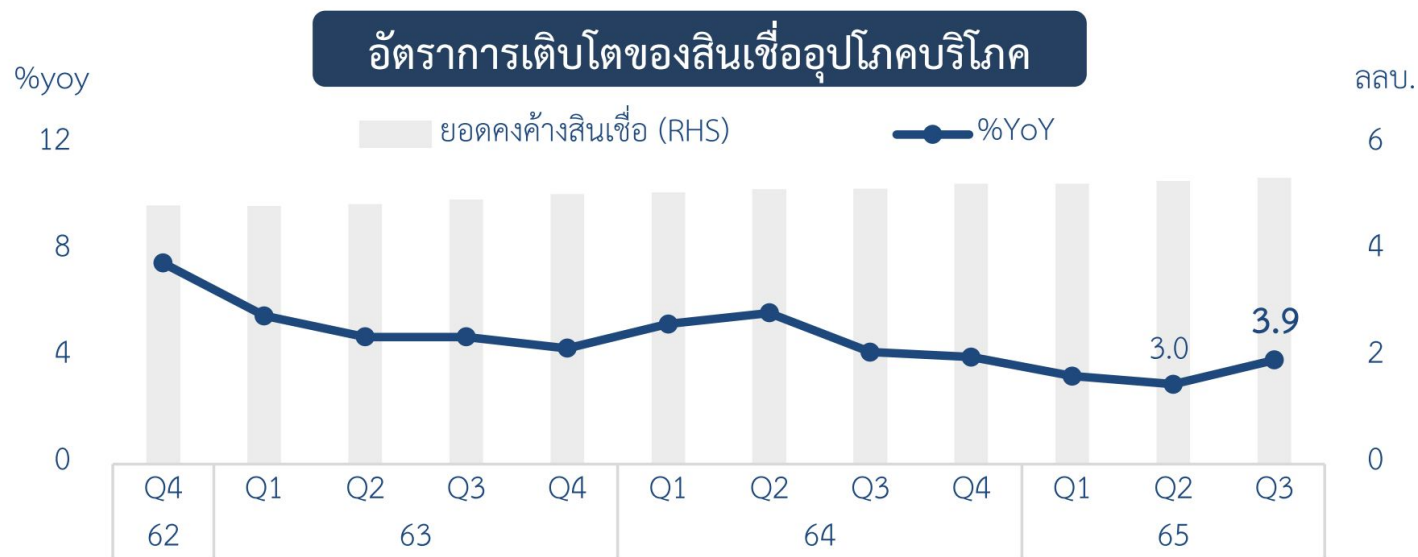
Customer Segmentation and Association Rules
for Personal Loan Offerings.



อาจารย์ที่ปรึกษา ผศ.ดร.ศติประภา หิริโอตปิ
นำเสนอโดย นางสาวณัฐริดา ลากรนชัย รหัสนักศึกษา 620710405

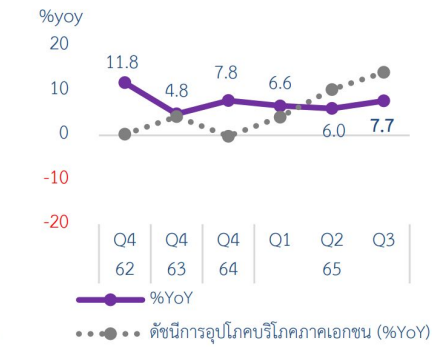
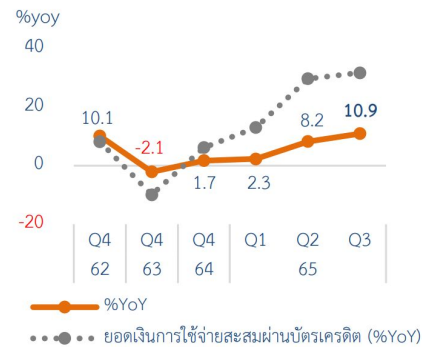
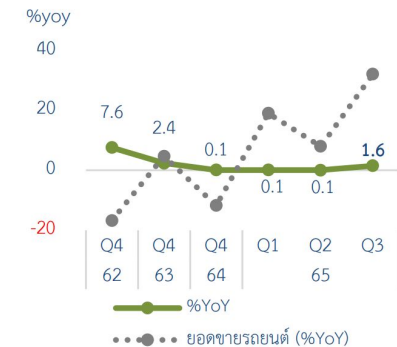
การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

บทนำ



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

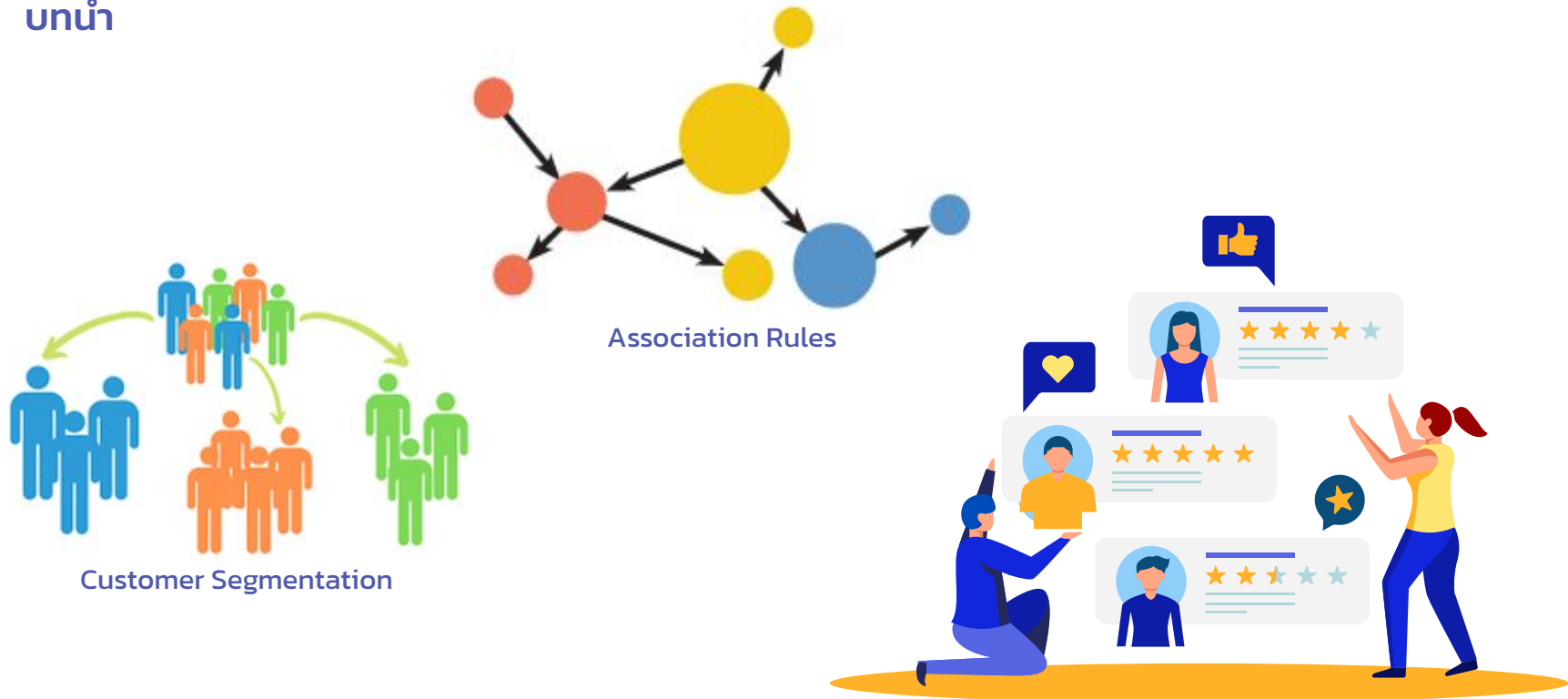
บทนำ



อัตราขยายตัวเทียบระยะเดียวกันปีก่อน | ตัวเลขในวงเล็บ แสดงสัดส่วนต่อสินเชื่อรวม

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินค้าส่วนบุคคล

บทนำ



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

วัตถุประสงค์

1. เพื่อแบ่งกลุ่มลูกค้าที่ใช้บริการทางการเงิน
2. เพื่อหากฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ขอบเขตการวิจัย

ข้อมูลที่ใช้ในการศึกษาเก็บรวบรวมมาจากเว็บไซต์ Kaggle ซึ่งเป็นข้อมูลลูกค้าที่ใช้บริการทางการเงิน UniversalBank ทั้งหมด 5,000 ตัวอย่าง ชุดข้อมูลนี้ประกอบด้วย 13 ตัวแปร

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ทราบแนวทางและวิธีการในการแบ่งกลุ่มลูกค้าอย่างเหมาะสม
2. ได้รับความรู้เกี่ยวกับการหากฎความสัมพันธ์ และสร้างความเข้าใจในพฤติกรรมในการเลือกซื้อหรือใช้บริการของลูกค้าได้

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ชุดข้อมูลที่ใช้ในการศึกษา

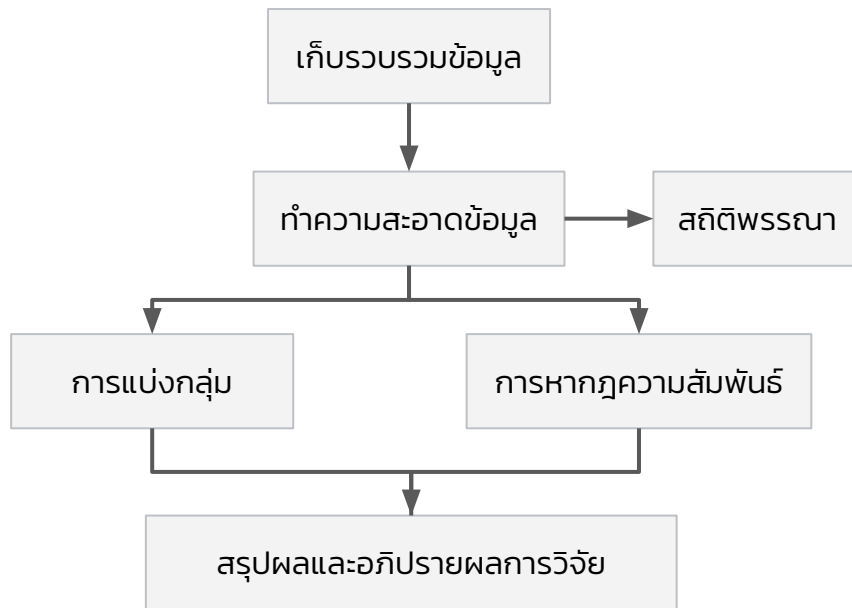
ข้อมูลลูกค้าที่ใช้บริการทางการเงินกับ UniversalBank มีจำนวน 5,000 ตัวอย่าง และจำนวนตัวแปร 13 ตัวแปร

	Age	Experience	Income	ZIP Code	Family	CCAvg	Education	Mortgage	Personal Loan	Securities Account	CD Account	Online	CreditCard
0	25	1	49	91107	4	1.6	1	0	0	1	0	0	0
1	45	19	34	90089	3	1.5	1	0	0	1	0	0	0
2	39	15	11	94720	1	1.0	1	0	0	0	0	0	0
3	35	9	100	94112	1	2.7	2	0	0	0	0	0	0
4	35	8	45	91330	4	1.0	2	0	0	0	0	0	1
...
4995	29	3	40	92697	1	1.9	3	0	0	0	0	1	0
4996	30	4	15	92037	4	0.4	1	85	0	0	0	1	0
4997	63	39	24	93023	2	0.3	3	0	0	0	0	0	0
4998	65	40	49	90034	3	0.5	2	0	0	0	0	1	0
4999	28	4	83	92612	3	0.8	1	0	0	0	0	1	1

5000 rows × 13 columns

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

แผนผังขั้นตอนการวิเคราะห์ข้อมูล



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

Exploratory Data Analysis

	Age	Experience	Income	ZIP Code	Family	CCAvg	Education	Mortgage	Personal Loan	Securities Account	CD Account	Online	CreditCard
89	25	-1	113	94303	4	2.30	3	0	0	0	0	0	1
226	24	-1	39	94085	2	1.70	2	0	0	0	0	0	0
315	24	-2	51	90630	3	0.30	3	0	0	0	0	1	0
451	28	-2	48	94132	2	1.75	3	89	0	0	0	1	0
524	24	-1	75	93014	4	0.20	1	0	0	0	0	1	0
536	25	-1	43	92173	3	2.40	2	176	0	0	0	1	0
540	25	-1	109	94010	4	2.30	3	314	0	0	0	1	0
576	25	-1	48	92870	3	0.30	3	0	0	0	0	0	1
583	24	-1	38	95045	2	1.70	2	0	0	0	0	1	0

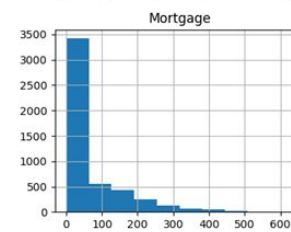
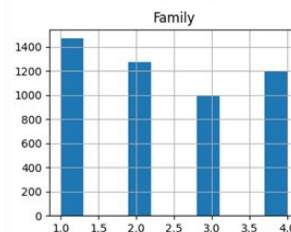
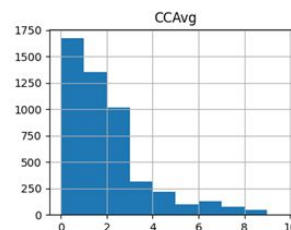
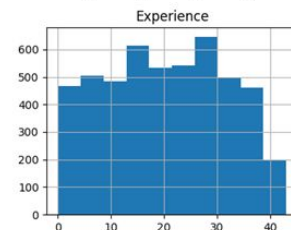
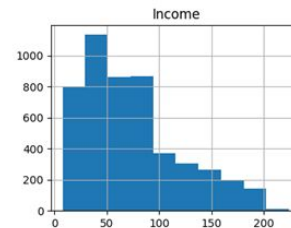
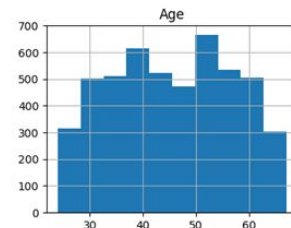
ตัวอย่าง ข้อมูลลูกค้าที่ใช้บริการทางการเงินกับ UniversalBank

พบว่า ตัวแปรจำนวนประสบการณ์การทำงาน (Experience) มีค่าติดลบจำนวน 52 ตัวอย่าง ผู้วิจัยจัดการข้อมูลส่วนนี้ด้วยการนำข้อมูลติดลบออก ดังนั้น จะได้ข้อมูลนำไปใช้ในการวิเคราะห์ต่อจำนวน 4,948 ตัวอย่าง

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ค่าสถิติพรรณนาและแผนภูมิฮิสโตแกรมของตัวแปรเชิงปริมาณของข้อมูลลูกค้า

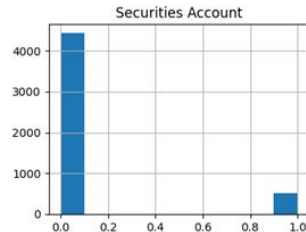
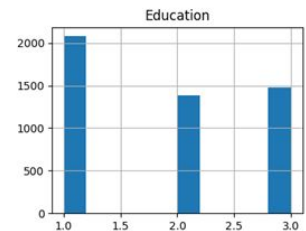
ตัวแปร	count	mean	std	min	median	max
Age	4948	45.5572	11.3207	24	46	67
Experience	4948	20.3310	11.3120	0	20	43
Family	4948	2.3915	1.1484	1	2	4
Income	4948	73.8145	46.1126	8	64	224
CCAvg	4948	1.9359	1.7477	0	1.5	10
Mortgage	4948	56.6344	101.8289	0	0	635



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ค่าสถิติพรรณนาและแผนภูมิฮิสโตแกรมของตัวแปรเชิงคุณภาพของข้อมูลลูกค้า

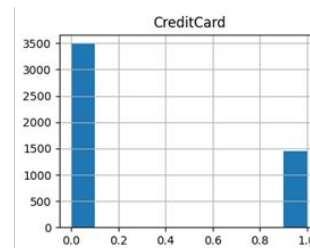
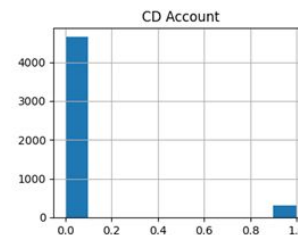
ตัวแปร	count	ร้อยละ	min	max
Education	4948		1	3
1: ระดับอนุปริญญาหรือปริญญาบัณฑิต	2080	42.04		
2: ระดับปริญญาโท	1387	28.03		
3: ระดับปริญญาเอก	1481	29.93		
Securities Account	4948		0	1
0: ไม่มี	4432	89.57		
1: มี	516	10.43		



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ค่าสถิติพรรณนาและแผนภูมิฮิสโตแกรมของตัวแปรเชิงคุณภาพของข้อมูลลูกค้า (Con't)

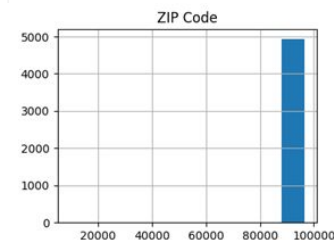
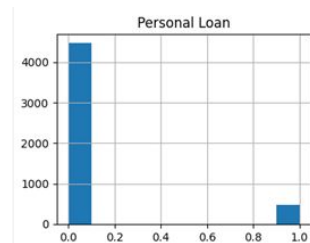
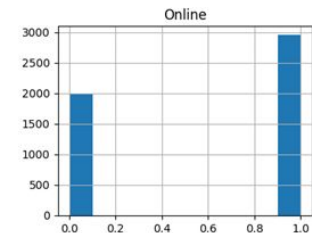
ตัวแปร	count	ร้อยละ	min	max
CD Account	4948		0	1
0: ไม่มี	4646	93.90		
1: มี	302	6.10		
CreditCard	4948		0	1
0: ไม่ใช้บริการ	3493	70.59		
1: ใช้บริการ	1455	29.41		



การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ค่าสถิติพรรณนาและแผนภูมิฮิสโตแกรมของตัวแปรเชิงคุณภาพของข้อมูลลูกค้า (Con't)

ตัวแปร	count	ร้อยละ	min	max
Online	4948		0	1
0: ไม่ใช้บริการ	1994	40.30		
1: ใช้บริการ	2954	59.70		
Personal Loan	4948		0	1
0: ไม่ยอมรับ	4468	90.30		
1: ยอมรับ	480	9.70		
ZIP Code	4948		9307	96651



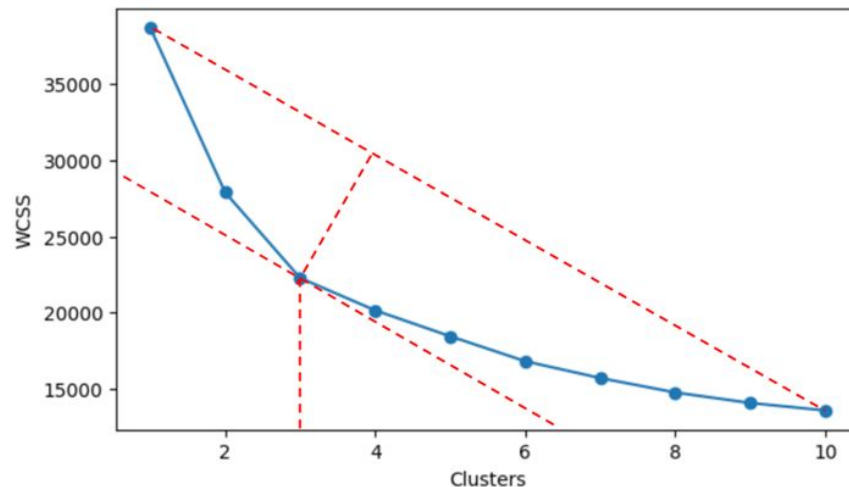
การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ผลการดำเนินงานการแบ่งกลุ่ม

	Age	Experience	Income	Family	CCAvg	Education	Mortgage
0	-1.816072	-1.709074	-0.538182	1.400757	-0.192215	-1.046782	-0.556228
1	-0.049224	-0.117679	-0.863506	0.529926	-0.249439	-1.046782	-0.556228
2	-0.579278	-0.471322	-1.362335	-1.211736	-0.535558	-1.046782	-0.556228
3	-0.932648	-1.001787	0.567918	-1.211736	0.437247	0.144176	-0.556228
4	-0.932648	-1.090198	-0.624935	1.400757	-0.535558	0.144176	-0.556228
...
4943	-1.462703	-1.532252	-0.733376	-1.211736	-0.020543	1.335134	-0.556228
4944	-1.374360	-1.443841	-1.275582	1.400757	-0.878901	-1.046782	0.278590
4945	1.540939	1.650538	-1.080388	-0.340905	-0.936125	1.335134	-0.556228
4946	1.717624	1.738949	-0.538182	0.529926	-0.821677	0.144176	-0.556228
4947	-1.551045	-1.443841	0.199218	0.529926	-0.650006	-1.046782	-0.556228

4948 rows × 7 columns

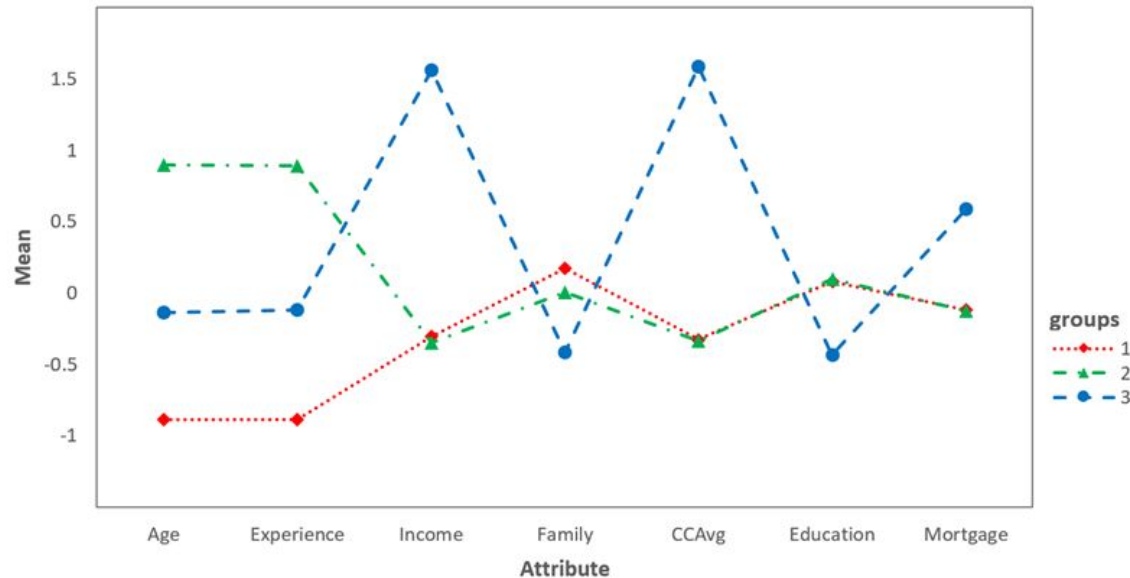
ตัวอย่าง ข้อมูลหลังจากแปลงให้อยู่ในรูป Standardized



พิจารณาเลือกค่า k ด้วย Elbow Method (k=3)

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ผลการดำเนินงานการแบ่งกลุ่ม



โปรไฟล์พล็อตแสดงลักษณะของค่าเฉลี่ยในแต่ละกลุ่ม

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ผลการดำเนินงานการหาความสัมพันธ์

	Personal Loan	Securities Account	CD Account	Online	CreditCard	group 1	group 2	group 3
0	0	1	0	0	0	1.0	0.0	0.0
1	0	1	0	0	0	1.0	0.0	0.0
2	0	0	0	0	0	1.0	0.0	0.0
3	0	0	0	0	0	1.0	0.0	0.0
4	0	0	0	0	1	1.0	0.0	0.0
...
4943	0	0	0	1	0	1.0	0.0	0.0
4944	0	0	0	1	0	1.0	0.0	0.0
4945	0	0	0	0	0	0.0	1.0	0.0
4946	0	0	0	1	0	0.0	1.0	0.0
4947	0	0	0	1	1	1.0	0.0	0.0

4948 rows × 8 columns

ตัวอย่าง ชุดข้อมูลสำหรับการหาความสัมพันธ์

	Personal Loan	Securities Account	CD Account	Online	CreditCard	group 1	group 2	group 3
0	False	True	False	False	False	True	False	False
1	False	True	False	False	False	True	False	False
2	False	False	False	False	False	True	False	False
3	False	False	False	False	False	True	False	False
4	False	False	False	False	True	True	False	False
...
4943	False	False	False	True	False	True	False	False
4944	False	False	False	True	False	True	False	False
4945	False	False	False	False	False	False	True	False
4946	False	False	False	True	False	False	True	False
4947	False	False	False	True	True	True	False	False

4948 rows × 8 columns

ตัวอย่าง ข้อมูลหลังจากแปลงให้อยู่ในรูปค่า Boolean

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ผลการดำเนินงานการหากฎความสัมพันธ์

min support = 0.01 และ min confidence = 0.7

	support	itemsets
0	0.097009	(Personal Loan)
1	0.104285	(Securities Account)
2	0.061035	(CD Account)
3	0.597009	(Online)
4	0.294058	(CreditCard)
...
64	0.017381	(CD Account, CreditCard, Securities Account, O...
65	0.013945	(CD Account, CreditCard, group 1, Online)
66	0.013339	(group 2, CD Account, CreditCard, Online)
67	0.017987	(CD Account, CreditCard, group 3, Online)
68	0.012935	(CD Account, group 3, Online, Personal Loan, C...

69 rows × 2 columns

ตัวอย่าง รายการความถี่ที่เกิดขึ้นบ่อย

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(CD Account)	(Online)	0.061035	0.597009	0.057195	0.937086	1.569635	0.020757	6.405438
1	(CD Account)	(CreditCard)	0.061035	0.294058	0.048504	0.794702	2.702533	0.030557	3.438620
2	(Personal Loan, Securities Account)	(CD Account)	0.012126	0.061035	0.010307	0.850000	13.926490	0.009567	6.259768
3	(Personal Loan, CD Account)	(Online)	0.028294	0.597009	0.025465	0.900000	1.507515	0.008573	4.029911
4	(Personal Loan, CD Account)	(group 3)	0.028294	0.171989	0.021625	0.764286	4.443814	0.016759	3.512775
5	(CD Account, group 3)	(Personal Loan)	0.027082	0.097009	0.021625	0.798507	8.231281	0.018998	4.481511
6	(CD Account, Securities Account)	(Online)	0.029709	0.597009	0.026677	0.897959	1.504097	0.008941	3.949313
7	(CreditCard, Securities Account)	(CD Account)	0.028294	0.061035	0.019806	0.700000	11.468874	0.018079	3.129884
8	(group 3, Securities Account)	(Online)	0.018998	0.597009	0.013339	0.702128	1.176076	0.001997	1.352899
9	(CD Account, CreditCard)	(Online)	0.048504	0.597009	0.045271	0.933333	1.563349	0.016313	6.044867
10	(CD Account, Online)	(CreditCard)	0.057195	0.294058	0.045271	0.791519	2.691710	0.028452	3.386128
11	(CD Account)	(CreditCard, Online)	0.061035	0.177445	0.045271	0.741722	4.180000	0.034440	3.184763
12	(CD Account, group 1)	(Online)	0.016774	0.597009	0.016370	0.975904	1.634655	0.006356	16.724131
13	(CD Account, group 2)	(Online)	0.017179	0.597009	0.015764	0.917647	1.537074	0.005508	4.893463
14	(CD Account, group 3)	(Online)	0.027082	0.597009	0.025061	0.925373	1.550016	0.008893	5.400081

ตัวอย่าง กฎความสัมพันธ์ที่ปรากฏร่วมกันบ่อย

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ผลการดำเนินงานการหากฎความสัมพันธ์

min support = 0.01 และ min confidence = 0.7

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	('CD Account', 'group 3')	('Personal Loan')	0.027082	0.097009	0.021625	0.798507	8.231281	0.018998	4.481511
1	('CD Account', 'group 3', 'Online')	('Personal Loan')	0.025061	0.097009	0.019604	0.782258	8.063777	0.017173	4.147070
2	('CD Account', 'CreditCard', 'group 3')	('Personal Loan')	0.019604	0.097009	0.014551	0.742268	7.651546	0.012650	3.503605
3	('CD Account', 'CreditCard', 'group 3', 'Online')	('Personal Loan')	0.017987	0.097009	0.012935	0.719101	7.412734	0.011190	3.214648

ตัวอย่าง กฎความสัมพันธ์ของการยอมรับสินเชื่อส่วนบุคคล (Personal Loan)

- ❑ IF 'CD Account' = True and 'group 3' = True THEN 'Personal Loan' = True
- ❑ IF 'CD Account' = True, 'group 3' = True, and 'Online' = True THEN 'Personal Loan' = True
- ❑ IF 'CD Account' = True, 'group 3' = True, and 'CreditCard' = True THEN 'Personal Loan' = True
- ❑ IF 'CD Account' = True, 'group 3' = True, 'CreditCard' = True, and 'Online' = True THEN 'Personal Loan' = True

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

สรุปและอภิปรายผลการวิจัย

		Mean						Mode
group	count	Age	Experience	Income	Family	CCAvg	Mortgage	Education
1	1,999	35.55	10.33	59.84	2.59	1.37	44.49	1
2	2,098	55.71	30.40	57.90	2.40	1.35	43.80	1
3	851	44.03	19.02	145.87	1.91	4.71	116.82	1

ตารางผลลัพธ์การแบ่งกลุ่มลูกค้าและค่ากลางตัวแปรในแต่ละกลุ่ม

จากข้อมูลชุดนี้ พบว่า ในกลุ่มลูกค้ากลุ่มที่ 3 ลูกค้าที่มีบัญชีเงินฝากประจำ (CD Account) กับธนาคารบางส่วนไม่ได้ยอมรับสินเชื่อส่วนบุคคลในแคมเปญก่อนหน้านี้ จำนวน 27 หลาย จากทั้งหมด 851 หลาย

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

ข้อเสนอแนะ

1. งานวิจัยครั้งนี้ใช้วิธีการพิจารณาเลือกค่า k ด้วย Elbow Method วิธีการเดียว ในการวิจัยครั้งต่อไป อาจเปลี่ยนวิธีการพิจารณาเลือกค่า k ด้วยวิธีการอื่น เช่น Silhouette Method, Gap Statistic เป็นต้น
2. สำหรับการหากฎความสัมพันธ์ การกำหนดค่าสนับสนุนขั้นต่ำและค่าความเชื่อมั่นขั้นต่ำนั้น ขึ้นอยู่กับผู้วิจัยและการใช้งาน ซึ่งจะได้ผลลัพธ์ที่แตกต่างกันไปตามชุดข้อมูลที่เลือกมาใช้งาน
3. หากมีการเพิ่มข้อมูลลูกค้าใหม่เข้ามา ควรจะศึกษาการแบ่งกลุ่มและการหากฎความสัมพันธ์ใหม่ เพื่อให้สอดคล้องกับพฤติกรรมของลูกค้าที่เปลี่ยนแปลงไปในอนาคต

การแบ่งกลุ่มลูกค้าและกฎความสัมพันธ์สำหรับการเสนอสินเชื่อส่วนบุคคล

วรรณกรรมที่เกี่ยวข้อง

- Musadig Aliyev, et al. (2020)
 - ศึกษาอัลกอริธึมการแบ่งกลุ่มที่แตกต่างกันสามแบบกับข้อมูลลูกค้าตามความถี่และภูมิหลังทางการเงิน โมเดลแรกใช้อัลกอริธึม K-Means โมเดลที่สองใช้การรวมกันของ DBSCAN และ K-Means และโมเดลที่สามใช้อัลกอริธึม Agglomerative clustering ผลการวิจัยพบว่า โมเดลที่สองสามารถแบ่งกลุ่มลูกค้าที่มีมูลค่าทางการเงินของลูกค้าได้มากที่สุด
- Malik Mubasher Hassan และ Tabasum Mirza (2020)
 - ศึกษาอัลกอริธึมการจัดประเภท Naïve Bayes สำหรับการทำโปรไฟล์ลูกค้า และอัลกอริธึมการจัดกลุ่ม BIRCH สำหรับการแบ่งกลุ่มลูกค้า ประโยชน์หลักของการทำโปรไฟล์ลูกค้าและการแบ่งกลุ่ม ได้แก่ การลูกค้า การรักษา/ขยายฐานลูกค้า และพัฒนาผลิตภัณฑ์เฉพาะกลุ่มลูกค้า
- Vesela Mihova และ Velisar Pavlov (2018)
 - การศึกษานี้ใช้การจัดกลุ่ม K-mean เพื่อระบุกลุ่มผู้กู้ที่รักดีสามกลุ่ม (platinum, gold and silver) ตามประวัติเครดิต 3 ตัวแปร ได้แก่ สถานะแยกที่สุดในรอบ 12 เดือน, วงเงินกู้ และเวลาที่ใช้บริการกับธนาคาร
- Mohammad Ali Farajian & Shahriar Mohammadi (2010)
 - อัลกอริธึม K-mean และกฎความสัมพันธ์ Apriori โดยอัลกอริธึม K-mean ใช้เพื่อระบุกลุ่มลูกค้าตามความถี่ และตัวทำนายการให้คะแนนพฤติกรรมทางการเงิน