การทำเหมืองข้อมูลเพื่อวิเคราะห์แบบประกันรถยนต์

Data Mining for Car Insurance Analysis

ธันย์ชนก สุระศันสนีย์

สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2560

Data Mining for Car Insurance Analysis

Thunchanok Surasunsanee

CO-OPERATIVE EDUCATION SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR

THE DEGREE OF BACHELOR OF SCIENCE (APPLIED MATHEMATICS)

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE

KING MONGKUT'S INSTIYUTE OF TECHNOLOGY LADKRABANG

ACADEMIC YEAR 2017

หัวข้อสหกิจศึกษา	การทำเหมืองข้อมูลเพื่อวิเคราะห์แบบประกันรถยนต์ Data Mining for Car Insurance Analysis
ชื่อนักศึกษา	นางสาวธันย์ชนก สุระศันสนีย์ รหัสนักศึกษา 57050073
ปริญญา	วิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)
ภาควิชา	คณิตศาสตร์
ปีการศึกษา	2560
อาจารย์ที่ปรึกษา	รศ.ดร.ฉัฐไชย์ ลีนาวงศ์

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้สหกิจศึกษานี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์) ประจำปีการศึกษา 2560

คณะกรรมการ	ลายมือชื่อ
ผศ.ดร.กาญจนา คำนึงกิจ ประธานกรรมการ	
ดร.สิริพร แฮนน่า วินเทอร์ กรรมการ	
รศ.ดร.ฉัฐไชย์ ลีนาวงศ์ กรรมการและอาจารย์ที่ปรึกษา	

ลิขสิทธิ์ของคณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

หัวข้อสหกิจศึกษา การทำเหมืองข้อมูลเพื่อวิเคราะห์แบบประกันรถยนต์

Data Mining for Car Insurance Analysis

ชื่อนักศึกษา นางสาวธันย์ชนก สุระศันสนีย์ รหัสนักศึกษา 57050073

ปริญญา วิทยาศาสตรบัณฑิต (คณิตศาสตร์ประยุกต์)

ภาควิชา คณิตศาสตร์

ปีการศึกษา 2560

อาจารย์ที่ปรึกษา รศ.ดร.ฉัฐไชย์ ลีนาวงศ์

บทคัดย่อ

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อศึกษาการทำเหมืองข้อมูล (Data Mining) เพื่อวิเคราะห์แบบประกันรถยนต์ โดยนำเทคนิค Classification มาช่วยในการสร้างโมเดลซึ่งมี 3 วิธี คือ Decision Tree, Naïve Bayes และ K-Nearest Neighbors โดยเริ่มตั้งแต่การรวบรวมข้อมูล การคัดเลือกข้อมูลและการลดขนาดข้อมูล ซึ่งทำให้สามารถลดข้อมูลที่นำไปใช้ในการสร้างโมเดลจาก 5,000 ข้อมูล เหลือเพียง 2,134 ข้อมูล โดยใช้โปรแกรม RapidMiner เข้ามาช่วยในการหาจำแนกแบ่งกลุ่มทั้ง 3 วิธีข้างต้น. ซึ่งมีค่าความแม่นยำดังนี้ 77.05%, 70.38%, 69.36% ตามลำดับ และเมื่อนำเข้าข้อมูลใหม่เพื่อมาทำนาย พบว่าโดยส่วนใหญ่จะได้ผลลัพธ์การทำนายเป็นแบบประกัน MO1 (ประกันชั้น 1)

Co – Operative Education Title	Data Mining for Car Insurance Analysis
Student	Miss.Thunchanok Surasunsanee 57050073
Degree	Bachelor of Science (Applied Mathematic)
Department	Mathematic
Academic Year	2017
Adviser	Assoc.Prof.Dr.Chartchai Leenawong

Abstract

This Co-Operative Education is intended to study data mining for car insurance analysis by using classification techniques. They are Decision Tree, Naïve Bayes and K-Nearest Neighbors. Starting with data collection, data selection and data reduction, the number of data is reduced from 5,000 to 2,134. Using the RapidMiner program to classify the car insurance buyers for insurance policy prediction by the above three methods, results in 77.05%, 70.38% and 69.36% accuracies, respectively. When new examples are input to be classified, the MO1 insurance is the most suggested policy by all the methods.

กิตติกรรมประกาศ

ตามที่ข้าพเจ้า นางสาว ธันย์ชนก สุระศันสนีย์ ได้ปฏิบัติงานสหกิจศึกษา ตั้งแต่วันที่ 12 ธันวาคม 2560 ถึง วันที่ 12 เมษายน 2561 ทำให้ข้าพเจ้าได้รับความรู้และประสบการณ์ต่างๆ จึงขอขอบพระคุณ รศ.ดร.ฉัฐไชย์ ลีนาวงศ์ อาจารย์ที่ปรึกษาสหกิจและขอบคุณพี่เลี้ยง และกรรมการสอบสหกิจศึกษาทุกท่าน ที่คอยให้การสนับสนุน ให้คำปรึกษา และคอยดูแลอย่างใกล้ชิด ทำให้การทำสหกิจศึกษาครั้งนี้ผ่านไปได้ด้วยดี

นอกจากนี้ยังมีบุคคลท่านอื่นๆ อีกที่ไม่ได้กล่าว ณ ที่นี้ ซึ่งให้ความกรุณาแนะนำ ในการจัดทำสหกิจศึกษาครั้งนี้ ข้าพเจ้าจึงขอขอบพระคุณทุกท่านที่ได้มีส่วนร่วมในการให้ข้อมูลและให้ความเข้าใจเกี่ยวกับชีวิตของการปฏิบัติงาน รวมถึงเป็นที่ปรึกษาในการจัดทำสหกิจศึกษาครั้งนี้ด้วย

ธันย์ชนก สุระศันสนีย์

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ก

บทคัดย่อภาษาอังกฤษ ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ช

สารบัญรูป ..... ช

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ 1

1.3 ขอบเขตของปัญหา 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 2

1.5 ขั้นตอนการดำเนินงาน 2

1.6 ระยะเวลาในการดำเนินงาน 3

บทที่ 2 ความรู้พื้นฐานที่เกี่ยวข้อง 4

2.1 การทำเหมืองข้อมูล (Data Mining) 4

2.1.1 ขั้นตอนการทำเหมืองข้อมูล 5

2.1.2 เทคนิคต่างๆ ของ Data Mining 5

2.1.3 ตัววัดประสิทธิภาพของโมเดล 19

2.1.4 ประโยชน์ของ Data Mining 20

สารบัญ (ต่อ)

หน้า

2.2 ความรู้ทั่วไปด้านประกันภัยรถยนต์ 20

2.2.1 การประกันภัยรถยนต์ภาคบังคับ 20

2.2.2 การประกันภัยรถยนต์ภาคสมัครใจ 20

2.3 คำศัพท์ของด้านประกัน 23

บทที่ 3 วิธีการดำเนินการ 26

3.1 ลักษณะข้อมูลและการเตรียมข้อมูล 26

3.2 การคัดเลือกข้อมูลและการลดขนาดของข้อมูล 26

3.2.1 การคัดเลือกข้อมูล 26

3.2.2 การลดขนาดข้อมูล 26

3.3 การเลือกเทคนิค 27

3.4 ขั้นตอนการดำเนินงาน 27

3.4.1 ขั้นตอนการทำงานของ Decision Tree 28

3.4.2 ขั้นตอนการทำงานของ Naïve Bayes 29

3.4.3 ขั้นตอนการทำงานของ K-Nearest Neighbors (K-NN) 30

3.5 การทดสอบโมเดล 31

3.5.1 วิธี Decision Tree 31

3.5.2 วิธี Naïve Bayes 32

3.5.3 วิธี K-Nearest Neighbors (K-NN) 32

3.6 ผลการวิเคราะห์และผลการทำนายข้อมูลใหม่ 32

สารบัญ (ต่อ)

หน้า

บทที่ 4 ผลการวิจัยและการอภิปรายผล 33

4.1 ผลลัพธ์ของ Decision Tree 33

4.2 ผลลัพธ์ Naïve Bayes 35

4.3 ผลลัพธ์ K-Nearest Neighbors (K-NN) 36

4.4 การทดสอบข้อมูลใหม่ 38

4.5 ผลสรุป 53

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ 54

5.1 สรุปผลการวิจัย 54

5.2 ข้อเสนอแนะ 54

เอกสารอ้างอิง 55

ภาคผนวก 56

ภาคผนวก ก 57

ภาคผนวก ข 62

สารบัญตาราง

ตารางที่ **หน้า**

1.1 ขั้นตอนการดำเนินงาน 3

2.1 ตัวอย่างข้อมูล Decision Tree 9

2.2 ตัวอย่างข้อมูล Naïve Bayes 15

2.3 ตัวอย่างข้อมูล K-Nearest Neighbors 16

2.4 ข้อมูลหลังจากคำนวณ 17

2.5 ข้อมูลที่เรียงลำดับจากน้อยไปมาก 18

2.6 ระยะทางที่น้อยที่สุด 5 อันดับแรก 19

2.7 ตารางเปรียบเทียบประเภทประกันภัยรถยนต์ 23

4.1 ผลลัพธ์ของ Decision Tree 34

4.2 ผลลัพธ์ของ Naïve Bayes 35

4.3 ผลลัพธ์ของ K-Nearest Neighbors 37

4.4 ผลลัพธ์ของการทดสอบข้อมูลใหม่วิธี Decision Tree 38

4.5 ผลลัพธ์ของการทดสอบข้อมูลใหม่วิธี Naïve Bayes 43

4.6 ผลลัพธ์ของการทดสอบข้อมูลใหม่วิธี K-Nearest Neighbors 48

4.7 ตารางผลสรุป 53

สารบัญรูป

รูปที่ **หน้า**

2.1 ขั้นตอนการทำงานของ Data Mining 4

2.2 การ Clustering ข้อมูล 6

2.3 ขั้นตอนการทำงาน Classification 7

2.4 แสดงโมเดล Decision Tree 13

3.1 ข้อมูลที่นำเข้า 27

3.2 การจำแนกข้อมูล Decision Tree 28

3.3 แสดงการทำงานของ Decision Tree 28

3.4 หน้าแสดงผลลัพธ์ของ Decision Tree 29

3.5 การเลือกฟังก์ชันเพื่อใช้สำหรับการสร้างโมเดลของเทคนิค Naïve Bayesian 29

3.6 แสดงการทำงานของ Naïve Bayesian 30

3.7 การเลือกฟังก์ชันเพื่อใช้สำหรับการสร้างโมเดลของเทคนิค K-Nearest Neighbors 30

3.8 แสดงการทำงานของ K-Nearest Neighbors 31

3.9 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี Decision Tree 31

3.10 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี Naïve Bayes 32

3.11 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี K-Nearest Neighbors 32

บทที่ 1

บทนำ

ในบทนี้จะกล่าวถึงที่มาและความสำคัญของปัญหา วัตถุประสงค์ ขอบเขตของการทำงาน ประโยชน์ที่คาดว่าจะได้รับ ขั้นตอนการทำงานและระยะเวลาของการทำงาน ซึ่งจะมีรายละเอียดดังนี้

1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบันนี้รถยนต์ได้เข้ามามีบทบาทในชีวิตประจำวัน ดังนั้นรถยนต์เป็นสิ่งสำคัญในการดำรงชีวิต จึงทำให้มีปริมาณรถยนต์เพิ่มขึ้นทุกปี จากปริมาณรถยนต์ที่เพิ่มขึ้นทำให้เกิดปัญหาการจราจรที่ติดขัดและมีความเสี่ยงของการเกิดอุบัติเหตุ สาเหตุที่ทำให้เกิดอุบัติเหตุนั้นอาจเกิดจากความใจร้อนของผู้ขับขี่ ความประมาท และการฝ่าฝืนกฎจราจร ดังนั้นจึงได้มีการออกกฎหมายประกันภัย พ.ร.บ.ในปี พ.ศ. 2535 ขึ้นมาซึ่งได้กล่าวไว้ว่า "รถทุกคันที่จดทะเบียนกับกรมการขนส่งทางบกจะต้องทำและมีไว้เป็นหลักประกันให้กับคนในรถทุกคันหรือผู้ที่ใช้รถใช้ถนนว่าจะได้รับสิทธิความคุ้มครองจากเงินกองกลางที่รถทุกคันได้ทำ พ.ร.บ. ว่าจะได้รับความคุ้มครอง เงินค่ารักษาพยาบาลจากการเกิดอุบัติ" ซึ่งเรียกว่าการประกันภัยรถยนต์ภาคบังคับ อาจมีความคุ้มครองที่ยังไม่เพียงพอต่อการชดเชยค่าความเสียหายที่เกิดขึ้นจากอุบัติเหตุจึงทำให้มีประกันภัยภาคสมัครใจขึ้นมา สามารถแบ่งประกันภัยภาคสมัครใจออกเป็น 5 ประเภท คือ ประกันภัยประเภท 1, 2, 3, 2+ และ 3+ ซึ่ง 5 ประเภทนี้ความคุ้มครองก็จะแตกต่างกันออกไป โดยแต่ละบริษัทรับประกันภัยให้ผู้บริโภคได้มีการจัดความคุ้มครองของการรับประกันภัยรถยนต์หลากหลายแบบให้ตรงกับความต้องการในการใช้งานและเหมาะสมกับความเสี่ยงที่อาจจะเกิดขึ้นจึงทำให้เล็งเห็นว่าหากนำข้อมูลในอดีตมาทำการวิเคราะห์จะช่วยให้สามารถเลือกซื้อประเภทประกันภัยรถยนต์ภาคสมัครใจ ซึ่งจะเป็นประโยชน์ต่อการออกแบบผลิตภัณฑ์ที่ตรงกับความต้องการ โดยงานวิจัยนี้ได้เลือกใช้เทคนิคของ Data Mining มาช่วยในการศึกษาครั้งนี้โดยเทคนิค Data Mining มีหลายเทคนิคคือ Association Rule, Clustering และ Classification โดยแต่ละเทคนิคนั้นก็ให้ผลลัพธ์ที่แตกต่างกันออกไปซึ่งงานวิจัยชิ้นนี้ได้เลือกเทคนิค Classification มาช่วยในการทำนายโมเดลและช่วยในการตัดสินใจเลือกซื้อประเภทประกันภัยรถยนต์

1.2 วัตถุประสงค์

1. เพื่อหาแนวโน้มในการเลือกซื้อประเภทประกันภัยรถยนต์
2. เพื่อวิเคราะห์ความสัมพันธ์ระหว่างข้อมูลด้วยโปรแกรม RapidMiner Studio

1.3 ขอบเขตของปัญหา

ข้อมูลที่นำมาศึกษานั้นจะเป็นประกันภัยรถยนต์ภาคสมัครใจใช้ข้อมูลรถที่มีจำนวน 4 ล้อ ประเภทรถจะเป็น Saloon, Van/Bus, Pick Up และเป็นข้อมูลที่อยู่ในระหว่างการคุ้มครองจากบริษัทแห่งหนึ่ง

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้นำความรู้ด้าน Data Mining มาใช้กับข้อมูลจริงในสหกิจ
2. นำผลจากการทำ Data Mining มาวางแผนเพื่อสร้างกลยุทธ์ในด้านการขาย

1.5 ขั้นตอนการดำเนินงาน

1. เลือกหัวข้อปัญหา
2. ศึกษาข้อมูลและเก็บรวบรวมข้อมูลขององค์กร
3. นำข้อมูลมาทำการวิเคราะห์
4. รายงานผลการดำเนินงานที่ทำกับองค์กร
5. ส่งงานให้องค์กรทำการตรวจสอบ
6. แก้ไขตามคำแนะนำ
7. ตรวจสอบความถูกต้องของการวิเคราะห์ข้อมูลและเนื้อหาทั้งหมด
8. นำเสนอสหกิจศึกษา
9. จัดทำรูปเล่มสหกิจ

1.6 ระยะเวลาในการดำเนินงาน

ระยะเวลาการศึกษาตั้งแต่วันที่ 12 ธันวาคม 2560 ถึง 12 เมษายน 2561

ตารางที่ 1.1 ขั้นตอนการดำเนินงาน

ขั้นตอนการดำเนินการ	ระยะเวลาในการดำเนินการ					
	ปี 2560	ปี 2561				
	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.
1. เลือกหัวข้อปัญหา						
2. ศึกษาข้อมูลและเก็บรวบรวมข้อมูลขององค์กร						
3. นำข้อมูลมาทำการวิเคราะห์						
4. รายงานผลการดำเนินงานที่ทำกับองค์กร						
5. ส่งงานให้องค์กรทำการตรวจสอบ						
6. แก้ไขตามคำแนะนำ						
7. ตรวจสอบความถูกต้องของข้อมูลและเนื้อหาทั้งหมด						
8. นำเสนอสหกิจศึกษา						
9. จัดทำรูปเล่มสหกิจ						

หมายเหตุ : เนื่องจากสถานประกอบการที่ไปทำสหกิจขอให้ไม่เปิดเผยชื่อ ดังนั้นรายละเอียดต่างๆ ที่เกี่ยวข้องกับสถานประกอบการและใบสมัครการตอบรับต่างๆ จะไม่รวมอยู่ในรายงานนี้

สำหรับเนื้อหาในบทต่อๆไปนั้นจะกล่าวถึง ความรู้พื้นฐานที่เกี่ยวข้องในบทที่ 2 วิธีการดำเนินงานในบทที่ 3 ผลการวิจัยและการอภิปรายผลในบทที่ 4 และสรุปผลการวิจัยและข้อเสนอแนะในบทที่ 5 ซึ่งจะมีการอธิบายรายละเอียดในแต่ละบทที่กล่าวมาตามลำดับ

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในงานวิจัยฉบับนี้ได้ศึกษา การทำเหมืองข้อมูล (Data Mining) ความหมายทางด้านประกันภัยรถยนต์และคำศัพท์ทางด้านประกัน โดยแยกเป็นประเด็น ดังนี้

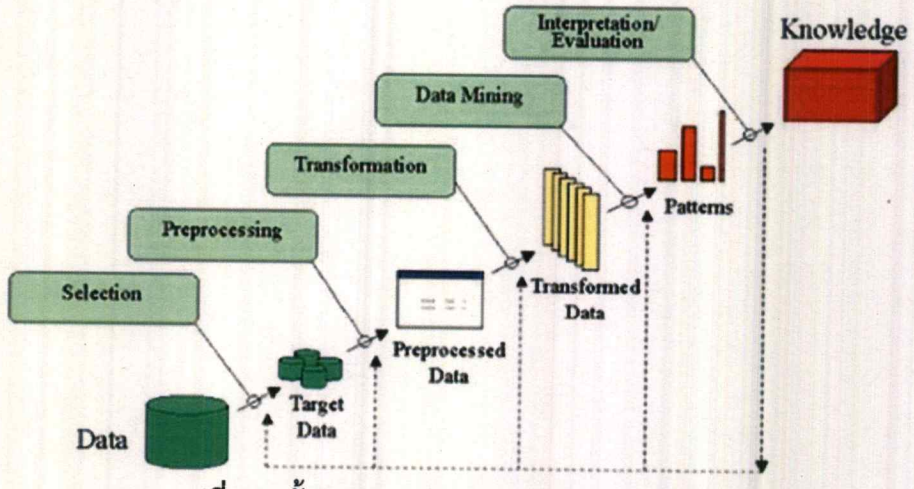
2.1 การทำเหมืองข้อมูล (Data Mining)

2.2 ความรู้ทั่วไปด้านประกันภัยรถยนต์

2.3 คำศัพท์ทางด้านประกันภัย

2.1 การทำเหมืองข้อมูล (Data Mining)

Data Mining คือ กระบวนการการของการค้นพบองค์ความรู้ที่น่าสนใจจากจำนวนข้อมูลที่มีมากมายที่ถูกเก็บไว้ในฐานข้อมูล Data Warehouses หรือที่เก็บข้อมูลอื่นๆดังนั้นโดยสรุปแล้วกระบวนการ Data Mining นี้เปรียบเสมือนการทำเหมืองแร่ที่ใช้เครื่องจักรคัดแยกแร่ที่เป็นที่ต้องการออกจากกองหิน กรวด ดินที่ปะปนมากับสายแร่ เพียงแต่ในกระบวนการ Data Mining สิ่งที่ได้จากกองข้อมูลมหาศาลคือ ความรู้ (Knowledge) ซ่อนอยู่ในกองข้อมูล ความรู้นี้จะช่วยให้เข้าใจลักษณะของข้อมูล และเข้าใจปัจจัยที่ทำให้เกิดลักษณะบางอย่างขึ้นในข้อมูลบางกลุ่ม ซึ่งจะสามารถช่วยทำนายแนวโน้มของข้อมูลใหม่ที่เกิดขึ้นในอนาคตได้ รวมถึงเข้าใจความสัมพันธ์ที่เชื่อมโยงข้อมูลแต่ละกลุ่มย่อยเข้าด้วยกัน ซึ่งจากกระบวนการสืบค้นความรู้ จะได้ดังรูปที่ 2.1



รูปที่ 2.1 ขั้นตอนการทำงานของ Data Mining

จากรูปที่ 2.1 สามารถสรุปเป็นกระบวนการทำ Data Mining ได้ 4 ขั้นตอนดังนี้

2.1.1 ขั้นตอนการทำเหมืองข้อมูล

ขั้นตอนที่ 1 เตรียมข้อมูล (Data Preparation) : ถ้าข้อมูลไม่อยู่ในรูปแบบที่ถูกต้องหรือเหมาะสม จะต้องมีการปรับข้อมูลให้อยู่ในรูปแบบที่โปรแกรม Data Mining จะเรียกใช้งานได้

ขั้นตอนที่ 2 ลดขนาดของข้อมูล (Data Reduction) : การจะหาโมเดลหรือ pattern ที่ข้อมูลส่วนใหญ่แสดงลักษณะเหล่านั้นออกมาเหมือนกัน จำเป็นต้องใช้ข้อมูลตัวอย่างจำนวนมาก ถ้าข้อมูลน้อยเกินไปอาจจะหาลักษณะร่วมเหล่านั้นไม่พบ แต่ในทางตรงกันข้ามถ้าข้อมูลมีปริมาณมากเกินไป การค้นหาโมเดลหรือ pattern จากกลุ่มข้อมูลขนาดใหญ่ต้องใช้เวลามาก ซึ่งถ้าลดจำนวนข้อมูลลงด้วยสัดส่วนที่ถูกต้อง โมเดลที่ได้ยังคงเป็นเช่นเดิมในขณะที่โปรแกรมใช้เวลาในการค้นหาโมเดลสั้นลง

ขั้นตอนที่ 3 ค้นหาโมเดลหรือความสัมพันธ์จากข้อมูล (Data Modeling / Discovery) : กระบวนการค้นหาโมเดลหรือความสัมพันธ์จะเริ่มจากข้อมูลเริ่มต้นจำนวนไม่มากนักจากนั้นนำผลที่ได้จากกระบวนการค้นหา (Learning Process / Method) ไปยืนยันกับข้อมูลทดสอบ ถ้าผลที่ได้ยังไม่พอใจอาจจะต้องปรับค่าพารามิเตอร์บางตัวของ Learning Method และเริ่มกระบวนการค้นหาใหม่กับข้อมูลจำนวนมากขึ้น จนกว่าผลที่ได้มีความถูกต้องอยู่ในระดับที่ยอมรับได้ จึงจะจบกระบวนการค้นหา

ขั้นตอนที่ 4 ตรวจสอบและวิเคราะห์ผล (Solution Analyses) : โมเดลหรือความสัมพันธ์ที่หามาได้ในขั้นตอนที่ 3 จะต้องถูกนำมาทดสอบอัตราความผิดพลาดและวิเคราะห์ความซับซ้อนของรูปแบบโมเดลถ้าอัตราความผิดพลาดยังสูงเกินไป อาจจะต้องย้อนกลับไปที่ขั้นตอนที่ 3 อีกครั้ง เพื่อปรับปรุงโมเดลให้ถูกต้องยิ่งขึ้น ในทำนองเดียวกัน ถ้าโมเดลที่หามาได้มีรูปแบบที่ซับซ้อนเกินไปจนยากต่อการทำความเข้าใจ อาจจะต้องย้อนกระบวนการกลับไปที่ขั้นตอนที่ 3 เพื่อให้หาโมเดลใหม่ที่มีความถูกต้องเท่าเดิมแต่มีรูปแบบที่ซับซ้อนน้อยลง

2.1.2 เทคนิคต่างๆของ Data Mining

การแก้ปัญหาของงานชนิดต่างๆ โดยใช้วิธี Data Mining ในแต่ละงานก็จะมีเทคนิคที่จะนำมาใช้ได้อย่างเหมาะสม โดยเทคนิคมีดังนี้

1. Association Rule

เป็นเทคนิคหนึ่งของ Data Mining ที่สำคัญ และสามารถนำไปประยุกต์ใช้ได้จริงกับงานต่างๆหลักการทำงานของวิธีนี้ คือ การค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลขนาดใหญ่ที่มีอยู่เพื่อนำไปใช้ในการวิเคราะห์หรือทำนายปรากฏการณ์ต่างๆหรือมาจากการวิเคราะห์การซื้อสินค้าของ

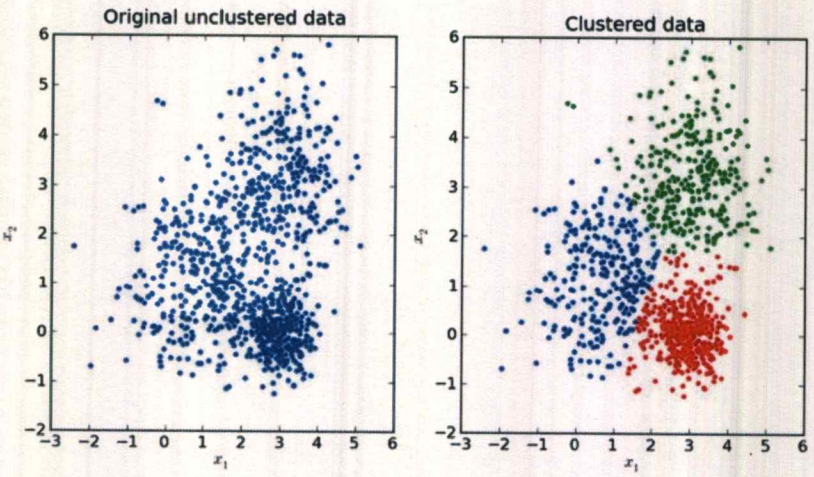
ลูกค้า เรียกว่า “Market Basket Analysis” ซึ่งประเมินจากข้อมูลในตารางที่รวบรวมไว้ ผลการวิเคราะห์ที่ได้จะเป็นคำตอบของปัญหา ซึ่งการวิเคราะห์แบบนี้เป็นการใช้ “กฎความสัมพันธ์” (Association Rule) เพื่อหาความสมพันธ์ของข้อมูล

ตัวอย่างการนำเทคนิคนี้ไปประยุกต์ใช้กับงานจริง ได้แก่ ระบบแนะนำหนังสือให้กับลูกค้าแบบอัตโนมัติของ Amazon ข้อมูลการสั่งซื้อทั้งหมดของ Amazon ซึ่งมีขนาดใหญ่มากจะถูกนำมาประมวลผลเพื่อหาความสัมพันธ์ของข้อมูล คือ ลูกค้าที่ซื้อหนังสือเล่มหนึ่งมักจะซื้อหนังสือเล่มใดพร้อมกันเสมอความสัมพันธ์ที่ได้จากกระบวนการนี้จะสามารถนำไปใช้คาดเดาได้ว่าควรแนะนำหนังสือเล่มใดเพิ่มเติมให้กับลูกค้าที่เพิ่งซื้อหนังสือจากร้าน ตัวอย่างเช่น buys (x, database) -> buys (x, data mining) [80%, 60%] หมายความว่า เมื่อซื้อหนังสือ database แล้วมีโอกาสที่จะซื้อหนังสือ data mining ด้วย 60% และมีการซื้อทั้งหนังสือ database และหนังสือ data mining พร้อมๆกัน 80%

2. Clustering

เป็นกระบวนการสร้างโมเดลเพื่อแบ่งข้อมูลเป็นแบบกลุ่มๆโดยที่ไม่รู้ล่วงหน้าว่าจะมีทั้งหมดกี่กลุ่ม โดยการจัดกลุ่มข้อมูลดังกล่าว ได้จากการพิจารณาคุณสมบัติในหลายๆมิติของข้อมูลถ้ารายการในข้อมูลมีลักษณะ คล้ายคลึงกันเป็นกลุ่มเดียวกันได้ เช่น เมื่อพิจารณาจากการกระจายข้อมูลจะเห็นได้ว่าหากพิจารณาการกระจายข้อมูลใน 2 มิติ ข้อมูลนั้นควรแบ่งได้เป็น 3 กลุ่ม ความรู้ใหม่ที่ได้คือในสเปซของข้อมูลทั้งหมด จะมีกลุ่มที่ต่างกันเพียง 3 กลุ่มเท่านั้น

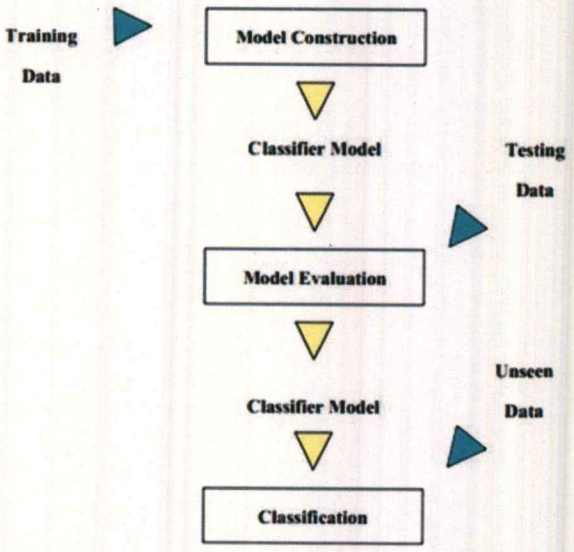
ยกตัวอย่างการนำไปใช้บริษัทผลิตรถยนต์แห่งหนึ่งมีข้อมูลที่ผ่านมาเกี่ยวกับข้อมูลรายได้ของลูกค้า ซึ่งมีตัวเลขเงินเดือนที่แตกต่างกันออกไป ไม่สามารถกำหนดได้ว่าลูกค้ามีรายได้ในระดับใด มีระดับสูงหรือระดับต่ำ จะใช้ค่าใดเป็นตัวแยกแยะระดับ แต่หากพิจารณาการกระจายข้อมูลก็อาจจะเห็นเป็นกลุ่มได้ในทำนองเดียวกัน



รูปที่ 2.2 การ Clustering ข้อมูล

3. Classification & Prediction

เป็นกระบวนการสร้างโมเดล จัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ ตัวอย่างเช่น จัดกลุ่มนักเรียนว่า ดีมาก ดี ปานกลาง ไม่ดี โดยพิจารณาจากประวัติและผลการเรียนหรือแบ่งประเภทของลูกค้าว่าเชื่อถือได้หรือไม่ โดยพิจารณาจากข้อมูลที่มีอยู่ กระบวนการ Classification นี้แบ่งออกเป็น 3 ขั้นตอน



รูปที่ 2.3 ขั้นตอนการทำงาน Classification

1. Model Construction (Learning) เป็นขั้นการสร้างโมเดลจากการเรียนรู้ข้อมูลที่มีการกำหนดคลาสไว้เรียบร้อยแล้ว (Training Data) ซึ่งโมเดลที่ได้อาจแสดงในรูปของแบบต้นไม้ (Decision Tree) หรือ แบบนิวรอลเน็ต (Neural Net) หรืออื่นๆ ขึ้นกับว่าจะนำเทคนิคใดมาใช้
2. Model Evaluation (Accuracy) เป็นขั้นการประเมินความถูกต้องโดยอาศัยข้อมูลที่ใช้ทดสอบ (Testing Data) ซึ่งคลาสที่แท้จริงของข้อมูลที่ใช้ทดสอบนี้จะถูกนำมาเปรียบเทียบกับคลาสที่หามาได้จากการทำนายจาก Model เพื่อทดสอบความถูกต้อง
3. Model Usage (Classification) เป็นการนำโมเดลที่ได้ไปใช้เพื่อแก้ปัญหาในการจำแนกข้อมูลที่ไม่เคยเห็นมาก่อน (Unseen Data) โดยจะทำการทำนายกลุ่มให้กับข้อมูลที่ไม่เคยเห็นมาก่อนหน้านี้

เทคนิคที่ใช้ทำ Data Mining ในแบบ Classification หรือ Predictive จะมีวิธีหลักๆ 3 วิธีดังนี้

1. Decision Tree เป็นเทคนิคที่ให้ผลลัพธ์ในลักษณะของโครงสร้างต้นไม้ ซึ่งเมื่อมีข้อมูลที่ต้องการจะจัดกลุ่ม ก็จะนำ attribute ต่างๆ ของข้อมูลนั้นไปเทียบกับ Decision Tree ตาม

เส้นทางใน tree จนกระทั่งคลาสปลายทางซึ่งก็คือ กลุ่มของข้อมูลที่เหมือนกัน โดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น

“If Income = High and Married = No THEN Risk = Poor”

“If Income = High and Married = Yes THEN Risk = Good”

ขั้นตอนในการสร้าง Decision Tree

1. เลือก Attribute ที่ทำหน้าที่เป็น Root Node

2. สร้างเส้นเชื่อมโยงไปยังโหนดลูก จำนวนเส้นเชื่อมโยงจะเท่ากับจำนวนค่าที่เป็นไปได้ทั้งหมดของ Attribute ที่เป็น Root Node

3. ถ้าโหนดลูกเป็นกลุ่มของข้อมูลที่อยู่ในคลาสเดียวกันทั้งหมดให้หยุดสร้างต้นไม้

4. แต่ถ้าโหนดลูกมีข้อมูลของหลายคลาสปะปนกันอยู่ ต้องสร้าง Subtree เพื่อจำแนกข้อมูลต่อไป โดยเลือก Subtree มาทำหน้าที่ Root Node ของ Subtree มาทำซ้ำในขั้นตอน 2, 3

ปัญหาคือควรจะเลือกแอตทริบิวต์ ใดเป็น Root Node ดังนั้นเกณฑ์การตัดสินใจ คือค่า Gain จะพิจารณาจากแอตทริบิวต์ที่มีค่า Gain สูงสุดเป็น Root Node

สูตร info (T) = $-\sum_{i=1}^{n}\left[\frac{freq(C_i,T)}{|T|}\right]\times\log_2\left[\frac{freq(C_i,T)}{|T|}\right]$

เมื่อ info (T) เป็นฟังก์ชันที่ระบุปริมาณข้อมูลที่ต้องการเพื่อให้สามารถจำแนกคลาสที่ต้องการ

เมื่อ $|T|$ คือ จำนวนข้อมูลทั้งหมดใน Training Dataset freq (C_j,T) คือ ความถี่ที่ข้อมูลใน T ปรากฏเป็นคลาส C_i

และ Gain (X) = info (T) – $info_x$ (T)

เมื่อ Gain เป็นค่าที่บอกถึงระดับความสามารถของการจำแนกคลาสในแต่ละแอตทริบิวต์

กำหนดให้ T แทน เซตของ Training Set

X แทน แอตทริบิวต์ที่ถูกเลือกให้เป็นตัวจำแนกข้อมูล

ตัวอย่าง

ตารางที่ 2.1 ตัวอย่างข้อมูล วิธี Decision Tree

No	Outlook	Temperature	Humidity	windy	Play
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Mild	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

จากข้อมูลในตารางที่ 2.1 ประกอบด้วย 5 แอตทริบิวต์ คือ

- **Outlook** แสดงสภาพภูมิอากาศ ประกอบด้วย 3 ค่า คือ Sunny, Overcast, Rain
- **Temperature** แสดงอุณหภูมิ ประกอบด้วย 3 ค่า คือ Hot, Mild, Cool
- **Humidity** แสดงค่าความชื้นในอากาศ ประกอบด้วย 2 ค่า คือ High, Normal
- **Windy** แสดงว่าเป็นวันที่ลมแรงหรือไม่ ประกอบด้วย 2 ค่า คือ Weak, Strong
- **Play** แสดงการจัดแข่งขันกีฬา ซึ่งเป็นคลาส ประกอบด้วย 2 ค่า คือ Yes, No

การสร้างโมเดล Decision Tree จะทำการคัดเลือกแอตทริบิวต์ที่มีความสัมพันธ์กับคลาสมากที่สุดขึ้นมาเป็นโหนดบนสุดของ Tree (Root Node) หลังจากนั้นก็จะหาแอตทริบิวต์ถัดไปเรื่อยๆ

ในการหาความสัมพันธ์ของแอตทริบิวต์นี้จะใช้ตัววัด ที่เรียกว่า Information Gain (IG) ขั้นตอนการทำมีดังนี้

1. คำนวณหาค่า Info (T) = $- p_{yes}\log_2 p_{yes} - p_{no}\log_2 p_{no}$

$$= -\frac{9}{14}\log_2\left(\frac{9}{14}\right)-\frac{5}{14}\log_2\left(\frac{5}{14}\right)=0.940$$

2. คำนวณค่า IG ของแอตทริบิวต์ Outlook, Temp, Humidity, Windy จะได้ดังนี้

Outlook

$$\text{Info}_{\text{Outlook=Sunny}}(T) = -\frac{2}{5}\log_2\left(\frac{2}{5}\right)-\frac{3}{5}\log_2\left(\frac{3}{5}\right)=0.971$$

$$\text{Info}_{\text{Outlook=Overcast}}(T) = -\frac{4}{4}\log_2\left(\frac{4}{4}\right)-0=0$$

$$\text{Info}_{\text{Outlook=Rain}}(T) = -\frac{3}{5}\log_2\left(\frac{3}{5}\right)-\frac{2}{5}\log_2\left(\frac{2}{5}\right)=0.971$$

$$\text{Info}_{\text{Outlook}}(T) = \left(\frac{5}{14}\times 0.971\right)+\left(\frac{4}{14}\times 0\right)+\left(\frac{5}{14}\times 0.971\right)=0.693$$

Gain (Outlook) = 0.940 – 0.693 = 0.247

Temp

$$\text{Info}_{\text{Temp=hot}}(T) = -\frac{2}{4}\log_2\left(\frac{2}{4}\right)-\frac{2}{4}\log_2\left(\frac{2}{4}\right)=1$$

$$\text{Info}_{\text{Temp=Mild}}(T) = -\frac{4}{6}\log_2\left(\frac{4}{6}\right)-\frac{2}{6}\log_2\left(\frac{2}{6}\right)=0.918$$

$$\text{Info}_{\text{Temp=Cool}}(T) = -\frac{3}{4}\log_2\left(\frac{3}{4}\right)-\frac{1}{4}\log_2\left(\frac{1}{4}\right)=0.623$$

$$\text{Info}_{\text{Temp}}(T) = \left(\frac{4}{14}\times 1\right)+\left(\frac{6}{14}\times 0.918\right)+\left(\frac{4}{14}\times 0.623\right)=0.857$$

Gain (Temp) = 0.940 – 0.857 = 0.083

Humidity

$$\text{Info}_{\text{Humidity=high}}(T) = -\frac{3}{7}\log_2\left(\frac{3}{7}\right)-\frac{4}{7}\log_2\left(\frac{3}{7}\right)=0.985$$

$$\text{Info}_{\text{Humidity=normal}}(T) = -\frac{6}{7}\log_2\left(\frac{6}{7}\right)-\frac{1}{7}\log_2\left(\frac{6}{7}\right)=0.5916$$

$$\text{Info}_{\text{Humidity}}(T) = \left(\frac{7}{14}\times 0.985\right)+\left(\frac{7}{14}\times 0.5916\right)=0.7883$$

Gain (Humidity) = 0.940 – 0.7883 = 0.151

Wind

$$\text{Info}_{\text{Wind=weak}}(T) = -\frac{6}{8}\log_2\left(\frac{6}{8}\right)-\frac{2}{8}\log_2\left(\frac{2}{8}\right)=0.8112$$

$$\text{Info}_{\text{Wind=strong}}(T) = -\frac{3}{6}\log_2\left(\frac{3}{6}\right)-\frac{3}{6}\log_2\left(\frac{3}{6}\right)=1$$

$$\text{Info}_{\text{Wind}}(T) = \left(\frac{8}{14}\times 0.8112\right)+\left(\frac{6}{14}\times 1\right)=0.892$$

Gain (Wind) = 0.940 – 0.892 = 0.048

จากการคำนวณจะเลือกค่า Gain ที่มากที่สุดนั้นก็คือ Outlook ทำให้ได้ว่า Outlook เป็น Root Node และจะเห็นว่าข้อมูลที่อยู่ในโหนดที่แอตทริบิวต์ Outlook = Overcast มีคลาสเดียวกันหมดคือ play = yes ดังนั้นโหนดนี้ไม่ต้องทำการแตกกิ่งต่อไปอีก แต่โหนดอื่นๆต้องทำการแตกกิ่งออกไปจนข้อมูลในแต่ละโหนดมีคลาสคำตอบเดียวกัน

3. จะทำการหา info (Outlook = Sunny) $= -\frac{2}{5}\log_2\left(\frac{2}{5}\right)-\frac{3}{5}\log_2\left(\frac{3}{5}\right)=0.971$

หาค่า $\text{info}_{\text{Temp}}$ (Outlook = Sunny)

$$\text{Info}_{\text{Temp=Hot}}(\text{Outlook = Sunny}) = 0 - \log_2 1 = 0$$

$$\text{Info}_{\text{Temp=Mild}}(\text{Outlook = Sunny}) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right)-\frac{1}{2}\log_2\left(\frac{1}{2}\right)=1$$

$$\text{Info}_{\text{Temp=Cool}}(\text{Outlook = Sunny}) = -1\log_2 1 - 0 = 0$$

$$\text{Info}_{\text{Temp}}(\text{Outlook = Sunny}) = \left(\frac{2}{5}\times 0\right)+\left(\frac{2}{5}\times 1\right)+\left(\frac{1}{5}\times 0\right)=0.4$$

Gain (Temp) = 0.971 – 0.4 = 0.571

หาค่า $\text{info}_{\text{Humidity}}$ (Outlook = Sunny)

$$\text{Info}_{\text{Humidity=High}}(\text{Outlook = Sunny}) = 0 - 1\log_2 1 = 0$$

$$\text{Info}_{\text{Humidity=Normal}}(\text{Outlook = Sunny}) = -1\log_2 1 - 0 = 0$$

$$\text{Info}_{\text{Humidity}}(\text{Outlook = Sunny}) = \left(\frac{3}{5}\times 0\right)+\left(\frac{2}{5}\times 0\right)=0$$

Gain (Humidity) = 0.971 – 0 = 0.971

หาค่า $\text{info}_{\text{Wind}}$ (Outlook = Sunny)

$$\text{Info}_{\text{Wind=Weak}}(\text{Outlook = Sunny}) = -\frac{1}{3}\log_2\left(\frac{1}{3}\right)-\frac{2}{3}\log_2\left(\frac{2}{3}\right)=0.918$$

$$\text{Info}_{\text{Wind=Strong}}(\text{Outlook = Sunny}) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right)-\frac{1}{2}\log_2\left(\frac{1}{2}\right)=1$$

$$\text{Info}_{\text{Wind}}(\text{Outlook = Sunny}) = \left(\frac{3}{5}\times 0.918\right)+\left(\frac{2}{5}\times 1\right)=0.9508$$

Gain (Wind) = 0.971 – 0.9508 = 0.0202

จากการคำนวณจะเลือกค่า Gain ที่มากที่สุดนั้นก็คือ Humidity ทำให้ได้ว่า Humidity เป็น Root Node จากขั้นตอนนี้จะเห็นว่าโหนด Outlook = Suuny และ Humidity = High มีคลาส no และ ข้อมูลที่อยู่โหนด Outlook = Sunny และ Humidity = Normal มีคลาส yes แสดงว่าสามารถหยุดทำการแตกกิ่งต่อจากโหนดเหล่านี้ได้

4. ต่อไปจะพิจารณา Info (Outlook = Rain) = $-\frac{3}{5}\log_2\left(\frac{3}{5}\right)-\frac{2}{5}\log_2\left(\frac{2}{5}\right)=0.971$

หาค่า $info_{Temp}$ (Outlook = Rain)

$Info_{Temp=Hot}$ (Outlook = Rain) = 0

$Info_{Temp=Mild}$ (Outlook = Rain) = $-\frac{2}{3}\log_2\left(\frac{2}{3}\right)-\frac{1}{3}\log_2\left(\frac{1}{3}\right)=0.918$

$Info_{Temp=Cool}$ (Outlook = Rain) = $-\frac{1}{2}\log_2\left(\frac{1}{2}\right)-\frac{1}{2}\log_2\left(\frac{1}{2}\right)=1$

$Info_{Temp}$(Outlook = Rain) = $\left(\frac{3}{5}\times 0.918\right)+\left(\frac{2}{5}\times 1\right)=0.9508$

Gain (Temp) = 0.971 – 0.9508 = 0.0202

หาค่า $info_{Humidity}$ (Outlook = Rain)

$Info_{Humidity=High}$ (Outlook = Rain) = $-\frac{1}{2}\log_2\left(\frac{1}{2}\right)-\frac{1}{2}\log_2\left(\frac{1}{2}\right)=1$

$Info_{Humidity=Normal}$ (Outlook = Rain) = $-\frac{2}{3}\log_2\left(\frac{2}{3}\right)-\frac{1}{3}\log_2\left(\frac{1}{3}\right)=0.918$

$Info_{Humidityl}$ (Outlook = Rain) = $\left(\frac{2}{5}\times 1\right)+\left(\frac{3}{5}\times 0.918\right)=0.950$

Gain (Humidity) = 0.971 – 0.950 = 0.021

หาค่า $info_{Wind}$ (Outlook = Rain)

$Info_{Wind=Weak}$ (Outlook = Rain) = $-1\log_2-0=0$

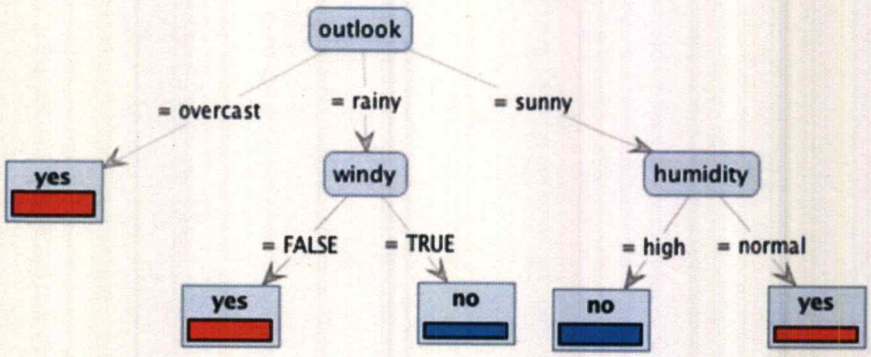
$Info_{Wind=Strong}$ (Outlook = Rain) = $0-1\log_2 1=0$

$Info_{Wind}$ (Outlook = Rain) = $\left(\frac{3}{5}\times 0\right)+\left(\frac{2}{5}\times 0\right)=0$

Gain (Wind) = 0.971 – 0 = 0.971

จากการคำนวณจะเลือกค่า Gain ที่มากที่สุดนั้นก็คือ Wind ทำให้ได้ว่า Wind เป็น Root Node จากขั้นตอนนี้จะเห็นว่าโหนด Outlook = Rain และ Wind = Strong มีคลาส no และ ข้อมูลที่อยู่โหนด Outlook = Rain และ Wind = Weak มีคลาส yes เมื่อโหนดลูกเป็นกลุ่มของ

ข้อมูลที่อยู่ในคลาสเดียวกันทั้งหมด จึงหยุดการคำนวณทั้งหมดนี้คือขั้นตอนการสร้างโมเดล decision tree จะได้ดังรูป



รูปที่ 2.4 แสดงโมเดล Decision Tree

2. Naïve Bayesian Classification

การจำแนกข้อมูลโดยใช้กฎของเบย์

- เป็นการจำแนกประเภทโดยใช้หลักสถิติในการพยากรณ์ความน่าจะเป็นของสมาชิก

- เป็นการจำแนกโดยนำความรู้เรื่องความน่าจะเป็น (Probability) เพื่อตัดสินใจในการจำแนกว่าข้อมูลที่พยากรณ์นั้นควรอยู่ในคลาสใด การจำแนกแบบนี้จะรู้จักกันในอีกชื่อคือ Naïve Bayesian

- เป็นอีกเทคนิคหนึ่งที่นำมาใช้ในการทำเหมืองข้อมูลสำหรับการจำแนกข้อมูลและสามารถคาดการณ์ผลลัพธ์ได้และสามารถอธิบายได้ด้วยการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระแต่ละตัวกับตัวแปรตามเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์

วิธี Naive Bayes หลักการของวิธีการนี้จะใช้การคำนวณความน่าจะเป็น

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

P (A|B) คือ ค่า conditional probability หรือค่าความน่าจะเป็นที่เกิดเหตุการณ์ B ขึ้นก่อนและจะมีเหตุการณ์ A ตามมา

P (A $\cap$ B) คือ ค่า joint probability หรือค่าความน่าจะเป็นที่เหตุการณ์ A และเหตุการณ์ B เกิดขึ้นร่วมกัน

P (B) คือ ค่าความน่าจะเป็นที่เหตุการณ์ B เกิดขึ้น

ในลักษณะเดียวกันเราจะเขียน P (B|A) หรือค่าความน่าจะเป็นที่เหตุการณ์ A เกิดขึ้นก่อนและเหตุการณ์ B เกิดขึ้นตามมาทีหลังได้เป็น

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

จากทั้ง 2 แบบจะเห็นว่ามีค่า P (A ∩ B) ที่เหมือนกันอยู่ดังนั้นเราสามารถเขียนสมการของ P(A ∩ B) ได้เป็นดังนี้

$$P(B \mid A) = \frac{P(A \mid B) \times P(B)}{P(A)}$$

และนี่คือสมการที่เรียกว่า Bayes Theorem หรือทฤษฎีของเบย์ ในการนำไปใช้งานทางด้าน Data Mining เปลี่ยนสัญลักษณ์ A และ B เสียใหม่ให้เป็น A และ C โดยที่ A คือ แอตทริบิวต์ (Attribute) และ C คือ ค่าคลาส (Class) ดังสมการด้านล่าง

$$P(C \mid A) = \frac{P(A \mid C) \times P(C)}{P(A)}$$

จากสมการของ Bayes จะมี 3 ส่วนที่สำคัญ คือ

Posterior probability หรือ P (C|A) คือ ค่าความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์เป็น A จะมีคลาส C

Likelihood หรือ P (A|C) คือ ค่าความน่าจะเป็นที่ข้อมูล Training Data ที่มีคลาส C และมีแอตทริบิวต์ A โดยที่ $A = a_1 \cap a_2 \ldots \cap a_M$ โดยที่ M คือจำนวนแอตทริบิวต์ใน Training Data

Prior probability หรือ P (C) คือ ค่าความน่าจะเป็นของคลาส C

แต่การที่แอตทริบิวต์ $A = a_1 \cap a_2 \ldots \cap a_M$ ที่เกิดขึ้นใน Training Data อาจจะมีจำนวนน้อยมากหรือไม่มีรูปแบบของแอตทริบิวต์แบบนี้เกิดขึ้นเลย ดังนั้นจึงได้ใช้หลักการที่ว่าแต่ละแอตทริบิวต์เป็น independent ต่อกันทำให้สามารถเปลี่ยนสมการ P (A|C) ได้เป็น

$$P(A \mid C) = P(a_1 \mid C) \times P(a_2 \mid C) \times \ldots \times P(a_M \mid C)$$

ตัวอย่าง

การสร้างเป็นโมเดล Naive Bayes ในนี้จะใช้ข้อมูลเดียวกันกับ Decision Tree นั่นคือข้อมูล weather

P (Play = yes) = 0.64

P (Play = No) = 0.36

ตารางที่ 2.2 ตัวอย่างข้อมูล วิธี Naïve Bayes

Attribute	play = yes	play = no
Outlook = Sunny	0.22	0.6
Outlook = Overcast	0.45	0
Outlook = Rainy	0.33	0.4
Temperature = Hot	0.22	0.4
Temperature = Mild	0.45	0.4
Temperature = Cool	0.33	0.2
Humidity = High	0.33	0.8
Humidity = Normal	0.67	0.2
Windy = Strong	0.33	0.6
Windy = Weak	0.67	0.4

จากตารางนี้ก็คือโมเดลของ Naïve Bayes ต่อมาจะมาดูการใช้งานโมเดลนี้ จะเอาข้อมูล instance แรกจากในตารางที่ 2.2 มาทำนายด้วยโมเดล Naïve Bayes แรกนี้ประกอบด้วย

แอตทริบิวต์ **outlook** = sunny

แอตทริบิวต์ **temperature** = hot

แอตทริบิวต์ **humidity** = high

แอตทริบิวต์ **windy** = Weak

ต้องคำนวณค่าความน่าจะเป็นที่มีแอตทริบิวต์เหล่านี้แล้วตอบคลาส play = yes ได้ดังนี้

- P (play =yes|A) = P (outlook = sunny|play = yes) * P (temperature = hot|play = yes) * P (humidity = high|play = yes) * P (windy = Weak|play = yes) * P (play = yes)

= 0.22*0.22*0.33*0.67*0.64 = 0.0068

หลังจากนั้นจะคำนวณค่าความน่าจะเป็นที่มีแอตทริบิวต์เหล่านี้แล้วตอบคลาส play = no ได้ดังนี้

- P (play = no|A) = P (outlook = sunny|play = no) * P (temperature = hot|play = no) * P (humidity = high|play = no) * P (windy = Weak|play = no) * P (play = no)

= 0.60*0.40*0.80*0.40*0.36 = 0.0276

เมื่อเปรียบเทียบค่าความน่าจะเป็นที่ได้จาก 2 คลาสแล้วพบว่าค่า P (play = no|A) =0.0276 มีค่ามากกว่า P (play = yes|A) = 0.0068 ดังนั้นโมเดลของจึงทำนายว่าข้อมูล instance นี้มีค่าคลาส play = no

3. K-Nearest Neighbors (K-NN) คือ วิธีการในการจัดแบ่งคลาส โดยเทคนิคนี้จะตัดสินใจว่าคลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ๆ ได้บ้างโดยการตรวจสอบจำนวนบางจำนวน ("K" ใน K-Nearest Neighbor) ของกรณีหรือเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด หลักการทำงานของ K-NN จะคล้ายกับการแบ่งกลุ่ม (Clustering) คือ จะทำการวัดระยะห่างระหว่างข้อมูลที่ต้องการทำนายกับข้อมูลที่อยู่ใกล้เคียงในจำนวน K ตัว สำหรับคำตอบที่ทำนายได้คือคลาสที่พบมากที่สุดของข้อมูลที่เป็นเพื่อนบ้านใน k ตัว

ขั้นตอนการทำงานของวิธี K-NN

1. กำหนดค่า K (K = จำนวนข้อมูลใกล้เคียงที่ต้องการนำมาพิจารณา)
2. หาค่าระยะห่างระหว่างข้อมูลกับข้อมูลใหม่
3. กำหนดคลาสให้กับข้อมูลใหม่ โดยพิจารณาจากข้อมูลที่มีระยะห่างน้อยที่สุด K ตัว

ตัวอย่าง

ตารางที่ 2.3 ตัวอย่างข้อมูล วิธี K-Nearest Neighbors

No	Petallength	Petalwidth	Class
1	1.40	0.20	Iris-setosa
2	1.70	0.30	Iris-setosa
3	1.60	0.40	Iris-setosa
4	1.90	0.40	Iris-setosa
5	1.00	0.20	Iris-setosa
6	4.70	1.40	Iris-versicolor
7	4.50	1.50	Iris-versicolor
8	4.00	1.30	Iris-versicolor
9	3.50	1.00	Iris-versicolor
10	4.20	1.50	Iris-versicolor
11	5.80	2.20	Iris-virginca
12	6.60	2.10	Iris-virginca
13	4.50	1.70	Iris-virginca
14	6.30	1.80	Iris-virginca
15	6.70	2.00	Iris-virginca

มีข้อมูลเข้ามาใหม่ควรจะอยู่ Class ไหน

No	Petallength	Petalwidth	Class
16	4.25	1.60	?

ขั้นตอนที่ 1 : กำหนดค่า K ในนี้กำหนดให้ K=5

ขั้นตอนที่ 2 : คำนวณหาระยะห่างจากข้อมูลใหม่กับข้อมูลทั้งหมด โดยใช้สูตร Euclidean ได้ดังนี้

$$D=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$$

เช่น

Point No1 (p1) : (1.40,0.20)

Point No 16 (p16) : (4.25,1.60)

Distance (p1,p16) = $\sqrt{(1.40-4.25)^2+(0.20-1.60)^2}$ = 3.18

Point No2 (p2) : (1.70,0.30)

Point No 16 : (4.25,1.60)

Distance (p2,p16) = $\sqrt{(1.70-4.25)^2+(0.30-1.60)^2}$ = 2.86

จะได้ว่า

ตารางที่ 2.4 ข้อมูลหลังจากคำนวณ

Point		(4.25, 1.60)
1	(1.40, 0.20)	3.18
2	(1.70, 0.30)	2.86
3	(1.60, 0.40)	2.91
4	(1.90, 0.40)	2.64
5	(1.00, 0.20)	3.54
6	(4.70, 1.40)	0.49
7	(4.50, 1.50)	0.27
8	(4.00, 1.30)	0.39
9	(3.50, 1.00)	0.96
10	(4.20, 1.50)	0.11
11	(5.80, 2.20)	1.66
12	(6.60, 2.10)	2.4
13	(4.50, 1.70)	0.27

Point		(4.25, 1.60)
14	(6.30, 1.80)	2.06
15	(6.70, 2.00)	2.48

เรียงลำดับข้อมูลตามระยะห่างจากน้อยสุดไปหามากสุด จะได้ว่า

ตารางที่ 2.5 ข้อมูลที่เรียงลำดับจากน้อยไปหามาก

Point		(4.25, 1.60)
10	(4.20, 1.50)	0.11
7	(4.50, 1.50)	0.27
13	(4.50, 1.70)	0.27
8	(4.00, 1.30)	0.39
6	(4.70, 1.40)	0.49
9	(3.50, 1.00)	0.96
11	(5.80, 2.20)	1.66
14	(6.30, 1.80)	2.06
12	(6.60, 2.10)	2.4
15	(6.70, 2.00)	2.48
4	(1.90, 0.40)	2.64
2	(1.70, 0.30)	2.86
3	(1.60, 0.40)	2.91
1	(1.40, 0.20)	3.18
5	(1.00, 0.20)	3.54

เลือก 5 อันดับแรก จะได้ดังนี้

ตารางที่ 2.6 ระยะทางที่น้อยที่สุด 5 อันดับแรก

Point		(4.25, 1.60)	Class
10	(4.20, 1.50)	0.11	Iris-versicolor
7	(4.50, 1.50)	0.27	Iris-versicolor
13.	(4.50, 1.70)	0.27	Iris-virginca
8	(4.00, 1.30)	0.39	Iris-versicolor
6	(4.70, 1.40)	0.49	Iris-versicolor

จากการเรียงลำดับ จะเห็นได้ว่า Iris-versicolor มีจำนวนมากที่สุดจึงสรุปได้ว่า No.16 อยู่ที่ Class Iris-versicolor

2.1.3 ตัววัดประสิทธิภาพของโมเดล

- Confusion Matrix เป็นการทำนายไว้ว่าถูกผิดเท่าไร

- Precision ดูสิ่งที่เรา Predict ออกมา แล้วทายถูกได้กี่เปอร์เซ็นต์ผลการคำนวณได้ทายถูกกี่ตัวจากผลที่ได้ออกมาจากตัวอย่าง

- Recall จำนวนที่ทำนายถูกกี่ตัว : ความถูกต้องของแต่ละคำตอบ ในการทำนายว่าสิ่งที่ต้องการคืออะไร ต้องการทำนายจำนวนคนที่ป่วยจะดูได้จาก Recall ออกมาว่าตัวไหนให้คำตอบมากกว่าน้อยกว่า

อีกตัวจะมาช่วยในการทำนายชื่อ

- F-Measure ช่วยหาค่าเฉลี่ยของ Precision และ Recall

- Accuracy จำนวนข้อมูลที่ทำนายถูกของทุกคลาส

- ROC Curve มีค่าเข้าใกล้ 1 จะแสดงว่ามีประสิทธิภาพดีกว่า จากกราฟ

Validation การแบ่งข้อมูลเพื่อทดสอบประสิทธิภาพของโมเดล วิธีการทดสอบโมเดลแบ่งออกเป็น 3 ตัว

- Self-consistency Test (Use Training Set) เอา Training Data มาเป็นตัว Test เลย คือ ใช้ข้อมูลเดิมแล้วมันมีความถูกต้องมากน้อยแค่ไหน

- Split test แบ่งออกเป็น 2 ส่วน คือ เป็นโมเดล และส่วนทดสอบ เช่น 70% สร้างโมเดล 30% ใช้ทดสอบ หรือ 80 : 20 ใช้ข้อมูล 2 ชุด เป็นTraining Data สำหรับสร้าง Model และ Testing Data สำหรับทดสอบ ถ้า Sampling มาดีก็ได้ข้อมูลทดสอบดี Split Test จะดีเมื่อมีข้อมูลมากๆหมื่นหรือแสนตัวขึ้นไป

- Cross - Validation test ใช้ค่อนข้างเยอะเหมือนกัน การทำ Split หลายๆรอบ การทำคล้ายๆ กับ Split แต่แบ่งจำนวน N ชุดเท่าๆกัน เช่น N=5, N=10 แล้วทำงานโดยการสร้างโมเดลทั้งหมด N ตัว จนครบ Split Test ทำรอบเดียวแล้วจบไป คือเก็บ 1, 2 ไว้ แต่ใน Cross จะเอา 3 ไปเป็นตัวทดสอบ แล้วใส่ 2 กลับเข้าไป แล้วเอา 1 เป็นตัวทดสอบ จะพบว่าทุกตัวจะถูกนำมาทดสอบหมด เราจะได้ความถูกต้องเฉลี่ยออกมาในแต่ละรอบ แต่ไม่เหมาะสมกับการทำแบบนี้กับกรณีที่ข้อมูลมากๆไม่เหมาะ

2.1.4 ประโยชน์ของ Data Mining

1. ช่วยชี้แนวทางการตัดสินใจและคาดการณ์ผลลัพธ์ที่จะได้จากการตัดสินใจ
2. เพิ่มความเร็วในการวิเคราะห์ฐานข้อมูลขนาดใหญ่
3. ค้นหาส่วนประกอบที่ซ่อนอยู่ภายในเอกสาร รวมถึงความสัมพันธ์ของส่วนประกอบต่าง ๆ ด้วย
4. เชื่อมโยงหน่วยงานต่าง ๆ ภายในองค์กร
5. การจัดกลุ่มข้อมูล เช่น จัดกลุ่มลูกค้าทั้งหมดของบริษัทประกันภัยที่ประสบอุบัติเหตุ ลักษณะเดียวกันเพื่อดำเนินการต่างๆ ตามนโยบายของบริษัท

2.2 ความรู้ทั่วไปด้านประกันภัยรถยนต์

การประกันภัยรถยนต์จำแนกได้ 2 ประเภท ได้แก่

2.2.1 การประกันภัยรถยนต์ภาคบังคับ (Compulsory Third Party Insurance)

หมายถึง การที่เจ้าของรถแต่ละคันต้องจัดให้มีประกันภัยตามความคุ้มครองที่กำหนดไว้ในพระราชบัญญัติคุ้มครองผู้ประสบภัยจากรถ พ.ศ. 2535 ได้แก่ การประกันภัยคุ้มครองผู้ประสบภัยจากรถ หรือเรียกสั้น ๆ ว่า "การประกันภัย พ.ร.บ." เป็นการประกันภัยตามพระราชบัญญัติคุ้มครองผู้ประสบภัยจากรถ พ.ศ. 2535 ที่กำหนดให้เจ้าของรถซึ่งใช้รถหรือมีรถไว้เพื่อใช้ต้องจัดให้มีการประกันภัย พ.ร.บ. (เว้นแต่ รถที่ พ.ร.บ.ฯ ระบุยกเว้นไว้ เช่น รถของสำนักพระราชวัง รถสำหรับเฉพาะองค์พระมหากษัตริย์ รถของกระทรวง ทบวง กรม เป็นต้น) มิเช่นนั้นจะมีโทษปรับไม่เกิน 10,000 บาท

2.2.2 การประกันภัยรถยนต์ภาคสมัครใจ (Voluntary Motor Insurance)

หมายถึง การที่เจ้าของรถแต่ละคันตัดสินใจทำประกันภัยรถยนต์ของตนด้วยความสมัครใจเพื่อคุ้มครองความเสียหายที่อาจจะเกิดขึ้นกับตัวรถ และ/หรือความรับผิดตามกฎหมายของ ผู้เอาประกันภัยที่มีต่อบุคคลภายนอก

การประกันภัยรถยนต์ภาคสมัครใจ

แบ่งความคุ้มครองเป็น 5 ประเภทดังนี้

การประกันภัยรถยนต์ประเภท 1 ให้ความคุ้มครองที่ครอบคลุมและสูงที่สุดเมื่อเทียบกับประเภทอื่น ๆ โดยให้ความคุ้มครอง 3 หมวดหลัก ดังนี้

1. หมวดการคุ้มครองความรับผิดต่อบุคคลภายนอก

- ความเสียหายต่อชีวิต ร่างกาย หรืออนามัยของบุคคลภายนอก เฉพาะส่วนที่เกินวงเงินสูงสุดตามการประกันภัยรถยนต์ภาคบังคับ (พ.ร.บ.) ทั้งนี้ ไม่รวมถึงผู้ขับขี่ที่เป็นฝ่ายจะต้องรับผิดตามกฎหมาย ตลอดจนลูกจ้างในทางการที่จ้าง บิดา มารดา คู่สมรส และบุตรของผู้ขับขี่นั้นหรือเป็นผู้เก็บรักษา ควบคุม ครอบครอง
- ความเสียหายต่อทรัพย์สินบุคคลภายนอก ทั้งนี้ ไม่รวมถึงที่ผู้เอาประกันภัย ผู้ขับขี่รถยนต์คันเอาประกันภัยในขณะเกิดอุบัติเหตุ คู่สมรส บิดา มารดา บุตรของผู้เอาประกันภัยหรือผู้ขับขี่นั้นเป็นเจ้าของ หรือเป็นผู้เก็บรักษา ควบคุมครอบครอง

2. หมวดการคุ้มครองรถยนต์สูญหายไฟไหม้ไม่ว่าจะเสียหายทั้งคันหรือบางส่วน เช่น กระจกมองข้างรถถูกขโมย สะเก็ดไฟมาโดนรถไฟไหม้ทั้งคัน เป็นต้น

3. หมวดการคุ้มครองความเสียหายต่อรถยนต์ที่เอาประกันภัยอันเกิดจากอุบัติเหตุ เช่น กันชนหน้ารถแตกหักจากการขับรถชนต้นไม้ ประตูรถได้รับความเสียหายจากการถูกเฉี่ยวชนกิ่งไม้หล่นใส่หลังคารถ เป็นต้น

การประกันภัยรถยนต์ประเภท 2 ให้ความคุ้มครองรองลงมาจากประเภท 1 โดยมีความคุ้มครองเพียง 2 หมวดหลักดังนี้

1. หมวดการคุ้มครองความรับผิดต่อบุคคลภายนอก

- ความเสียหายต่อชีวิต ร่างกาย หรืออนามัยของบุคคลภายนอก เฉพาะส่วนที่เกินวงเงินสูงสุดตามการประกันภัยรถยนต์ภาคบังคับ (พ.ร.บ.) ทั้งนี้ ไม่รวมถึงผู้ขับขี่ที่เป็นฝ่ายจะต้องรับผิดตามกฎหมาย ตลอดจนลูกจ้างในทางการที่จ้าง บิดา มารดา คู่สมรส และบุตรของผู้ขับขี่นั้น
- ความเสียหายต่อทรัพย์สินบุคคลภายนอก ทั้งนี้ ไม่รวมถึงที่ผู้เอาประกันภัย ผู้ขับขี่รถยนต์คันเอาประกันภัยในขณะเกิดอุบัติเหตุ คู่สมรส บิดามารดา บุตรของผู้เอาประกันภัยหรือผู้ขับขี่นั้นเป็นเจ้าของ หรือเป็นผู้เก็บรักษา ควบคุมครอบครอง

2. หมวดการคุ้มครองรถยนต์สูญหาย ไฟไหม้

การประกันภัยรถยนต์ประเภท 3 ให้ความคุ้มครองเฉพาะหมวดการคุ้มครองความรับผิดต่อบุคคลภายนอกเพียงอย่างเดียว ซึ่งประกอบด้วย ความเสียหายต่อชีวิต ร่างกาย หรืออนามัยของบุคคลภายนอก เฉพาะส่วนที่เกินวงเงินสูงสุดตามการประกันภัยรถยนต์ภาคบังคับ (พ.ร.บ.) และความเสียหายต่อทรัพย์สินบุคคลภายนอก

การประกันภัยรถยนต์ประเภท 4 ถูกพัฒนาขึ้นเพื่อให้ความคุ้มครองต่อทรัพย์สินของบุคคลภายนอกซึ่งไม่ได้รับความคุ้มครองในการประกันภัย พ.ร.บ. แต่จำนวนเงินจำกัดความรับผิดตามกฎหมายสำหรับความเสียหายต่อทรัพย์สินของบุคคลภายนอก ในการประกันภัยประเภทนี้จำกัดจำนวนเงินไว้ไม่เกิน 100,000 บาท ต่ออุบัติเหตุแต่ละครั้งเท่านั้น

การประกันภัยรถยนต์ประเภท 5 การประกันภัยรถยนต์แบบคุ้มครองเฉพาะภัย หรือที่เรียกกันว่าการประกันภัยรถยนต์ประเภท 5 ซึ่งเปิดความคุ้มครองให้แต่ละบริษัทประกันวินาศภัยสามารถเลือกขยายความคุ้มครองได้โดยกำหนดบังคับให้ทุกกรมธรรม์ประกันภัยต้องมีความคุ้มครองความรับผิดตามกฎหมายต่อบุคคลภายนอกเป็นความคุ้มครองขั้นต่ำในปัจจุบันมีแบบประกันภัยที่เป็นที่นิยม 2 แบบ คือ

ก. แบบที่ขยายความคุ้มครองเพิ่มขึ้นจากการประกันภัยรถยนต์ประเภท 2 ซึ่งเรียกกันทั่วไปว่า 2+ ให้ความคุ้มครองคล้ายคลึงประเภท 2 และเพิ่มความเสียหายต่อตัวรถยนต์ในจำนวนเงินที่จำกัดอันเนื่องมาจากการชนกับยานพาหนะทางบกและต้องมีคู่กรณีเท่านั้น และ

ข. แบบที่ขยายความคุ้มครองเพิ่มขึ้นจากการประกันภัยรถยนต์ประเภท 3 ซึ่งเรียกกันทั่วไปว่า 3+ ให้ความคุ้มครองคล้ายคลึงประเภท 3 และเพิ่มความเสียหายต่อตัวรถยนต์ในจำนวนเงินที่จำกัดอันเนื่องมาจากการชนกับยานพาหนะทางบกและต้องมีคู่กรณีเท่านั้น

นอกเหนือจากความคุ้มครองหลักในการประกันภัยทั้ง 5 ประเภทข้างต้นแล้ว ผู้เอาประกันภัยยังสามารถเลือกซื้อความคุ้มครองเพิ่มเติมตามเอกสารแนบท้ายได้อีก 3 แบบ คือ

- การประกันภัยอุบัติเหตุส่วนบุคคล-ร.ย.01 ให้ความคุ้มครองการเสียชีวิต สูญเสียอวัยวะ ทุพพลภาพถาวร ทุพพลภาพชั่วคราวของผู้ขับขี่และผู้โดยสารซึ่งอยู่ในหรือกำลังขับขี่ หรือกำลังขึ้น หรือกำลังลงจากรถยนต์คันเอาประกันภัย ซึ่งเป็นผลมาจากอุบัติเหตุ
- ค่ารักษาพยาบาล-ร.ย.02 ให้ความคุ้มครองค่ารักษาพยาบาลของผู้ขับขี่และผู้โดยสารซึ่งอยู่ในหรือกำลังขับขี่ หรือกำลังขึ้น หรือกำลังลงจากรถยนต์คันเอาประกันภัย ซึ่งเป็นผลมาจากอุบัติเหตุ
- การประกันตัวผู้ขับขี่-ร.ย.03 ให้ความคุ้มครองผู้เอาประกันภัย หรือบุคคลใดซึ่งขับขี่รถยนต์โดยความยินยอมของผู้เอาประกันภัยกรณีที่เกิดอุบัติเหตุจนเป็นเหตุให้บุคคลดังกล่าวถูกควบคุมตัวในคดีอาญา ซึ่งบริษัทประกันภัยจะดำเนินการโดยไม่ชักช้าตามจำนวนเงินที่พนักงานสอบสวน พนักงานอัยการหรือศาลกำหนดจนกว่าคดีจะถึงที่สุด

ตารางที่ 2.7 ตารางเปรียบเทียบประเภทประกันภัยรถยนต์

	1	2+	2	3+	3
คุ้มครองคู่กรณี					
บุคคล					
ทรัพย์สิน					
คุ้มครองตัวรถ					
ไฟไหม้					
รถหาย					
การชน					
คุ้มครองคนในรถ					
อุบัติเหตุการชน					
ค่ารักษาพยาบาล					
ประกันตัวผู้ขับขี่					

หมายเหตุ : จากอุบัติเหตุที่มีคู่กรณีเป็นยานพาหนะทางบกซึ่งหมายถึง รถเดินด้วยกำลังเครื่องยนต์โดยใช้พลังงานเชื้อเพลิง เช่น น้ำมันก๊าซหรือกำลังไฟฟ้าและรวมไปถึงรถไฟรถราง

2.3 คำศัพท์ทางด้านประกันภัย

- **คำขอเอาประกันภัย (Application Form)**

ใบคำขอเอาประกันภัยเป็นเอกสารสำคัญที่ทำขึ้นเพื่อแสดงความประสงค์ว่าจะเอาประกันภัยอย่างใดอย่างหนึ่งไว้กับผู้รับประกันภัยผู้เอาประกันภัยมีหน้าที่ที่จะต้องเปิดเผยความจริง (Disclosure) โดยไม่ต้องรอให้ผู้รับประกันภัย สอบถามและหากมีข้อความสอบถามใดๆผู้เอาประกันภัยจะต้องตอบตามความจริงทั้งหมด(Representation) มิฉะนั้นสัญญาประกันภัยอาจตกเป็นโมฆียะ ซึ่งผู้รับประกันภัยสามารถบอกเลิกได้

- **กรมธรรม์ (Policy)**

เอกสารซึ่งออกโดยผู้รับประกันภัย (The insurer) โดยมีข้อความตรงกับความประสงค์ในใบคำขอ เพื่อใช้เป็นหลักฐานต่อไปซึ่งระบุถึงสาระสำคัญของข้อตกลง เงื่อนไข และความคุ้มครองตามสัญญาประกันภัย

- **ผู้รับประกันภัย** (The Insurer)

คู่สัญญาประกันภัย (โดยทั่วไปคือบริษัทประกันภัย) ที่ได้รับใบอนุญาตตามกฎหมายจากกรมการประกันภัย กระทรวงพาณิชย์ ผู้รับประกันภัยมีสิทธิในการรับเบี้ยประกันภัย และมีหน้าที่พิจารณารับประกันภัย ชดใช้ค่าสินไหมทดแทนเมื่อเกิดวินาศภัยขึ้นตามที่ระบุไว้ในสัญญา ในการชดใช้นั้น อาจชดใช้เป็นเงินสด การซ่อมแซมให้อยู่ในสภาพเดิม หรือการหาของชิ้นใหม่มาแทนที่ได้รับความเสียหายก็ได้

- **ผู้เอาประกันภัย** (The Insured)

คู่สัญญาประกันภัยซึ่งมีหน้าที่เปิดเผยข้อความจริงต่อผู้รับประกันภัย ตลอดจนมีหน้าที่ชำระเบี้ย ประกันภัย และเมื่อเกิดความเสียหายขึ้นในส่วนที่เอาประกันภัยไว้ ผู้เอาประกันภัยก็มีสิทธิในการเรียกร้องค่าสินไหมทดแทน ตามความเสียหายที่เกิดขึ้นจริง

- **ค่าสินไหมทดแทน** (Claim Amount)

ความเสียหายที่ผู้เอาประกันภัยเรียกร้องให้ผู้รับประกันภัยชดใช้ โดยความเสียหายดังกล่าวเป็นผลมาจากภัยตามที่ระบุไว้ในกรมธรรม์ และมีจำนวนตามที่เสียหายจริง

- **เบี้ยประกันภัย** (Premium)

จำนวนเงินที่ผู้เอาประกันภัยจะต้องส่งใช้ให้กับผู้รับประกันภัย เนื่องจากสัญญาประกันภัยเป็นสัญญา ต่างตอบแทน ถ้าผู้เอาประกันภัยไม่ชำระเบี้ยประกันภัย และเกิดความเสียหายขึ้น ผู้รับประกันภัยก็อาจปฏิเสธการจ่ายค่าสินไหมทดแทนได้จนกว่าผู้เอาประกันภัยจะชำระเบี้ยประกันภัยตามหน้าที่ของตน

- **ทุนเอาประกันภัยหรือจำนวนเงินที่เอาประกันภัย** (Sum Insured)

จำนวนเงินสูงสุดที่ผู้รับประกันภัยจะต้องชดใช้ เมื่อมีความเสียหายเกิดขึ้นตามสัญญา

- **ค่าเสียหายส่วนแรก** (Deductible)

ค่าใช้จ่ายในส่วนที่ผู้เอาประกันภัย (The Insured) ต้องรับผิดชอบเองในความเสียหายที่เกิดขึ้นแต่ละครั้ง เช่นหากกรมธรรม์ประกันภัยรถยนต์ มีการกำหนดความเสียหายส่วนแรกไว้ที่ 2,000 บาท/ครั้ง ในกรณีที่เกิดอุบัติเหตุขึ้นในแต่ละครั้ง หากความเสียหาย เท่ากับ 2,000 บาท หรือน้อยกว่าจะไม่ได้รับการชดเชยจากทางบริษัทประกันภัย หากแต่ผู้เอาประกันภัยจะต้องเป็นผู้รับผิดชอบเอง หากความเสียหายมากกว่า 2,000 บาท ผู้เอาประกันภัยจะจ่ายเพียงแค่ 2,000 บาท เท่านั้น ทั้งนี้บริษัทประกันภัยจะเป็นผู้รับผิดชอบส่วนที่เพิ่มขึ้นเอง ค่าเสียหายส่วนแรกจะมีส่วนในการทำให้มูลค่าเบี้ยประกันภัยลดลงได้ตามจำนวนที่ระบุ ทั้งนี้ ค่าเสียหายส่วนแรก ส่วนหนึ่งมีขึ้นเพื่อให้ผู้เอาประกันภัยมีความระมัดระวังมากขึ้น (เพราะการเกิดอุบัติเหตุและความเสียหายขึ้นนั้น ย่อมหมายถึงการเสียค่าเสียหายส่วนแรกด้วย)

- **การต่ออายุ (RENEWAL)**

  การดำเนินการเพื่อให้กรมธรรม์มีผลบังคับต่อเนื่องสืบไปจากวันที่กรมธรรม์สิ้นผลบังคับ

ซึ่งในบทที่ 3 จะกล่าวถึงวิธีการดำเนินการ มีวิธีการคัดเลือกข้อมูล วิธีการจัดการข้อมูลและการลดขนาดของข้อมูล

บทที่ 3

วิธีการดำเนินงานวิจัย

ในบทที่ 3 นี้จะกล่าวถึงการรวบรวมข้อมูล ขั้นตอนในคัดเลือกข้อมูล เทคนิคที่จะเอามาใช้และการพยากรณ์ โดยมีขั้นตอน ดังนี้

3.1 ลักษณะข้อมูลและการเตรียมข้อมูล

งานวิจัยครั้งนี้ได้รวบรวมข้อมูลลูกค้าโดยมีข้อมูลต่างๆ เช่น State code, Policy Number, model, make, Body_Type, age, province, tons, passgr, noseat, cc, vehcls, cover และ type เป็นต้น มีทั้งหมด 5,000 ข้อมูล ข้อมูลที่จัดเก็บมีทั้งหมด 15 แอตทริบิวต์ ซึ่งข้อมูลพวกนี้จะถูกจัดเก็บอยู่ในลักษณะของไฟล์ Excel

3.2 การคัดเลือกข้อมูลและการลดขนาดของข้อมูล

3.2.1 การคัดเลือกข้อมูล

การคัดเลือกข้อมูล ผู้วิจัยจะทำการเลือกข้อมูลดังนี้

1. จะเลือก State code ที่เป็น IF จะทำการคัด CA (Cancle) จะไม่นำข้อมูลที่มีการยกเลิกมาใช้ในการวิจัยครั้งนี้

2. จะไม่เอาข้อมูลที่เป็นประกันภัยภาคบังคับ (พ.ร.บ.) มาทำในการวิเคราะห์ครั้งนี้เพราะรถทุกคันต้องทำประกันภัยภาคบังคับทุกคัน

3. งานวิจัยนี้จะใช้ประเภทรถ Saloon, Van/Bus และ Pick Up

4. ผู้วิจัยจะเลือกน้ำหนักที่ไม่เกิน 3 ตัน

5. ได้ทำการปรับเปลี่ยนช่อง Capacity ให้อยู่ในลักษณะเดียวกันทั้งหมด

3.2.2 การลดขนาดข้อมูล

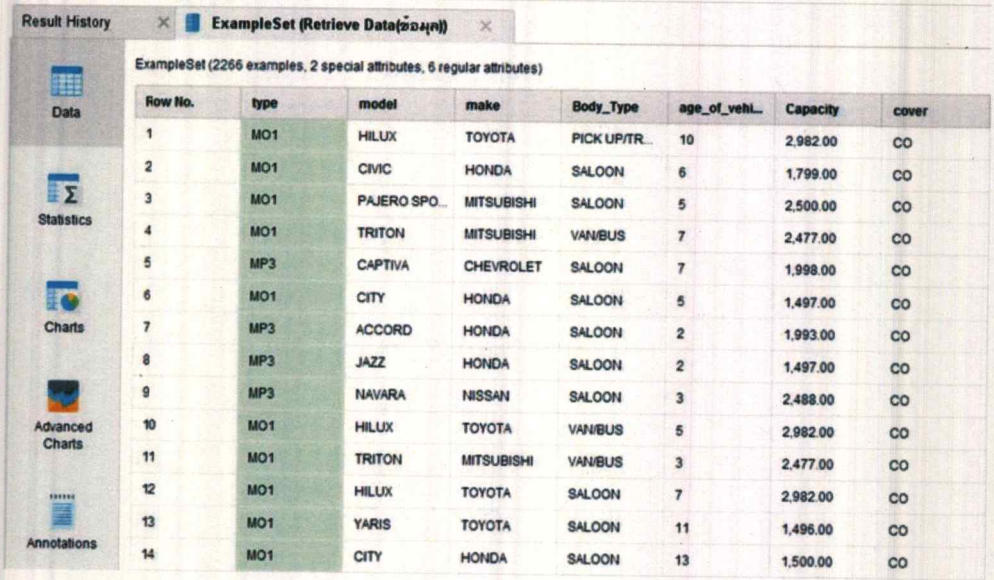
การลดขนาดข้อมูล เพื่อให้ได้ข้อมูลที่สามารถนำเข้าไปทดสอบในโปรแกรม RapidMiner จึงทำการลดขนาดข้อมูลที่ไม่จำเป็นออกไปจากที่มี 15 แอตทริบิวต์ จะทำการลดแอตทริบิวต์เพราะบางแอตทริบิวต์ไม่ส่งผลต่อการวิเคราะห์ทำให้เหลือแอตทริบิวต์ที่จำเป็นเพียง 8 แอตทริบิวต์ ได้แก่ Policy Number, model, make, Body_Type, age, Capacity, cover และ type จากการคัดเลือกข้อมูลทำให้เหลือข้อมูลเพียง 2,134 ข้อมูล จากนั้นนำข้อมูลมา save เป็นไฟล์ .csv เพื่อที่จะทำการทดสอบในโปรแกรม RapidMiner ต่อไป

3.3 การเลือกเทคนิค

จากข้อมูลที่มีอยู่นั้นงานวิจัยนี้จะทำการจำแนกข้อมูลซึ่งจะต้องมี Training Data และ Testing Data และทำการจำแนกข้อมูลทำให้เลือก Classification (สามารถอ่านเพิ่มเติมได้ที่บท 2) ซึ่งในเทคนิคนี้จะมีหลายวิธีดังนี้ 1.Decision Tree 2.Naïve Bayesian และ 3.K-NN

3.4 ขั้นตอนการดำเนินงาน

นำข้อมูลเข้าไปในโปรแกรมRapidจะได้ดังนี้



Result History | ExampleSet (Retrieve Data(ข้อมูล))

Data | Statistics | Charts | Advanced Charts | Annotations

ExampleSet (2266 examples, 2 special attributes, 6 regular attributes)

Row No.	type	model	make	Body_Type	age_of_vehi...	Capacity	cover
1	MO1	HILUX	TOYOTA	PICK UP/TR...	10	2,982.00	CO
2	MO1	CIVIC	HONDA	SALOON	6	1,799.00	CO
3	MO1	PAJERO SPO...	MITSUBISHI	SALOON	5	2,500.00	CO
4	MO1	TRITON	MITSUBISHI	VAN/BUS	7	2,477.00	CO
5	MP3	CAPTIVA	CHEVROLET	SALOON	7	1,998.00	CO
6	MO1	CITY	HONDA	SALOON	5	1,497.00	CO
7	MP3	ACCORD	HONDA	SALOON	2	1,993.00	CO
8	MP3	JAZZ	HONDA	SALOON	2	1,497.00	CO
9	MP3	NAVARA	NISSAN	SALOON	3	2,488.00	CO
10	MO1	HILUX	TOYOTA	VAN/BUS	5	2,982.00	CO
11	MO1	TRITON	MITSUBISHI	VAN/BUS	3	2,477.00	CO
12	MO1	HILUX	TOYOTA	SALOON	7	2,982.00	CO
13	MO1	YARIS	TOYOTA	SALOON	11	1,496.00	CO
14	MO1	CITY	HONDA	SALOON	13	1,500.00	CO

รูปที่ 3.1 ข้อมูลที่นำเข้า

การนำเข้าข้อมูลมีขั้นตอนดังนี้ กด Add Data >> My Computer >> เลือกไฟล์งาน >> Next >> Next >> ทำการเปลี่ยนค่าให้ Number Policy เป็น id และ type เป็น label >> Next >>Finish >> โปรแกรมจะแสดงหน้าจอดังรูปที่ 3.1

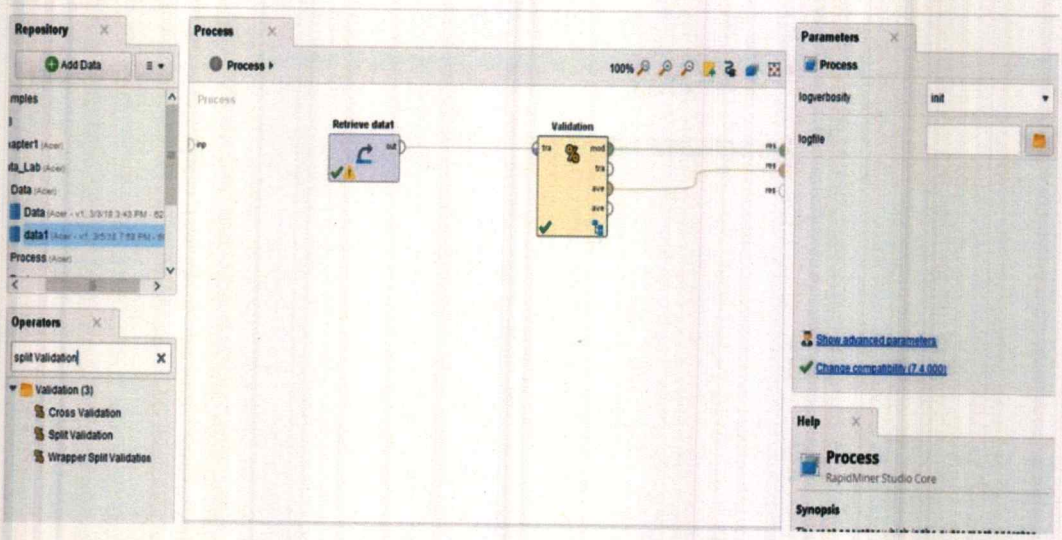
จากรูปสามารถแบ่งแอตทริบิวต์ได้ออกตามหน้าที่การใช้ได้ 2 แบบหลัก คือ

1. แอตทริบิวต์ทั่วไป (Attribute) เป็นแอตทริบิวต์ปกติที่จะใช้ในการสร้างโมเดลหรือเรียกว่าเป็นฟีเจอร์ (feature) หรือตัวแปรต้น (independent variable)

2. ลาเบล (label) แบบนี้จะเป็นแอตทริบิวต์ชนิดพิเศษที่มักจะใช้แสดงคำตอบของสิ่งที่เราต้องการจะสร้างโมเดลมาทำนายหรือเรียกว่า คลาส (Class) หรือตัวแปรตาม (Dependent Variable)

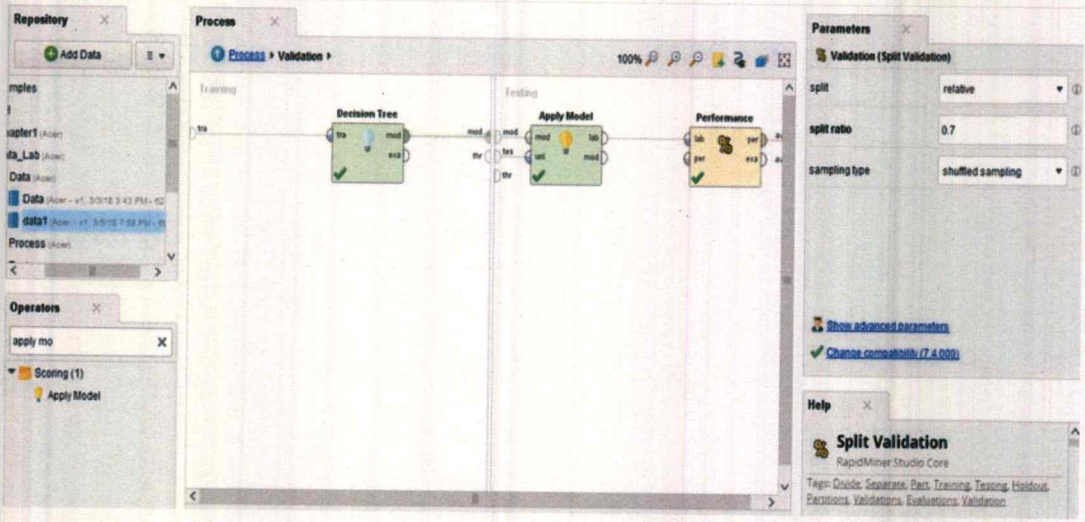
จากรูปจะเห็นได้ว่า แอตทริบิวต์ที่เป็น Label นั้นจะแสดงด้วยคอลัมน์สีเขียวและแอตทริบิวต์ทั่วไปแสดงด้วยคอลัมน์ที่เป็นสีเทา

3.4.1 ขั้นตอนการทำงานของ Decision Tree



รูปที่ 3.2 การจำแนกข้อมูล วิธี Decision Tree

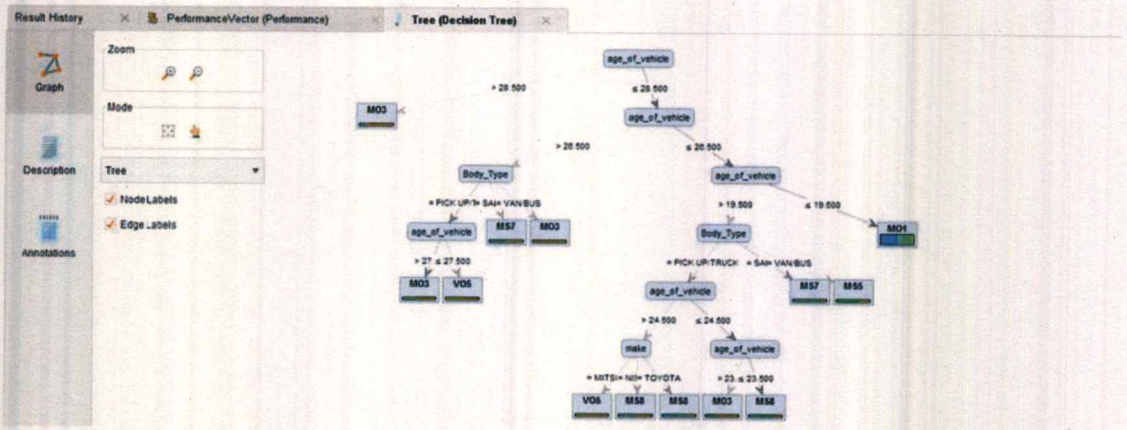
ในรูปที่ 3.2 นี้จะอธิบายการใช้ Operator : Split Validation ใช้สำหรับแบ่งข้อมูลเพื่อสร้างโมเดลและใช้ในการทดสอบประสิทธิภาพของโมเดล ในงานวิจัยนี้จะแบ่งข้อมูลออกเป็น 80% สร้างโมเดล และ 20% ทดสอบโมเดล



รูปที่ 3.3 แสดงการทำงานของ Decision Tree

ในรูปที่ 3.3 นี้จะอธิบายการใช้ Operator : Apply Model ใช้สำหรับทำนายค่าข้อมูลใหม่ Operator : Performance (Classification) ใช้สำหรับคำนวณตัววัดประสิทธิภาพของโมเดล

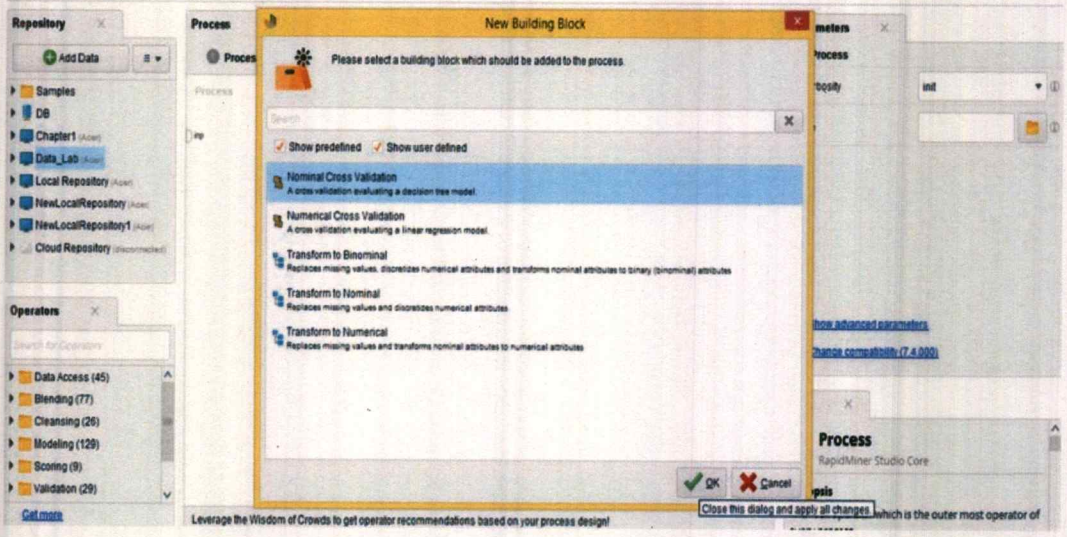
ผลลัพธ์ของ Decision Tree



รูปที่ 3.4 หน้าแสดงผลลัพธ์ของ Decision Tree

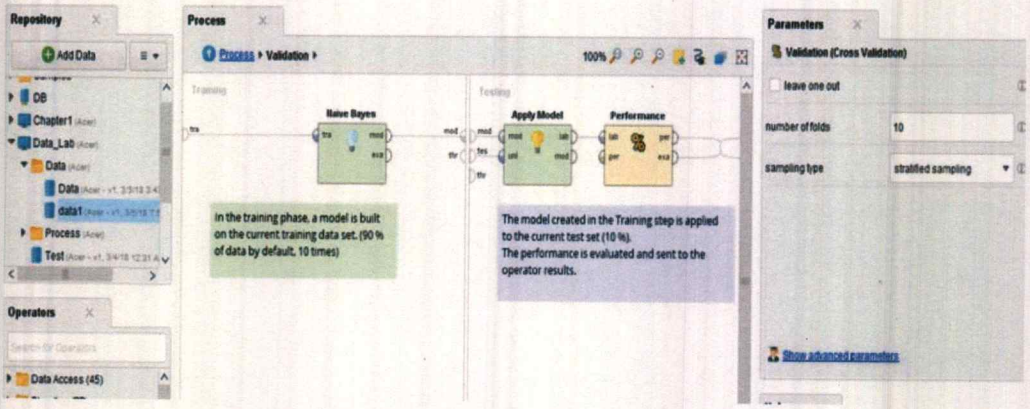
จากผลการทดสอบของเทคนิค Decision Tree มีค่าความแม่นยำ 77.05%

3.4.2 ขั้นตอนการทำงานของ Naïve Bayesian



รูปที่ 3.5 การเลือกฟังก์ชันเพื่อใช้สำหรับการสร้างโมเดลของเทคนิค Naïve Bayesian

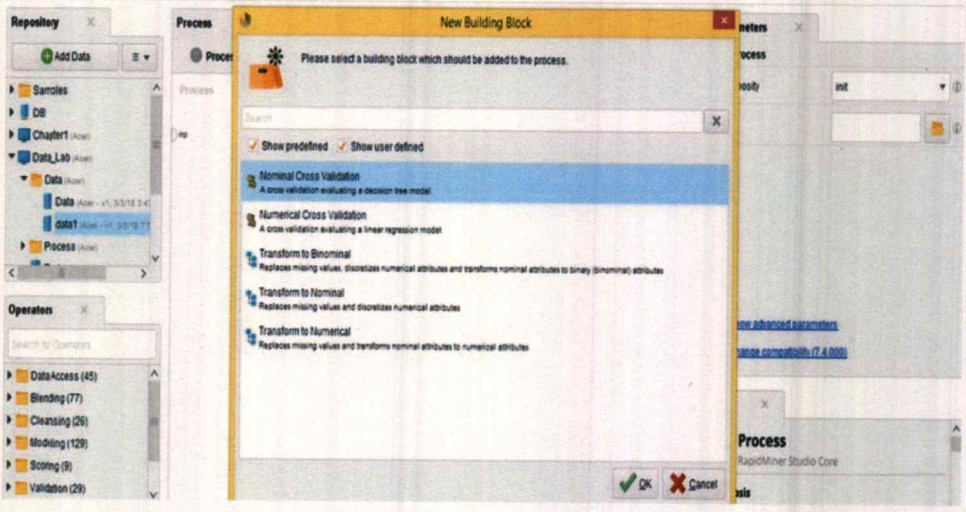
ขั้นตอนมีดังนี้เลือก Edit >> Insert Building Block และหน้าจอแสดงผลจะได้ดังรูปที่ 3.5 และเลือก Nominal Cross Validation ใช้สำหรับแบ่งข้อมูลเพื่อสร้างโมเดล และใช้ในการทดสอบประสิทธิภาพของโมเดล



รูปที่ 3.6 แสดงการทำงานของ Naïve Bayesian

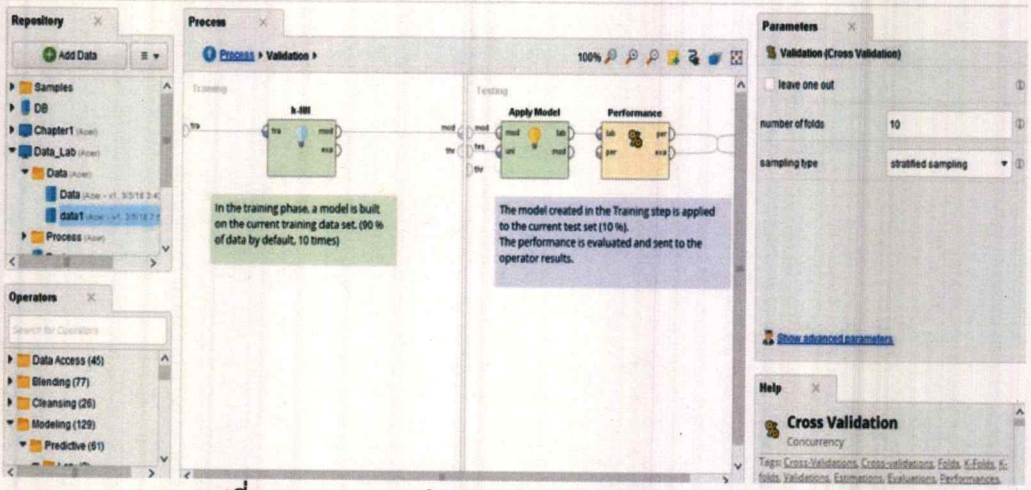
สำหรับการทดสอบประสิทธิภาพของโมเดลจะเลือกแบบ 12-fold ดังนั้นจากผลการทดสอบของเทคนิค Naïve Bayesian Classification มีค่าความแม่นยำ accuracy: 70.38% +/- 2.44% (mikro: 70.38%) (ดูผลลัพธ์ทั้งหมดต่อได้ที่ บทที่ 4)

3.4.3 ขั้นตอนการทำงานของ K-Nearest Neighbors (K-NN)



รูปที่ 3.7 การเลือกฟังก์ชันเพื่อใช้สำหรับการสร้างโมเดลของเทคนิค K-Nearest Neighbors

ขั้นตอนมีดังนี้เลือก Edit >> Insert Building Block และหน้าจอแสดงผลจะได้ดังรูปที่ 3.7 และเลือก Nominal Cross Validation ใช้สำหรับแบ่งข้อมูลเพื่อสร้างโมเดล และใช้ในการทดสอบประสิทธิภาพของโมเดล



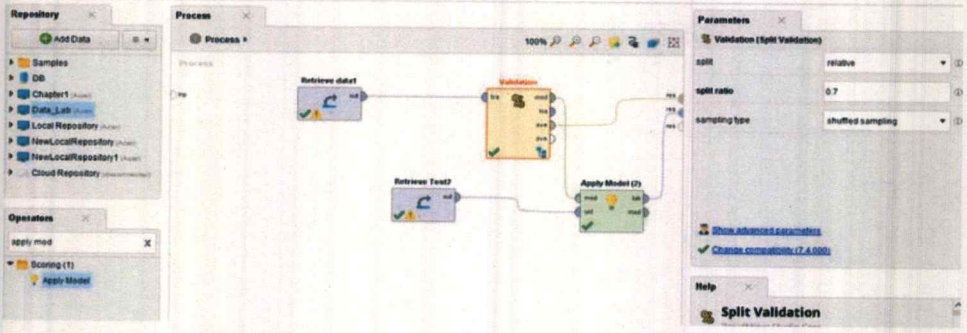
รูปที่ 3.8 แสดงการทำงานของ K – Nearest Neighbors

สำหรับการทดสอบประสิทธิภาพของโมเดลจะเลือกแบบ 12-fold ดังนั้นจากผลการทดสอบของเทคนิค K-Nearest Neighbors มีค่าความแม่นยำ accuracy: 69.36% +/- 3.04% (mikro: 69.35%) (ดูผลลัพธ์ทั้งหมดต่อได้ที่ บทที่ 4)

3.5 การทดสอบโมเดล

เป็นการนำโมเดลที่สร้างมาทำนายค่าของข้อมูลใหม่จากวิธีต่างๆ โดยข้อมูลใหม่ที่จะมาทำนายนั้นมีจำนวน 200 ข้อมูล เป็นการทำนายว่า make, model, Body_Type, age, Capacity เหมาะสมกับรูปแบบใด

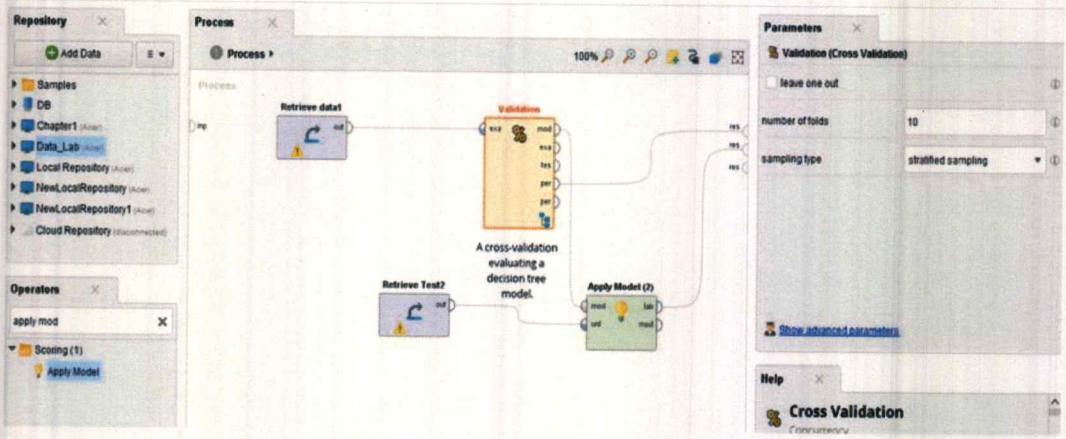
3.5.1 วิธี Decision Tree



รูปที่ 3.9 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี Decision Tree

นำไฟล์ Test มา Save เป็น .CVS ในไฟล์นี้จะมีข้อมูลจำนวน 200 ข้อมูล และจะใช้ Operator : Apply Model เพื่อทำนายข้อมูลใหม่ ซึ่งสามารถดูผลลัพธ์ได้ที่บท 4

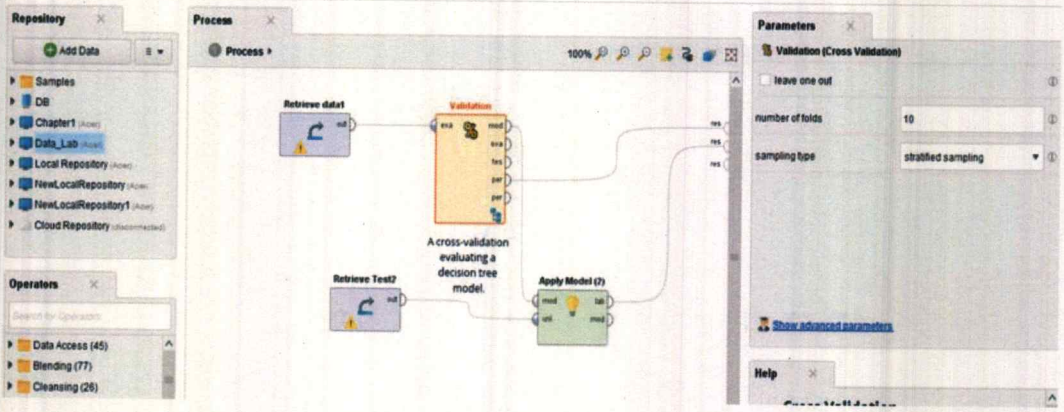
3.5.2 วิธี Naïve Bayesian



รูปที่ 3.10 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี Naïve Bayes

นำไฟล์ Test มา Save เป็น .CVS ในไฟล์นี้จะมีข้อมูลจำนวน 200 ข้อมูล และจะใช้ Operator : Apply Model เพื่อทำนายข้อมูลใหม่ ซึ่งสามารถดูผลลัพธ์ได้ที่บท 4

3.5.3 วิธี K-Nearest Neighbors



รูปที่ 3.11 ขั้นตอนการนำข้อมูลใหม่มาทดสอบวิธี K-Nearest Neighbors

นำไฟล์ Test มา Save เป็น .CVS ในไฟล์นี้จะมีข้อมูลจำนวน 200 ข้อมูล และจะใช้ Operator : Apply Model เพื่อทำนายข้อมูลใหม่ ซึ่งสามารถดูผลลัพธ์ได้ที่บท 4

3.6 ผลการวิเคราะห์และผลการทำนายข้อมูลใหม่

การจำแนกกลุ่มมี 3 วิธี 1.Decision Tree 2.Naïve Bayesian 3.K–Neareat Neighbors จะได้ว่าแต่ละแบบให้ผลลัพธ์ที่มีค่าความแม่นยำดังนี้ 77.05%, 70.38%, 69.36% ตามลำดับ และสามารถดูผลลัพธ์ต่อได้ในบทที่ 4

บทที่ 4

ผลการวิจัยและการอภิปรายผล

ในบทที่ 4 จะเป็นการกล่าวถึงผลลัพธ์ที่ได้ทำการสร้างโมเดลและการทำนายข้อมูลใหม่ซึ่งแบ่งผลลัพธ์เป็น 2 ส่วน คือ 1.ข้อมูลที่เป็น Training Data และส่วนที่ 2.ข้อมูลจำนวนที่เป็นข้อมูลทดสอบ (Testing Data) ซึ่ง Training Data นั้นจะบอกค่าความแม่นยำนั้นก็คือ Class Precision และ Class Recall

4.1. ผลลัพธ์ของ Decision Tree

จากตารางที่ 4.1

- P.MO1 และ True MO1 คือ พบ 215 ข้อมูลใน MO1 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS7 และ True MS7 คือ พบ 109 ข้อมูลใน MS7 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS5 และ True MS5 คือ พบ 2 ข้อมูลใน MS5 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS7 และ True MS8 คือ พบ 49 ข้อมูลที่โมเดลทำนายว่า MS7 แต่ข้อมูลจริง MS8
- P.MS7 และ True VO5 คือ พบ 15 ข้อมูลที่โมเดลทำนายว่า MS7 แต่ข้อมูลจริง VO5

ตารางที่ 4.1 ผลลัพธ์ของ Decision Tree

	true MO1	true MP3	true MS8	true MS7	true MS5	true MS9	true MS6	true VO5	true MO3	true VO2	true MO2	class precision
pred. MO1	215	6	0	0	0	0	0	0	0	0	0	97.29%
pred. MP3	2	2	0	0	0	0	0	0	0	0	0	50.00%
pred. MS8	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. MS7	0	0	49	109	4	11	4	15	6	0	0	55.05%
pred. MS5	0	0	0	0	2	0	0	0	0	0	0	100.00%
pred. MS9	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. MS6	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. VO5	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. MO3	0	0	0	0	0	0	0	1	0	0	0	0.00%
pred. VO2	0	0	0	0	0	0	0	0	0	1	0	100.00%
pred. MO2	0	0	0	0	0	0	0	0	0	0	0	0.00%
class recall	99.08%	25.00%	0.00%	100.00%	33.33%	0.00%	0.00%	0.00%	0.00%	100.00%	0.00%	

ค่า Accuracy = จำนวนข้อมูลที่ทำนายถูกของทุกคลาส/จำนวนข้อมูลทั้งหมด

= (215+2+109+2+1) / 427 = 0.7704918*100 = 77.05%

Class Recall คือ อัตราส่วนของข้อมูลที่โมเดลทำนายถูก

เช่น Class Recall ของ MO1 = จำนวนข้อมูลที่ทำนายถูกของคลาส MO1 / จำนวนข้อมูลทั้งหมดของคลาส MO1

= 215/217 = 0.9908*100 = 99.08%

Class Recall ของ MS7 = 109/109 = 1*100 = 100.00%

Class Recall ของ MS5 = 2/6 = 0.33333*100 = 33.33%

Class Precision คือ อัตราส่วนที่ทำนายถูกจากข้อมูลที่ทำนายว่าเป็นคลาสที่พิจารณา

เช่น Class Precision ของ MO1 = จำนวนข้อมูลที่ทำนายถูกของคลาส MO1/ จำนวนข้อมูลทั้งหมดที่ทำนายของคลาส MO1

= 215/221 = 0.9728506*100 = 97.29%

Class Precision ของ MS7 = 109/198 = 0.55050505*100 = 55.05%

4.2 ผลลัพธ์ของ Naïve Bayes

จากตารางที่ 4.2

- P.MO1 และ True MO1 คือ พบ 1,043 ข้อมูลใน MO1 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS7 และ True MS7 คือ พบ 313 ข้อมูลใน MS7 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MO3 และ True MS7 คือ พบ 11 ข้อมูลที่โมเดลทำนายว่า MO3 แต่ข้อมูลจริง MS7
- P.VO5 และ True MO3 คือ พบ 4 ข้อมูลที่โมเดลทำนายว่า VO5 แต่ข้อมูลจริง MO3

ตารางที่ 4.2 ผลลัพธ์ของ Naïve Bayes

	true MO1	true MP3	true MS8	true MS7	true MS5	true MS9	true MS6	true VO5	true MO3	true VO2	true MO2	class precision
pred. MO1	1043	33	0	2	0	0	0	0	0	0	0	96.75%
pred. MP3	40	18	0	0	0	0	0	0	0	0	0	31.03%
pred. MS8	0	0	83	127	1	12	2	11	3	0	0	34.73%
pred. MS7	0	0	176	313	2	16	11	40	10	1	0	55.01%
pred. MS5	0	0	1	2	21	3	0	2	1	0	0	70.00%
pred. MS9	2	0	7	14	0	5	0	2	2	1	0	15.15%
pred. MS6	0	0	5	9	0	0	1	1	1	0	0	5.88%
pred. VO5	0	0	11	30	0	2	1	1	4	0	1	2.00%
pred. MO3	2	0	12	11	1	4	2	9	16	0	0	28.07%
pred. VO2	1	0	0	1	0	0	0	0	0	1	0	33.33%
pred. MO2	0	0	0	0	0	0	0	0	0	0	0	0.00%
class recall	95.86%	35.29%	28.14%	61.49%	84.00%	11.90%	5.88%	1.52%	43.24%	33.33%	0.00%	

ค่า Accuracy ของ Naïve Bayes = 1,502/2,134*100 = 70.38%

Class Recall คือ อัตราส่วนของข้อมูลที่โมเดลทำนายถูก

เช่น Class Recall ของ MO1 = 1,043/1,088 = 0.9586397*100 = 95.86%

Class Recall ของ MS7 = 313/509 = 0.614931*100 = 61.49%

Class Recall ของ MS8 = 83/295 = 0.2813559*100 = 28.14%

Class Precision คือ อัตราส่วนที่ทำนายถูกจากข้อมูลที่ทำนายว่าเป็นคลาสที่พิจารณา

เช่น Class Precision ของ MO1 = 1,043/1,078 = 0.9675324*100 = 96.75%

Class Precision ของ MS7 = 313/569 = 0.5500878*100 = 55.01%

Class Precision ของ MS8 = 83/239 = 0.347280*100 = 34.73%

4.3 ผลลัพธ์ของ K-Nearest Neighbors (K-NN)

จากตารางที่ 4.3

- P.MO1 และ True MO1 คือ พบ 1,076 ข้อมูลใน MO1 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS7 และ True MS7 คือ พบ 322 ข้อมูลใน MS7 ที่โมเดลทำนายตรงกับข้อมูลจริง
- P.MS7 และ True MS8 คือ พบ 193 ข้อมูลที่โมเดลทำนายว่า MS7 แต่ข้อมูลจริง MS8
- P.MS7 และ True MO3 คือ พบ 11 ข้อมูลที่โมเดลทำนายว่า MS7 แต่ข้อมูลจริง MO3

ตารางที่ 4.3 ผลลัพธ์ของ K-NN

	true MO1	true MP3	true MS8	true MS7	true MS5	true MS9	true MS6	true VO5	true MO3	true VO2	true MO2	class precision
pred. MO1	1076	49	33	66	11	2	1	8	2	0	0	86.22%
pred. MP3	1	2	0	0	0	0	0	0	0	0	0	66.67%
pred. MS8	2	0	65	117	4	11	0	16	8	0	0	29.15%
pred. MS7	9	0	193	322	8	29	16	37	11	3	1	51.19%
pred. MS5	0	0	0	0	1	0	0	0	0	0	0	100.00%
pred. MS9	0	0	0	0	1	0	0	0	0	0	0	0.00%
pred. MS6	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. VO5	0	0	1	0	0	0	0	0	2	0	0	0.00%
pred. MO3	0	0	3	4	0	0	0	5	14	0	0	53.85%
pred. VO2	0	0	0	0	0	0	0	0	0	0	0	0.00%
pred. MO2	0	0	0	0	0	0	0	0	0	0	0	0.00%
class recall	98.90%	3.92%	22.03%	63.26%	4.00%	0.00%	0.00%	0.00%	37.84%	0.00%	0.00%	

ค่า Accuracy = 1,480 / 2,134 = 0.6935332*100 = 69.36%

Class Recall คือ อัตราส่วนของข้อมูลที่โมเดลทำนายถูก

เช่น Class Recall ของ MO1 = 1,076/1,088 = 0.9889705*100 = 98.90%

Class Recall ของ MS7 = 322/509 = 0.6326129*100 = 63.26%

Class Recall ของ MS8 = 65/295 = 0.220338*100 = 22.03%

Class Precision คือ อัตราส่วนที่ทำนายถูกจากข้อมูลที่ทำนายว่าเป็นคลาสที่พิจารณา

เช่น Class Precision ของ MO1 = 1,076/1,248= 0.862179*100 = 86.22%

Class Precision ของ MS7 = 322/629 = 0.511923*100 = 51.19%

Class Precision ของ MS8 = 65/223 = 0.2914798*100 = 29.15%

4.4 การทดสอบข้อมูลใหม่

การทดสอบข้อมูลใหม่นั้นจะมีข้อมูลใหม่มาทดสอบทั้งหมด 200 ข้อมูล จะทำการทดสอบทั้ง 3 วิธี ซึ่งขั้นตอนในการทดสอบได้กล่าวในบทที่ 3

ตารางที่ 4.4 ผลลัพธ์ของการทดสอบข้อมูลใหม่วิธี Decision Tree

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	Cover
1	MO1	FORTUNER	TOYOTA	SALOON	3	2,694.00	CO
2	MO1	FORTUNER	TOYOTA	SALOON	3	2,393.00	TP
3	MO1	HILUX	TOYOTA	SALOON	6	2,982.00	TP
4	MO1	HILUX	TOYOTA	SALOON	15	2,982.00	TP
5	MO1	RANGER	FORD	SALOON	4	2,200.00	CO
6	MO1	CRV	HONDA	SALOON	16	2,000.00	TP
7	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
8	MO1	HILUX	TOYOTA	PICK UP/TRUCK	7	2,982.00	CO
9	MO1	CX-5	MAZDA	SALOON	3	2,200.00	CO
10	MS7	E CLASS	BENZ	SALOON	21	2,295.00	TP
11	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
12	MO1	COMMUTER	TOYOTA	VAN/BUS	7	2,982.00	CO
13	MS5	HIACE	TOYOTA	VAN/BUS	22	2,438.00	TF
14	MO1	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	TF
15	MS8	HILUX	TOYOTA	PICK UP/TRUCK	20	2,982.00	TP
16	VO5	CKD	MITSUBISHI	PICK UP/TRUCK	25	3,494.00	TP
17	MS7	COROLLA	TOYOTA	SALOON	24	1,600.00	TP
18	MO1	FRONTIER	NISSAN	PICK UP/TRUCK	14	2,488.00	TP
19	MO1	D-MAX	ISUZU	PICK UP/TRUCK	14	2,999.00	CO
20	MO1	CKD	NISSAN	PICK UP/TRUCK	18	2,499.00	CO
21	MO1	D-MAX	ISUZU	PICK UP/TRUCK	11	2,999.00	TP
22	MO1	COROLLA	TOYOTA	SALOON	10	1,598.00	TP
23	MS7	COROLLA	TOYOTA	SALOON	20	1,468.00	TF
24	MO1	HILUX	TOYOTA	PICK UP/TRUCK	14	2,982.00	CO
25	MO1	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TP
26	MO1	HILUX	TOYOTA	SALOON	8	2,694.00	TP
27	MO1	D-MAX	ISUZU	SALOON	14	2,999.00	TP
28	MO1	CIVIC	HONDA	SALOON	6	2,000.00	TP
29	VO5	CKD	MITSUBISHI	PICK UP/TRUCK	26	3,494.00	TP
30	MO1	CKD	MITSUBISHI	PICK UP/TRUCK	15	3,494.00	TF
31	MO1	SOLUNA	TOYOTA	SALOON	16	1,500.00	TF
32	MO1	ACCORD	HONDA	SALOON	9	1,997.00	CO
33	MO1	CKD	NISSAN	SALOON	7	2,496.00	TF
34	MO1	JAZZ	HONDA	SALOON	14	1,497.00	TF
35	MO1	HILUX	TOYOTA	VAN/BUS	6	2,982.00	CO
36	MO1	D-MAX	ISUZU	VAN/BUS	7	2,999.00	CO
37	MO1	CAB-4	ISUZU	SALOON	11	2,500.00	TF
38	MO1	SOLUNA	TOYOTA	SALOON	15	1,500.00	CO

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	Cover
39	MO1	TEANA	NISSAN	SALOON	5	1,997.00	CO
40	MO1	CIVIC	HONDA	SALOON	12	1,800.00	TF
41	MS8	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TF
42	MO1	S40	VOLVO	SALOON	19	1,948.00	TP
43	MO1	CITY	HONDA	SALOON	6	1,497.00	TF
44	MO1	MU-X	ISUZU	SALOON	4	3,000.00	CO
45	MO1	JAZZ	HONDA	SALOON	2	1,497.00	TP
46	MO1	D-MAX	ISUZU	SALOON	6	2,499.00	CO
47	MO1	SOLUNA	TOYOTA	SALOON	15	1,496.00	TP
48	MO1	HILUX	TOYOTA	VAN/BUS	5	2,982.00	CO
49	MO1	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TP
50	MO1	TRITON	MITSUBISHI	PICK UP/TRUCK	8	2,477.00	CO
51	MO1	CKD	TOYOTA	PICK UP/TRUCK	13	2,499.00	TP
52	MO1	SOLUNA	TOYOTA	SALOON	5	1,500.00	CO
53	MO1	YARIS	TOYOTA	SALOON	4	1,200.00	CO
54	MO1	ALMERA	NISSAN	SALOON	4	1,200.00	CO
55	MS7	SOLUNA	TOYOTA	SALOON	21	2,000.00	TP
56	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
57	MO1	CIVIC	HONDA	SALOON	6	1,800.00	TF
58	MO1	XENON	TATA	PICK UP/TRUCK	8	2,850.00	TP
59	MO1	COOPER	MINI	SALOON	8	1,598.00	TF
60	MO1	JAZZ	HONDA	SALOON	7	1,500.00	TF
61	MO1	SOLUNA	TOYOTA	SALOON	6	1,497.00	TF
62	MO1	NAVARA	NISSAN	PICK UP/TRUCK	4	2,488.00	CO
63	MS7	E CLASS	BENZ	SALOON	22	2,295.00	TP
64	MS8	D-MAX	ISUZU	PICK UP/TRUCK	25	2,999.00	TP
65	MO1	HIACE	TOYOTA	VAN/BUS	11	2,494.00	TF
66	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TP
67	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	3,000.00	CO
68	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
69	MO1	COROLLA	TOYOTA	SALOON	6	1,798.00	CO
70	MO1	C CLASS	BENZ	SALOON	6	1,796.00	CO
71	MO1	D-MAX	ISUZU	PICK UP/TRUCK	7	2,999.00	TF
72	MO1	CIVIC	HONDA	SALOON	5	1,798.00	TF
73	MO1	RANGER	FORD	SALOON	6	2,200.00	TF
74	MO1	HILUX	TOYOTA	PICK UP/TRUCK	6	2,982.00	CO
75	MO1	2	MAZDA	SALOON	3	1,299.00	CO
76	MO1	COROLLA	TOYOTA	SALOON	15	1,598.00	TF
77	MO1	3	MAZDA	SALOON	9	1,600.00	CO
78	MO1	TRITON	MITSUBISHI	SALOON	4	2,500.00	TP
79	MO1	HILUX	TOYOTA	SALOON	11	2,494.00	CO
80	MO1	FIGHTER	MAZDA	PICK UP/TRUCK	12	2,982.00	TP
81	MO1	JAZZ	HONDA	SALOON	7	1,497.00	CO

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	Cover
82	MO1	D-MAX	ISUZU	SALOON	3	2,499.00	CO
83	MO1	CITY	HONDA	SALOON	6	1,497.00	TP
84	MO1	HILUX	TOYOTA	PICK UP/TRUCK	10	2,982.00	TP
85	MO1	SPACE WAGON	MITSUBISHI	SALOON	13	2,378.00	CO
86	MO1	SWIFT ECO	SUZUKI	SALOON	4	1,242.00	CO
87	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
88	MO1	SPACE WAGON	MITSUBISHI	SALOON	11	2,378.00	TF
89	MO1	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
90	MO1	YARIS	TOYOTA	SALOON	3	1,197.00	CO
91	MO1	COROLLA	TOYOTA	SALOON	3	1,798.00	CO
92	MO1	SOLUNA	TOYOTA	SALOON	3	1,497.00	CO
93	MO1	C CLASS	BENZ	SALOON	5	1,800.00	TP
94	MO1	JAZZ	HONDA	SALOON	5	2,000.00	TP
95	MO1	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
96	MO1	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
97	MO1	AVEO	CHEVROLET	SALOON	11	1,399.00	TP
98	MO1	HILUX	TOYOTA	SALOON	8	2,494.00	TF
99	MS7	CORONA	TOYOTA	SALOON	24	1,600.00	TP
100	MO1	INNOVA	TOYOTA	SALOON	9	1,998.00	TF
101	MO1	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
102	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TP
103	MO1	FORTUNER	TOYOTA	SALOON	12	2,982.00	TF
104	MO1	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TF
105	MO1	CIVIC	HONDA	SALOON	3	1,798.00	CO
106	MO1	YARIS	TOYOTA	SALOON	3	1,197.00	CO
107	MO1	BT-50	MAZDA	SALOON	5	2,198.00	TF
108	MO1	PAJERO SPORT	MITSUBISHI	SALOON	10	2,500.00	TP
109	MO1	CIVIC	HONDA	SALOON	5	1,800.00	CO
110	MO1	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
111	MO1	CKD	TOYOTA	SALOON	12	2,494.00	TP
112	MO1	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
113	MS7	CKD	NISSAN	SALOON	28	1,998.00	TP
114	MO1	CITY	HONDA	SALOON	7	1,497.00	TF
115	MO1	ACCORD	HONDA	SALOON	10	2,354.00	TF
116	MO1	PAJERO SPORT	MITSUBISHI	SALOON	6	2,477.00	TP
117	MS7	SUNNY	NISSAN	SALOON	22	1,498.00	TP
118	MO1	TEANA	NISSAN	SALOON	14	2,400.00	TF
119	MO1	SOLUNA	TOYOTA	SALOON	18	1,498.00	TP
120	MO1	SOLUNA	TOYOTA	SALOON	6	1,500.00	TP
121	MO1	CKD	TOYOTA	VAN/BUS	5	1,497.00	TP
122	MO1	SOLUNA	TOYOTA	SALOON	13	1,500.00	CO
123	MO1	HILUX	TOYOTA	SALOON	13	2,982.00	TP
124	MO1	HILUX	TOYOTA	SALOON	9	2,982.00	TF

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	Cover
125	MS8	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TP
126	MO1	FORTUNER	TOYOTA	SALOON	9	2,982.00	TF
127	MO1	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
128	MO1	CITY	HONDA	SALOON	3	1,497.00	TF
129	MO1	FIESTA	FORD	SALOON	5	1,499.00	TP
130	MO1	FORTUNER	TOYOTA	SALOON	3	2,982.00	TF
131	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
132	MO1	JAZZ	HONDA	SALOON	5	1,497.00	TP
133	MO1	SOLUNA	TOYOTA	SALOON	4	1,497.00	TP
134	MO1	TRITON	MITSUBISHI	PICK UP/TRUCK	4	2,477.00	TP
135	MO1	HILUX	TOYOTA	SALOON	8	2,494.00	TP
136	MO1	SPORT RIDER	TOYOTA	SALOON	18	2,000.00	TP
137	MO1	SOLUNA	TOYOTA	SALOON	15	1,500.00	TF
138	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TF
139	MO1	C CLASS	BENZ	SALOON	6	1,800.00	TF
140	MO1	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
141	MO1	COROLLA	TOYOTA	SALOON	15	1,598.00	TP
142	MO1	HILUX	TOYOTA	PICK UP/TRUCK	8	2,494.00	TP
143	MS8	BIG M	NISSAN	PICK UP/TRUCK	23	2,663.00	TP
144	MO1	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TF
145	MO1	CAMRY	TOYOTA	SALOON	15	2,362.00	TP
146	MO1	D-MAX	ISUZU	PICK UP/TRUCK	6	2,999.00	CO
147	MO1	CITY	HONDA	SALOON	9	1,500.00	TF
148	MS7	ACCORD	HONDA	SALOON	27	2,000.00	TP
149	MO1	LANCER	MITSUBISHI	SALOON	7	2,000.00	CO
150	MO1	D-MAX	ISUZU	PICK UP/TRUCK	15	2,999.00	TF
151	MO1	JAZZ	HONDA	SALOON	5	1,497.00	CO
152	MO1	MG3	MG	SALOON	3	1,498.00	CO
153	MO1	ACCORD	HONDA	SALOON	10	2,354.00	TP
154	MO1	D-MAX	ISUZU	SALOON	4	2,500.00	TP
155	MO1	H-1	HYUNDAI	VAN/BUS	6	2,497.00	CO
156	MO1	COROLLA	TOYOTA	SALOON	5	1,598.00	CO
157	MO1	CIVIC	HONDA	SALOON	11	1,799.00	CO
158	MO1	FORTUNER	TOYOTA	SALOON	11	2,982.00	CO
159	MO1	CIVIC	HONDA	SALOON	15	1,688.00	TP
160	MO1	D-MAX	ISUZU	PICK UP/TRUCK	2	2,999.00	CO
161	MO1	SOLUNA	TOYOTA	SALOON	11	1,500.00	CO
162	MS7	S40	VOLVO	SALOON	20	1,948.00	TP
163	MO1	NAVARA	NISSAN	SALOON	2	2,488.00	CO
164	MS8	CKD	MITSUBISHI	PICK UP/TRUCK	21	3,494.00	TP
165	MS7	C CLASS	BENZ	SALOON	23	2,159.00	TP
166	MO1	LANCER	MITSUBISHI	SALOON	10	1,600.00	TF
167	MO1	TRITON	MITSUBISHI	SALOON	7	2,351.00	TF

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	Cover
168	MO1	D-MAX	ISUZU	SALOON	2	1,850.00	CO
169	MO1	D-MAX	ISUZU	VAN/BUS	7	2,999.00	TF
170	MO3	CKD	MITSUBISHI	SALOON	29	1,298.00	TP
171	MO1	FORTUNER	TOYOTA	SALOON	2	2,400.00	CO
172	MO1	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
173	MO1	SOLUNA	TOYOTA	SALOON	3	1,497.00	TP
174	MO1	CITY	HONDA	SALOON	8	1,497.00	TP
175	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
176	MO1	TEANA	NISSAN	SALOON	13	2,300.00	TP
177	MO1	SUNNY	NISSAN	SALOON	14	1,769.00	TP
178	MO1	2	MAZDA	SALOON	5	1,498.00	TP
179	MO1	HILUX	TOYOTA	VAN/BUS	3	2,494.00	TF
180	MO1	HILUX	TOYOTA	PICK UP/TRUCK	13	2,494.00	TF
181	MO1	SOLUNA	TOYOTA	SALOON	6	1,497.00	TP
182	MS7	TROOPER	ISUZU	SALOON	21	3,168.00	TP
183	MO1	HILUX	TOYOTA	SALOON	7	3,000.00	CO
184	MO1	CITY	HONDA	SALOON	5	1,500.00	CO
185	MO1	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
186	MO1	SOLUNA	TOYOTA	SALOON	12	1,497.00	TP
187	MO1	CAMRY	TOYOTA	SALOON	3	1,500.00	CO
188	MO1	COROLLA	TOYOTA	SALOON	17	1,600.00	TF
189	MO1	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	CO
190	MO1	SOLUNA	TOYOTA	SALOON	4	1,497.00	CO
191	MO1	CAPTIVA	CHEVROLET	SALOON	5	2,384.00	CO
192	MO1	CRV	HONDA	SALOON	15	2,000.00	TF
193	MO1	HILUX	TOYOTA	PICK UP/TRUCK	10	2,494.00	TP
194	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
195	MO3	CKD	MITSUBISHI	PICK UP/TRUCK	24	3,494.00	TP
196	MO1	D-MAX	ISUZU	PICK UP/TRUCK	18	2,999.00	TP
197	MO1	D-MAX	ISUZU	SALOON	4	2,000.00	TP
198	MO1	HILUX	TOYOTA	VAN/BUS	17	2,494.00	TP
199	MO1	CKD	NISSAN	PICK UP/TRUCK	8	1,497.00	TP
200	MO1	NAVARA	NISSAN	PICK UP/TRUCK	8	2,488.00	TP

จากการนำข้อมูลใหม่ไปทดสอบนั้นสามารถสรุปผลลัพธ์ได้ดังนี้ MO1, MO3, MS5, MS7, MS8, VO5 มีจำนวน 175, 2, 1, 14, 6, 2 จำนวน ตามลำดับ ซึ่งจะเห็นได้ว่าวิธี Decision Tree นั้นได้ให้ผลลัพธ์ MO1 มากที่สุด

ตารางที่ 4.5 ผลลัพธ์ของการทดสอบข้อมูลใหม่วิธี Naïve Bayes

Row	Prediction	model	Make	Body_Type	age_of_vehicle	Capacity	Cover
1	MO1	FORTUNER	TOYOTA	SALOON	3	2,694.00	CO
2	MS7	FORTUNER	TOYOTA	SALOON	3	2,393.00	TP
3	MS7	HILUX	TOYOTA	SALOON	6	2,982.00	TP
4	MS7	HILUX	TOYOTA	SALOON	15	2,982.00	TP
5	MO1	RANGER	FORD	SALOON	4	2,200.00	CO
6	MS7	CRV	HONDA	SALOON	16	2,000.00	TP
7	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
8	MO1	HILUX	TOYOTA	PICK UP/TRUCK	7	2,982.00	CO
9	MO1	CX-5	MAZDA	SALOON	3	2,200.00	CO
10	MS7	E CLASS	BENZ	SALOON	21	2,295.00	TP
11	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
12	MO1	COMMUTER	TOYOTA	VAN/BUS	7	2,982.00	CO
13	MS5	HIACE	TOYOTA	VAN/BUS	22	2,438.00	TF
14	MS5	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	TF
15	MS7	HILUX	TOYOTA	PICK UP/TRUCK	20	2,982.00	TP
16	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	25	3,494.00	TP
17	MS8	COROLLA	TOYOTA	SALOON	24	1,600.00	TP
18	MS7	FRONTIER	NISSAN	PICK UP/TRUCK	14	2,488.00	TP
19	MO1	D-MAX	ISUZU	PICK UP/TRUCK	14	2,999.00	CO
20	MO1	CKD	NISSAN	PICK UP/TRUCK	18	2,499.00	CO
21	MS8	D-MAX	ISUZU	PICK UP/TRUCK	11	2,999.00	TP
22	MS7	COROLLA	TOYOTA	SALOON	10	1,598.00	TP
23	MO2	COROLLA	TOYOTA	SALOON	20	1,468.00	TF
24	MO1	HILUX	TOYOTA	PICK UP/TRUCK	14	2,982.00	CO
25	MS8	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TP
26	MS7	HILUX	TOYOTA	SALOON	8	2,694.00	TP
27	MS7	D-MAX	ISUZU	SALOON	14	2,999.00	TP
28	MS7	CIVIC	HONDA	SALOON	6	2,000.00	TP
29	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	26	3,494.00	TP
30	VO2	CKD	MITSUBISHI	PICK UP/TRUCK	15	3,494.00	TF
31	VO5	SOLUNA	TOYOTA	SALOON	16	1,500.00	TF
32	MO1	ACCORD	HONDA	SALOON	9	1,997.00	CO
33	MS8	CKD	NISSAN	SALOON	7	2,496.00	TF
34	VO5	JAZZ	HONDA	SALOON	14	1,497.00	TF
35	MO1	HILUX	TOYOTA	VAN/BUS	6	2,982.00	CO
36	MO1	D-MAX	ISUZU	VAN/BUS	7	2,999.00	CO
37	MO1	CAB-4	ISUZU	SALOON	11	2,500.00	TF
38	MO1	SOLUNA	TOYOTA	SALOON	15	1,500.00	CO
39	MO1	TEANA	NISSAN	SALOON	5	1,997.00	CO
40	VO5	CIVIC	HONDA	SALOON	12	1,800.00	TF
41	MS9	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TF
42	VO5	S40	VOLVO	SALOON	19	1,948.00	TP

Row	Prediction	model	Make	Body_Type	age_of_vehicle	Capacity	Cover
43	MP3	CITY	HONDA	SALOON	6	1,497.00	TF
44	MO1	MU-X	ISUZU	SALOON	4	3,000.00	CO
45	MS7	JAZZ	HONDA	SALOON	2	1,497.00	TP
46	MO1	D-MAX	ISUZU	SALOON	6	2,499.00	CO
47	MS7	SOLUNA	TOYOTA	SALOON	15	1,496.00	TP
48	MO1	HILUX	TOYOTA	VAN/BUS	5	2,982.00	CO
49	MS7	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TP
50	MO1	TRITON	MITSUBISHI	PICK UP/TRUCK	8	2,477.00	CO
51	MO3	CKD	TOYOTA	PICK UP/TRUCK	13	2,499.00	TP
52	MO1	SOLUNA	TOYOTA	SALOON	5	1,500.00	CO
53	MO1	YARIS	TOYOTA	SALOON	4	1,200.00	CO
54	MO1	ALMERA	NISSAN	SALOON	4	1,200.00	CO
55	MS7	SOLUNA	TOYOTA	SALOON	21	2,000.00	TP
56	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
57	MO1	CIVIC	HONDA	SALOON	6	1,800.00	TF
58	MS7	XENON	TATA	PICK UP/TRUCK	8	2,850.00	TP
59	VO5	COOPER	MINI	SALOON	8	1,598.00	TF
60	MO1	JAZZ	HONDA	SALOON	7	1,500.00	TF
61	MP3	SOLUNA	TOYOTA	SALOON	6	1,497.00	TF
62	MO1	NAVARA	NISSAN	PICK UP/TRUCK	4	2,488.00	CO
63	MS7	E CLASS	BENZ	SALOON	22	2,295.00	TP
64	MS9	D-MAX	ISUZU	PICK UP/TRUCK	25	2,999.00	TP
65	MS5	HIACE	TOYOTA	VAN/BUS	11	2,494.00	TF
66	MS7	SOLUNA	TOYOTA	SALOON	8	1,497.00	TP
67	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	3,000.00	CO
68	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
69	MO1	COROLLA	TOYOTA	SALOON	6	1,798.00	CO
70	MO1	C CLASS	BENZ	SALOON	6	1,796.00	CO
71	MS5	D-MAX	ISUZU	PICK UP/TRUCK	7	2,999.00	TF
72	MP3	CIVIC	HONDA	SALOON	5	1,798.00	TF
73	MS7	RANGER	FORD	SALOON	6	2,200.00	TF
74	MO1	HILUX	TOYOTA	PICK UP/TRUCK	6	2,982.00	CO
75	MO1	2	MAZDA	SALOON	3	1,299.00	CO
76	MS8	COROLLA	TOYOTA	SALOON	15	1,598.00	TF
77	MO1	3	MAZDA	SALOON	9	1,600.00	CO
78	MS7	TRITON	MITSUBISHI	SALOON	4	2,500.00	TP
79	MO1	HILUX	TOYOTA	SALOON	11	2,494.00	CO
80	MS8	FIGHTER	MAZDA	PICK UP/TRUCK	12	2,982.00	TP
81	MO1	JAZZ	HONDA	SALOON	7	1,497.00	CO
82	MO1	D-MAX	ISUZU	SALOON	3	2,499.00	CO
83	MS7	CITY	HONDA	SALOON	6	1,497.00	TP
84	MS7	HILUX	TOYOTA	PICK UP/TRUCK	10	2,982.00	TP
85	MO1	SPACE WAGON	MITSUBISHI	SALOON	13	2,378.00	CO

Row	Prediction	model	Make	Body_Type	age_of_vehicle	Capacity	Cover
86	MO1	SWIFT ECO	SUZUKI	SALOON	4	1,242.00	CO
87	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
88	MO1	SPACE WAGON	MITSUBISHI	SALOON	11	2,378.00	TF
89	MS7	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
90	MP3	YARIS	TOYOTA	SALOON	3	1,197.00	CO
91	MO1	COROLLA	TOYOTA	SALOON	3	1,798.00	CO
92	MP3	SOLUNA	TOYOTA	SALOON	3	1,497.00	CO
93	MS7	C CLASS	BENZ	SALOON	5	1,800.00	TP
94	MS7	JAZZ	HONDA	SALOON	5	2,000.00	TP
95	MS7	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
96	MS7	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
97	MS7	AVEO	CHEVROLET	SALOON	11	1,399.00	TP
98	MS6	HILUX	TOYOTA	SALOON	8	2,494.00	TF
99	VO5	CORONA	TOYOTA	SALOON	24	1,600.00	TP
100	MO1	INNOVA	TOYOTA	SALOON	9	1,998.00	TF
101	MS5	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
102	MS8	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TP
103	MS8	FORTUNER	TOYOTA	SALOON	12	2,982.00	TF
104	MS5	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TF
105	MO1	CIVIC	HONDA	SALOON	3	1,798.00	CO
106	MP3	YARIS	TOYOTA	SALOON	3	1,197.00	CO
107	MO1	BT-50	MAZDA	SALOON	5	2,198.00	TF
108	MS7	PAJERO SPORT	MITSUBISHI	SALOON	10	2,500.00	TP
109	MO1	CIVIC	HONDA	SALOON	5	1,800.00	CO
110	MS7	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
111	MS7	CKD	TOYOTA	SALOON	12	2,494.00	TP
112	MS7	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
113	MS7	CKD	NISSAN	SALOON	28	1,998.00	TP
114	MP3	CITY	HONDA	SALOON	7	1,497.00	TF
115	MO1	ACCORD	HONDA	SALOON	10	2,354.00	TF
116	MS7	PAJERO SPORT	MITSUBISHI	SALOON	6	2,477.00	TP
117	MS7	SUNNY	NISSAN	SALOON	22	1,498.00	TP
118	MO1	TEANA	NISSAN	SALOON	14	2,400.00	TF
119	MS7	SOLUNA	TOYOTA	SALOON	18	1,498.00	TP
120	MS7	SOLUNA	TOYOTA	SALOON	6	1,500.00	TP
121	VO5	CKD	TOYOTA	VAN/BUS	5	1,497.00	TP
122	MO1	SOLUNA	TOYOTA	SALOON	13	1,500.00	CO
123	MS7	HILUX	TOYOTA	SALOON	13	2,982.00	TP
124	MO1	HILUX	TOYOTA	SALOON	9	2,982.00	TF
125	MS9	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TP
126	MO1	FORTUNER	TOYOTA	SALOON	9	2,982.00	TF
127	MS7	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
128	MP3	CITY	HONDA	SALOON	3	1,497.00	TF

Row	Prediction	model	Make	Body_Type	age_of_vehicle	Capacity	Cover
129	MS7	FIESTA	FORD	SALOON	5	1,499.00	TP
130	MP3	FORTUNER	TOYOTA	SALOON	3	2,982.00	TF
131	MP3	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
132	MS7	JAZZ	HONDA	SALOON	5	1,497.00	TP
133	MS7	SOLUNA	TOYOTA	SALOON	4	1,497.00	TP
134	MS7	TRITON	MITSUBISHI	PICK UP/TRUCK	4	2,477.00	TP
135	MS7	HILUX	TOYOTA	SALOON	8	2,494.00	TP
136	MS7	SPORT RIDER	TOYOTA	SALOON	18	2,000.00	TP
137	VO5	SOLUNA	TOYOTA	SALOON	15	1,500.00	TF
138	MS5	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TF
139	MS7	C CLASS	BENZ	SALOON	6	1,800.00	TF
140	MS5	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
141	MS7	COROLLA	TOYOTA	SALOON	15	1,598.00	TP
142	MS7	HILUX	TOYOTA	PICK UP/TRUCK	8	2,494.00	TP
143	MO3	BIG M	NISSAN	PICK UP/TRUCK	23	2,663.00	TP
144	MS5	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TF
145	MS7	CAMRY	TOYOTA	SALOON	15	2,362.00	TP
146	MO1	D-MAX	ISUZU	PICK UP/TRUCK	6	2,999.00	CO
147	MO1	CITY	HONDA	SALOON	9	1,500.00	TF
148	MS7	ACCORD	HONDA	SALOON	27	2,000.00	TP
149	MO1	LANCER	MITSUBISHI	SALOON	7	2,000.00	CO
150	MS5	D-MAX	ISUZU	PICK UP/TRUCK	15	2,999.00	TF
151	MP3	JAZZ	HONDA	SALOON	5	1,497.00	CO
152	MO1	MG3	MG	SALOON	3	1,498.00	CO
153	MS7	ACCORD	HONDA	SALOON	10	2,354.00	TP
154	MS7	D-MAX	ISUZU	SALOON	4	2,500.00	TP
155	MO1	H-1	HYUNDAI	VAN/BUS	6	2,497.00	CO
156	MO1	COROLLA	TOYOTA	SALOON	5	1,598.00	CO
157	MO1	CIVIC	HONDA	SALOON	11	1,799.00	CO
158	MO1	FORTUNER	TOYOTA	SALOON	11	2,982.00	CO
159	VO5	CIVIC	HONDA	SALOON	15	1,688.00	TP
160	MO1	D-MAX	ISUZU	PICK UP/TRUCK	2	2,999.00	CO
161	MO1	SOLUNA	TOYOTA	SALOON	11	1,500.00	CO
162	MO2	S40	VOLVO	SALOON	20	1,948.00	TP
163	MP3	NAVARA	NISSAN	SALOON	2	2,488.00	CO
164	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	21	3,494.00	TP
165	MS9	C CLASS	BENZ	SALOON	23	2,159.00	TP
166	MO1	LANCER	MITSUBISHI	SALOON	10	1,600.00	TF
167	MS7	TRITON	MITSUBISHI	SALOON	7	2,351.00	TF
168	MP3	D-MAX	ISUZU	SALOON	2	1,850.00	CO
169	MS5	D-MAX	ISUZU	VAN/BUS	7	2,999.00	TF
170	MO3	CKD	MITSUBISHI	SALOON	29	1,298.00	TP
171	MO1	FORTUNER	TOYOTA	SALOON	2	2,400.00	CO

Row	Prediction	model	Make	Body_Type	age_of_vehicle	Capacity	Cover
172	MS8	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
173	MS7	SOLUNA	TOYOTA	SALOON	3	1,497.00	TP
174	MS8	CITY	HONDA	SALOON	8	1,497.00	TP
175	MP3	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
176	MS7	TEANA	NISSAN	SALOON	13	2,300.00	TP
177	MS6	SUNNY	NISSAN	SALOON	14	1,769.00	TP
178	MS7	2	MAZDA	SALOON	5	1,498.00	TP
179	MS5	HILUX	TOYOTA	VAN/BUS	3	2,494.00	TF
180	MS5	HILUX	TOYOTA	PICK UP/TRUCK	13	2,494.00	TF
181	MS7	SOLUNA	TOYOTA	SALOON	6	1,497.00	TP
182	MS6	TROOPER	ISUZU	SALOON	21	3,168.00	TP
183	MO1	HILUX	TOYOTA	SALOON	7	3,000.00	CO
184	MO1	CITY	HONDA	SALOON	5	1,500.00	CO
185	MS8	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
186	MS7	SOLUNA	TOYOTA	SALOON	12	1,497.00	TP
187	MO1	CAMRY	TOYOTA	SALOON	3	1,500.00	CO
188	VO5	COROLLA	TOYOTA	SALOON	17	1,600.00	TF
189	MO1	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	CO
190	MP3	SOLUNA	TOYOTA	SALOON	4	1,497.00	CO
191	MO1	CAPTIVA	CHEVROLET	SALOON	5	2,384.00	CO
192	VO2	CRV	HONDA	SALOON	15	2,000.00	TF
193	MS7	HILUX	TOYOTA	PICK UP/TRUCK	10	2,494.00	TP
194	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
195	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	24	3,494.00	TP
196	MS8	D-MAX	ISUZU	PICK UP/TRUCK	18	2,999.00	TP
197	MS7	D-MAX	ISUZU	SALOON	4	2,000.00	TP
198	MS7	HILUX	TOYOTA	VAN/BUS	17	2,494.00	TP
199	MO3	CKD	NISSAN	PICK UP/TRUCK	8	1,497.00	TP
200	MS8	NAVARA	NISSAN	PICK UP/TRUCK	8	2,488.00	TP

จากการนำข้อมูลใหม่ไปทดสอบนั้นสามารถสรุปผลลัพธ์ได้ดังนี้ MO1, MO2, MO3, MP3, MS5, MS6, MS7, MS8, MS9, VO2, VO5 มีจำนวน 67, 2, 4, 15, 13, 3, 66, 14, 4, 2, 10 จำนวนตามลำดับ ซึ่งจะเห็นได้ว่าวิธี Naïve Bayes นั้นได้ให้ผลลัพธ์ MO1 มากที่สุด

ตารางที่ 4.6 K-NN (K-Nearest Neighbors)

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	cover
1	MO1	FORTUNER	TOYOTA	SALOON	3	2,694.00	CO
2	MS7	FORTUNER	TOYOTA	SALOON	3	2,393.00	TP
3	MS7	HILUX	TOYOTA	SALOON	6	2,982.00	TP
4	MS8	HILUX	TOYOTA	SALOON	15	2,982.00	TP
5	MO1	RANGER	FORD	SALOON	4	2,200.00	CO
6	MS7	CRV	HONDA	SALOON	16	2,000.00	TP
7	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
8	MO1	HILUX	TOYOTA	PICK UP/TRUCK	7	2,982.00	CO
9	MO1	CX-5	MAZDA	SALOON	3	2,200.00	CO
10	MS7	E CLASS	BENZ	SALOON	21	2,295.00	TP
11	MO1	D-MAX	ISUZU	PICK UP/TRUCK	13	2,999.00	CO
12	MO1	COMMUTER	TOYOTA	VAN/BUS	7	2,982.00	CO
13	MS7	HIACE	TOYOTA	VAN/BUS	22	2,438.00	TF
14	MS5	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	TF
15	MS8	HILUX	TOYOTA	PICK UP/TRUCK	20	2,982.00	TP
16	MO3	CKD	MITSUBISHI	PICK UP/TRUCK	25	3,494.00	TP
17	MS7	COROLLA	TOYOTA	SALOON	24	1,600.00	TP
18	MS7	FRONTIER	NISSAN	PICK UP/TRUCK	14	2,488.00	TP
19	MO1	D-MAX	ISUZU	PICK UP/TRUCK	14	2,999.00	CO
20	MO1	CKD	NISSAN	PICK UP/TRUCK	18	2,499.00	CO
21	MS7	D-MAX	ISUZU	PICK UP/TRUCK	11	2,999.00	TP
22	MS7	COROLLA	TOYOTA	SALOON	10	1,598.00	TP
23	MS7	COROLLA	TOYOTA	SALOON	20	1,468.00	TF
24	MO1	HILUX	TOYOTA	PICK UP/TRUCK	14	2,982.00	CO
25	MS8	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TP
26	MS7	HILUX	TOYOTA	SALOON	8	2,694.00	TP
27	MS7	D-MAX	ISUZU	SALOON	14	2,999.00	TP
28	MS7	CIVIC	HONDA	SALOON	6	2,000.00	TP
29	MS8	CKD	MITSUBISHI	PICK UP/TRUCK	26	3,494.00	TP
30	MS9	CKD	MITSUBISHI	PICK UP/TRUCK	15	3,494.00	TF
31	MO1	SOLUNA	TOYOTA	SALOON	16	1,500.00	TF
32	MO1	ACCORD	HONDA	SALOON	9	1,997.00	CO
33	MO1	CKD	NISSAN	SALOON	7	2,496.00	TF
34	MS7	JAZZ	HONDA	SALOON	14	1,497.00	TF
35	MO1	HILUX	TOYOTA	VAN/BUS	6	2,982.00	CO
36	MO1	D-MAX	ISUZU	VAN/BUS	7	2,999.00	CO
37	MO1	CAB-4	ISUZU	SALOON	11	2,500.00	TF
38	MO1	SOLUNA	TOYOTA	SALOON	15	1,500.00	CO
39	MO1	TEANA	NISSAN	SALOON	5	1,997.00	CO
40	MO1	CIVIC	HONDA	SALOON	12	1,800.00	TF
41	MS7	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TF
42	MS7	S40	VOLVO	SALOON	19	1,948.00	TP

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	cover
43	MO1	CITY	HONDA	SALOON	6	1,497.00	TF
44	MO1	MU-X	ISUZU	SALOON	4	3,000.00	CO
45	MP3	JAZZ	HONDA	SALOON	2	1,497.00	TP
46	MO1	D-MAX	ISUZU	SALOON	6	2,499.00	CO
47	MS8	SOLUNA	TOYOTA	SALOON	15	1,496.00	TP
48	MO1	HILUX	TOYOTA	VAN/BUS	5	2,982.00	CO
49	MS8	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TP
50	MO1	TRITON	MITSUBISHI	PICK UP/TRUCK	8	2,477.00	CO
51	MS8	CKD	TOYOTA	PICK UP/TRUCK	13	2,499.00	TP
52	MO1	SOLUNA	TOYOTA	SALOON	5	1,500.00	CO
53	MO1	YARIS	TOYOTA	SALOON	4	1,200.00	CO
54	MO1	ALMERA	NISSAN	SALOON	4	1,200.00	CO
55	MS7	SOLUNA	TOYOTA	SALOON	21	2,000.00	TP
56	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
57	MO1	CIVIC	HONDA	SALOON	6	1,800.00	TF
58	MS8	XENON	TATA	PICK UP/TRUCK	8	2,850.00	TP
59	MO1	COOPER	MINI	SALOON	8	1,598.00	TF
60	MO1	JAZZ	HONDA	SALOON	7	1,500.00	TF
61	MO1	SOLUNA	TOYOTA	SALOON	6	1,497.00	TF
62	MO1	NAVARA	NISSAN	PICK UP/TRUCK	4	2,488.00	CO
63	MS7	E CLASS	BENZ	SALOON	22	2,295.00	TP
64	MS7	D-MAX	ISUZU	PICK UP/TRUCK	25	2,999.00	TP
65	MO1	HIACE	TOYOTA	VAN/BUS	11	2,494.00	TF
66	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TP
67	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	3,000.00	CO
68	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
69	MO1	COROLLA	TOYOTA	SALOON	6	1,798.00	CO
70	MO1	C CLASS	BENZ	SALOON	6	1,796.00	CO
71	MO1	D-MAX	ISUZU	PICK UP/TRUCK	7	2,999.00	TF
72	MO1	CIVIC	HONDA	SALOON	5	1,798.00	TF
73	MO1	RANGER	FORD	SALOON	6	2,200.00	TF
74	MO1	HILUX	TOYOTA	PICK UP/TRUCK	6	2,982.00	CO
75	MO1	2	MAZDA	SALOON	3	1,299.00	CO
76	MS8	COROLLA	TOYOTA	SALOON	15	1,598.00	TF
77	MO1	3	MAZDA	SALOON	9	1,600.00	CO
78	MS7	TRITON	MITSUBISHI	SALOON	4	2,500.00	TP
79	MO1	HILUX	TOYOTA	SALOON	11	2,494.00	CO
80	MS7	FIGHTER	MAZDA	PICK UP/TRUCK	12	2,982.00	TP
81	MO1	JAZZ	HONDA	SALOON	7	1,497.00	CO
82	MO1	D-MAX	ISUZU	SALOON	3	2,499.00	CO
83	MO1	CITY	HONDA	SALOON	6	1,497.00	TP
84	MS7	HILUX	TOYOTA	PICK UP/TRUCK	10	2,982.00	TP
85	MO1	SPACE WAGON	MITSUBISHI	SALOON	13	2,378.00	CO

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	cover
86	MO1	SWIFT ECO	SUZUKI	SALOON	4	1,242.00	CO
87	MO1	SOLUNA	TOYOTA	SALOON	7	1,500.00	CO
88	MO1	SPACE WAGON	MITSUBISHI	SALOON	11	2,378.00	TF
89	MS7	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
90	MO1	YARIS	TOYOTA	SALOON	3	1,197.00	CO
91	MO1	COROLLA	TOYOTA	SALOON	3	1,798.00	CO
92	MO1	SOLUNA	TOYOTA	SALOON	3	1,497.00	CO
93	MS7	C CLASS	BENZ	SALOON	5	1,800.00	TP
94	MS7	JAZZ	HONDA	SALOON	5	2,000.00	TP
95	MS7	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
96	MS7	HILUX	TOYOTA	PICK UP/TRUCK	13	2,982.00	TP
97	MS7	AVEO	CHEVROLET	SALOON	11	1,399.00	TP
98	MO1	HILUX	TOYOTA	SALOON	8	2,494.00	TF
99	MS7	CORONA	TOYOTA	SALOON	24	1,600.00	TP
100	MO1	INNOVA	TOYOTA	SALOON	9	1,998.00	TF
101	MO1	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
102	MS7	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TP
103	MO1	FORTUNER	TOYOTA	SALOON	12	2,982.00	TF
104	MS8	HILUX	TOYOTA	PICK UP/TRUCK	8	2,982.00	TF
105	MO1	CIVIC	HONDA	SALOON	3	1,798.00	CO
106	MO1	YARIS	TOYOTA	SALOON	3	1,197.00	CO
107	MO1	BT-50	MAZDA	SALOON	5	2,198.00	TF
108	MS7	PAJERO SPORT	MITSUBISHI	SALOON	10	2,500.00	TP
109	MO1	CIVIC	HONDA	SALOON	5	1,800.00	CO
110	MS7	SOLUNA	TOYOTA	SALOON	11	1,497.00	TP
111	MS7	CKD	TOYOTA	SALOON	12	2,494.00	TP
112	MS7	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
113	MO3	CKD	NISSAN	SALOON	28	1,998.00	TP
114	MO1	CITY	HONDA	SALOON	7	1,497.00	TF
115	MO1	ACCORD	HONDA	SALOON	10	2,354.00	TF
116	MO1	PAJERO SPORT	MITSUBISHI	SALOON	6	2,477.00	TP
117	MS7	SUNNY	NISSAN	SALOON	22	1,498.00	TP
118	MO1	TEANA	NISSAN	SALOON	14	2,400.00	TF
119	MS7	SOLUNA	TOYOTA	SALOON	18	1,498.00	TP
120	MO1	SOLUNA	TOYOTA	SALOON	6	1,500.00	TP
121	MS7	CKD	TOYOTA	VAN/BUS	5	1,497.00	TP
122	MO1	SOLUNA	TOYOTA	SALOON	13	1,500.00	CO
123	MS8	HILUX	TOYOTA	SALOON	13	2,982.00	TP
124	MO1	HILUX	TOYOTA	SALOON	9	2,982.00	TF
125	MS7	STRADA	MITSUBISHI	PICK UP/TRUCK	20	2,835.00	TP
126	MO1	FORTUNER	TOYOTA	SALOON	9	2,982.00	TF
127	MS7	HILUX	TOYOTA	PICK UP/TRUCK	11	2,982.00	TP
128	MO1	CITY	HONDA	SALOON	3	1,497.00	TF

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	cover
129	MS8	FIESTA	FORD	SALOON	5	1,499.00	TP
130	MO1	FORTUNER	TOYOTA	SALOON	3	2,982.00	TF
131	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
132	MO1	JAZZ	HONDA	SALOON	5	1,497.00	TP
133	MO1	SOLUNA	TOYOTA	SALOON	4	1,497.00	TP
134	MS7	TRITON	MITSUBISHI	PICK UP/TRUCK	4	2,477.00	TP
135	MS7	HILUX	TOYOTA	SALOON	8	2,494.00	TP
136	MS7	SPORT RIDER	TOYOTA	SALOON	18	2,000.00	TP
137	MS8	SOLUNA	TOYOTA	SALOON	15	1,500.00	TF
138	MO1	D-MAX	ISUZU	PICK UP/TRUCK	12	2,999.00	TF
139	MO1	C CLASS	BENZ	SALOON	6	1,800.00	TF
140	MO1	COMMUTER	TOYOTA	VAN/BUS	12	2,982.00	TF
141	MS8	COROLLA	TOYOTA	SALOON	15	1,598.00	TP
142	MS7	HILUX	TOYOTA	PICK UP/TRUCK	8	2,494.00	TP
143	MS7	BIG M	NISSAN	PICK UP/TRUCK	23	2,663.00	TP
144	MS8	D-MAX	ISUZU	PICK UP/TRUCK	8	2,999.00	TF
145	MS8	CAMRY	TOYOTA	SALOON	15	2,362.00	TP
146	MO1	D-MAX	ISUZU	PICK UP/TRUCK	6	2,999.00	CO
147	MO1	CITY	HONDA	SALOON	9	1,500.00	TF
148	MO3	ACCORD	HONDA	SALOON	27	2,000.00	TP
149	MO1	LANCER	MITSUBISHI	SALOON	7	2,000.00	CO
150	MS8	D-MAX	ISUZU	PICK UP/TRUCK	15	2,999.00	TF
151	MO1	JAZZ	HONDA	SALOON	5	1,497.00	CO
152	MO1	MG3	MG	SALOON	3	1,498.00	CO
153	MS8	ACCORD	HONDA	SALOON	10	2,354.00	TP
154	MO1	D-MAX	ISUZU	SALOON	4	2,500.00	TP
155	MO1	H-1	HYUNDAI	VAN/BUS	6	2,497.00	CO
156	MO1	COROLLA	TOYOTA	SALOON	5	1,598.00	CO
157	MO1	CIVIC	HONDA	SALOON	11	1,799.00	CO
158	MO1	FORTUNER	TOYOTA	SALOON	11	2,982.00	CO
159	MS8	CIVIC	HONDA	SALOON	15	1,688.00	TP
160	MO1	D-MAX	ISUZU	PICK UP/TRUCK	2	2,999.00	CO
161	MO1	SOLUNA	TOYOTA	SALOON	11	1,500.00	CO
162	MS7	S40	VOLVO	SALOON	20	1,948.00	TP
163	MO1	NAVARA	NISSAN	SALOON	2	2,488.00	CO
164	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	21	3,494.00	TP
165	MS8	C CLASS	BENZ	SALOON	23	2,159.00	TP
166	MO1	LANCER	MITSUBISHI	SALOON	10	1,600.00	TF
167	MO1	TRITON	MITSUBISHI	SALOON	7	2,351.00	TF
168	MO1	D-MAX	ISUZU	SALOON	2	1,850.00	CO
169	MO1	D-MAX	ISUZU	VAN/BUS	7	2,999.00	TF
170	MO3	CKD	MITSUBISHI	SALOON	29	1,298.00	TP
171	MO1	FORTUNER	TOYOTA	SALOON	2	2,400.00	CO

Row	Prediction	Model	Make	Body_Type	age_of_vehicle	Capacity	cover
172	MS8	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
173	MO1	SOLUNA	TOYOTA	SALOON-	3	1,497.00	TP
174	MO1	CITY	HONDA	SALOON	8	1,497.00	TP
175	MO1	SOLUNA	TOYOTA	SALOON	8	1,497.00	TF
176	MS8	TEANA	NISSAN	SALOON	13	2,300.00	TP
177	MS8	SUNNY	NISSAN	SALOON	14	1,769.00	TP
178	MO1	2	MAZDA	SALOON	5	1,498.00	TP
179	MS5	HILUX	TOYOTA	VAN/BUS	3	2,494.00	TF
180	MS8	HILUX	TOYOTA	PICK UP/TRUCK	13	2,494.00	TF
181	MS8	SOLUNA	TOYOTA	SALOON	6	1,497.00	TP
182	MS7	TROOPER	ISUZU	SALOON	21	3,168.00	TP
183	MO1	HILUX	TOYOTA	SALOON	7	3,000.00	CO
184	MO1	CITY	HONDA	SALOON	5	1,500.00	CO
185	MS8	D-MAX	ISUZU	PICK UP/TRUCK	5	2,999.00	TP
186	MS7	SOLUNA	TOYOTA	SALOON	12	1,497.00	TP
187	MO1	CAMRY	TOYOTA	SALOON	3	1,500.00	CO
188	MS7	COROLLA	TOYOTA	SALOON	17	1,600.00	TF
189	MO1	COMMUTER	TOYOTA	VAN/BUS	3	2,982.00	CO
190	MO1	SOLUNA	TOYOTA	SALOON	4	1,497.00	CO
191	MO1	CAPTIVA	CHEVROLET	SALOON	5	2,384.00	CO
192	MS8	CRV	HONDA	SALOON	15	2,000.00	TF
193	MS7	HILUX	TOYOTA	PICK UP/TRUCK	10	2,494.00	TP
194	MS7	COROLLA	TOYOTA	SALOON	24	2,000.00	TP
195	MS7	CKD	MITSUBISHI	PICK UP/TRUCK	24	3,494.00	TP
196	MS7	D-MAX	ISUZU	PICK UP/TRUCK	18	2,999.00	TP
197	MS7	D-MAX	ISUZU	SALOON	4	2,000.00	TP
198	MS7	HILUX	TOYOTA	VAN/BUS	17	2,494.00	TP
199	MS8	CKD	NISSAN	PICK UP/TRUCK	8	1,497.00	TP
200	MO1	NAVARA	NISSAN	PICK UP/TRUCK	8	2,488.00	TP

จากการนำข้อมูลใหม่ไปทดสอบนั้นสามารถสรุปผลลัพธ์ได้ดังนี้ MO1, MO3, MP3, MS5, MS7, MS8, MS9 มีจำนวน 108, 4, 1, 2, 56, 28, 1 จำนวน ตามลำดับ ซึ่งจะเห็นได้ว่าวิธี K-NN นั้นได้ให้ผลลัพธ์ MO1 มากที่สุด

สามารถสรุปได้ว่า จากการนำข้อมูลใหม่ไปทดสอบนั้นจะเห็นได้ว่า ผลลัพธ์ของวิธี Decision Tree, Naïve Bayes และ K-NN นั้นผลลัพธ์ที่มากที่สุดคือ MO1

4.5 ผลสรุป

ตารางที่ 4.7 ตารางผลสรุป

ค่าที่ได้ \ วิธี	Decision Tree	Naïve Bayes	K-NN
Accuracy	77.05%	70.38%	69.36%
Class Recall ของ MS7	100.00%	61.49%	63.26%
Class Precision ของ MS7	55.05%	55.01%	51.19%
Class Recall ของ MO1	99.08%	95.86%	98.90%
Class Precision ของ MO1	97.29%	96.75%	86.22%
ข้อมูลใหม่ที่ทำนายว่าเป็น MO1	175	67	108
ข้อมูลใหม่ที่ทำนายว่าเป็น MO3	2	4	4
ข้อมูลใหม่ที่ทำนายว่าเป็น MS7	14	66	56

จากตารางที่ 4.7 จะเห็นได้ว่าวิธีที่มีค่า Accuracy มากที่สุดคือ Decision Tree ซึ่งมีค่า 77.05% และข้อมูลใหม่ที่นำไปทำนายนั้นจะได้ว่าผลลัพธ์ของวิธี Decision Tree, Naïve Bayes และ K-NN นั้นผลลัพธ์ที่มากที่สุดคือ MO1

สำหรับเนื้อหาในบทต่อไปนั้นจะกล่าวถึงบทสรุปผลการวิจัยและข้อเสนอแนะต่างๆ

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้จะกล่าวถึงสรุปผลจากการทำงานวิจัยครั้งนี้และข้อเสนอแนะต่างๆ มีดังนี้

5.1 สรุปผลการวิจัย

งานวิจัยเรื่อง การทำเหมืองข้อมูลเพื่อวิเคราะห์แบบประกันภัยรถยนต์ มีวัตถุประสงค์เพื่อวิเคราะห์แบบประกันภัยรถยนต์ เพราะในปัจจุบันได้มีรูปแบบประกันภัยให้เลือกมากมายไม่ว่าจะเป็นประเภท 1, 2, 3, 2+, 3+ จึงได้นำข้อมูลมาศึกษาและผู้วิจัยก็ได้ทำการศึกษา Data Mining ทำให้เห็นวิธีที่เหมาะสมนั่นก็คือ Classification

ซึ่งเทคนิค Classification จะมีวิธี 3 วิธี คือ Decision Tree, Naïve Bayes และ K-Nearest Neighbors (K-NN) ต่อไปจะเป็นการเตรียมข้อมูลซึ่งจำนวนข้อมูลที่จะทำการศึกษานั้นมีทั้งหมด 5,000 ข้อมูล 15 แอตทริบิวต์ อยู่ในไฟล์ excel และได้มีการคัดเลือกและทำการลดขนาดข้อมูลนั้นคือไม่เอาประกันภัยภาคบังคับมาทำการวิเคราะห์ครั้งนี้และใช้ประเภทรถ Saloon, Van/Bus และ Pick Up ทำให้เหลือข้อมูลเพียง 2,134 ข้อมูล และเหลือเพียง 8 แอตทริบิวต์ และได้ทำการ save ข้อมูลเป็นไฟล์ .csv เพื่อที่จะใช้ในโปรแกรม RapidMiner Studio

ซึ่งทั้ง 3 วิธี ได้ให้ผลลัพธ์ที่มีค่าความถูกต้องมากที่สุดคือ Decision Tree มีค่าความแม่นยำคือ 77.05% ส่วนในขั้นตอนสุดท้ายนั้นเป็นการนำข้อมูลใหม่มาทำนายซึ่งมีข้อมูล 200 ข้อมูล จะทำการนำข้อมูลใหม่มาทำนายทั้ง 3 วิธี ผลลัพธ์ของวิธี Decision Tree, Naïve Bayes และ K-NN นั้นผลลัพธ์ที่มากที่สุดคือ MO1 นั่นก็คือประกันภัยชั้น 1

5.2 ข้อเสนอแนะ

1. ควรจะเพิ่มจำนวนข้อมูลในการพยากรณ์เพื่อความแม่นยำ การทำนายนั้นมีหลากหลายวิธีควรที่จะเลือกวิธีที่เหมาะสม

2. ข้อมูลนั้นเป็นสิ่งสำคัญในการทำ Data Mining ดังนั้นควรที่จะทำระบบคลังข้อมูลให้เกิดประโยชน์มากที่สุด

3. การทำ Data Mining ถ้าเรารู้จุดประสงค์ก็สามารถขยายฐานลูกค้าได้มากยิ่งขึ้น

4. งานวิจัยฉบับนี้เป็นเพียงการประยุกต์ใช้ในการทำ Data Mining เท่านั้น

เอกสารอ้างอิง

[1] นงคราญ คาวิชัย (2559). Practical Data Mining With RapidMiner Studio 7. มหาวิทยาลัยแม่โจ้, คณะวิทยาศาตร์, สาขาวิชาเทคโนโลยีสารสนเทศ. สืบค้นจาก https://www.slideshare.net/nongluk/practical-data-mining-with-rapid-miner-studio7

[2] สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) สืบค้นจาก http://www.oic.or.th/th/consumer

[3] กฤษณะ ไวยมัย, ธนารักษ์ ธีระมั่นคง, นิค เชอร์โคน และ สุภาวดี อิงศรีสว่าง (2014). Introduction to Data Mining Techniques. สืบค้นจากhttp://dataminingtrend.com/2014

[4] ทิพย์ธิดา วงศ์พิพันธ์. 2556. "การใช้เหมืองข้อมูลช่วยในการตัดสินใจการให้สินเชื่อ กรณีศึกษา : บริษัท กรุงไทยคาร์เร้นท์ แอนด์ ลิส จำกัด (มหาชน)." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีคอมพิวเตอร์และการสื่อสาร คณะวิศวกรรมศาสตร์, มหาวิทยาลัยธุรกิจบัณฑิตย์

ภาคผนวก

ภาคผนวก ก

แบบประกันรถยนต์

2 Save

ความคุ้มครอง		รหัส 110		รหัส 210		รหัส 320	
รถยนต์สูญหาย / ไฟไหม้	ต่อครั้ง	50,000	100,000	50,000	100,000	50,000	100,000
ความรับผิดต่อชีวิต ร่างกายบุคคลภายนอก	ต่อคน	300,000	300,000	300,000	300,000	300,000	300,000
	ต่อครั้ง	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000
ความรับผิดต่อทรัพย์สินบุคคลภายนอก	ต่อครั้ง	600,000	600,000	600,000	600,000	600,000	600,000
อุบัติเหตุส่วนบุคคล	ต่อคน	100,000 (5)	100,000 (5)	100,000 (3)	100,000 (3)	100,000 (3)	100,000 (3)
ค่ารักษาพยาบาล	ต่อคน	100,000 (5)	100,000 (5)	100,000 (3)	100,000 (3)	100,000 (3)	100,000 (3)
การประกันตัวผู้ขับขี่	ต่อครั้ง	200,000	200,000	200,000	200,000	200,000	200,000
เบี้ยประกันภัย (รวมภาษี อากร ยังไม่รวม พ.ร.บ.)	ต่อปี	**3,000**	**3,500**	**3,500**	**4,000**	**3,500**	**4,000**

<u>เงื่อนไขการรับประกันภัย</u>

- รถยนต์นั่ง (110)
- รถยนต์โดยสาร (210) ไม่เกิน 12 ที่นั่ง
- รถยนต์กระบะบรรทุก (320) นน ไม่เกิน 4 ตัน
- ไม่รับรถยนต์ที่มีการจดทะเบียนที่เป็นนิติบุคคล (ชื่อบริษัท)
- รับประกันภัยรถยนต์ทุกรุ่น/ยี่ห้อ
- มูลค่ารถยนต์ขณะที่เอาประกันภัย ต้องไม่ต่ำกว่าจำนวนเงินเอาประกันภัย
- อายุรถไม่เกิน 25 ปี นับจากวันที่จดทะเบียนรถยนต์
- ไม่คุ้มครองอุปกรณ์เสริมอื่น ๆ หรืออุปกรณ์ที่ไม่ใช่มาตรฐานของโรงงาน
- เป็นกรมธรรม์ Fixed Rate & Fixed Sum Insured ไม่มีส่วนลดทุกประเภท
- เงื่อนไขความคุ้มครอง และจำนวนเงินจำกัดความรับผิดต้องเป็นไปตาม Package ที่ BS กำหนด
- ไม่รับโอนส่วนลดประวัติดีทุกกรณี
- กรมธรรม์แบบเต็มปี (ไม่มีระยะสั้น หรือระยะยาว)

VMI - 3rd CLASS : THIRD PARTY ONLY

Veh.Code	STANDARD CLASSIC					STANDARD MASTER			
	Saloon		Van & Bus		Pick-up	Saloon		Van & Bus	
	110	120	210 <=20seats	220 <=20seats	320 <=4tons	110	120	210 <=20seats	220 <=20seats
TPBI	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	2,000,000	2,000,000	2,000,000	2,000,000
	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000
TPPD	2,500,000	2,500,000	2,500,000	2,500,000	2,500,000	5,000,000	5,000,000	5,000,000	5,000,000
PA	100,000 (7)	100,000 (7)	100,000 (12)	100,000 (12)	100,000 (3)	100,000 (7)	100,000 (7)	100,000 (12)	100,000 (12)
MED	100,000 (7)	100,000 (7)	100,000 (12)	100,000 (12)	100,000 (3)	100,000 (7)	100,000 (7)	100,000 (12)	100,000 (12)
BB	300,000	300,000	300,000	300,000	300,000	300,000	300,000	300,000	300,000
Net Prem.	3,723.00	4,189.00	5,213.00	6,533.00	4,560.00	4,468.00	4,840.00	5,864.00	7,468.00
+ Vat Stamp	**4,000**	**4,500**	**5,600**	**7,019**	**4,900**	**4,800**	**5,200**	**6,300**	**8,023**

เงื่อนไขการรับประกันภัย

- รถยนต์นั่ง (110)
- รถยนต์โดยสาร (210 , 220) ไม่เกิน 20 ที่นั่ง
- รถยนต์กระบะบรรทุก (320) นน ไม่เกิน 4 ตัน
- อนุโลมให้ผู้เอาประกันภัยที่เป็นนิติบุคคล (ซื้อบริษัท) ใช้รหัส 110 และ 210 ได้
- บริษัทฯ ขอสงวนสิทธิ์ในการพิจารณาต่ออายุกรณีหากมีการเรียกร้องสินไหมแต่ละปีเกิน 2 ครั้ง และ L/R เกิน 200% 2 ปี ติดต่อกันอาจพิจารณาไม่ต่ออายุ หรือปรับอัตราเบี้ยประกันภัยเพิ่มขึ้น
- อายุไม่เกิน 25 ปี นับจากวันที่จดทะเบียนรถยนต์
- เงื่อนไขความคุ้มครอง และจำนวนเงินจำกัดความรับผิดแบบ Master & Classic เท่านั้น
- กรมธรรม์แบบเต็มปี (ไม่มีระยะสั้น หรือระยะยาว)

ห้ามออกกรมธรรม์ย้อนหลังวันคุ้มครอง

PACKAGE	SINGLE PLAN I			SINGLE PLAN II			2 IN 1 HAPPY C+			2 IN 1 WOW			2 IN 1 HAPPY D		HAPPY D (2014)	
Veh.Code	Saloon	Van&Bus	Pick-up	Saloon	Van&Bus	Pick-up	Saloon	Van&Bus	Pick-up	Saloon	Van&Bus	Pick-up	Saloon	Pick-up	Saloon	Pick-up
	110	210 <=20seats	320 <=4tons	110	210 <=20seats	320 <=4tons	110	210 <=20seats	320 <=4tons	110	210 <=20seats	320 <=4tons	110	320 <=4tons	110	320 <=4tons
TPBI	300,000	300,000	300,000	1,000,000	1,000,000	1,000,000	300,000	300,000	300,000	1,000,000	1,000,000	1,000,000	300,000	300,000	300,000	300,000
	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000	10,000,000
TPPD	1,000,000	1,000,000	1,000,000	2,000,000	2,000,000	2,000,000	1,000,000	1,000,000	1,000,000	2,000,000	2,000,000	2,000,000	600,000	600,000	600,000	600,000
PA	70,000 (5)	70,000 (5)	70,000 (3)	100,000 (7)	100,000 (12)	100,000 (3)	70,000 (5)	70,000 (5)	70,000 (3)	100,000 (7)	100,000 (12)	100,000 (3)	50,000 (7)	50,000 (3)	100,000 (7)	100,000 (3)
MED	50,000 (5)	50,000 (5)	50,000 (3)	50,000 (7)	50,000 (12)	50,000 (3)	50,000 (5)	50,000 (5)	50,000 (3)	50,000 (7)	50,000 (12)	50,000 (3)	50,000 (7)	50,000 (3)	50,000 (7)	50,000 (3)
BB	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	200,000	100,000	100,000	100,000	100,000
Net Prem.	2,234	3,071	2,885	2,512	3,444	3,071	2,005	2,902	2,729	2,284	3,273	2,916	1,725	2,635	1,722	2,653
+ Vat Stamp	**2,400**	**3,300**	**3,100**	**2,700**	**3,700**	**3,300**	2,155	3,118	2,932	2,455	3,517	3,133	1,853	2,831	**1,850**	**2,850**
+ CMI							**2,800**	**4,300**	**3,900**	**3,100**	**4,700**	**4,100**	**2,499**	**3,799**		

เงื่อนไขการรับประกันภัย

- รถยนต์นั่ง (110)
- รถยนต์โดยสาร (210) ไม่เกิน 20 ที่นั่ง
- รถยนต์กระบะบรรทุก (320) นน ไม่เกิน 4 ตัน
- ไม่รับรถยนต์ที่มีการจดทะเบียนที่เป็นนิติบุคคล (ซื้อบริษัท)
- บริษัทฯ ขอสงวนสิทธิในการพิจารณาต่ออายุกรณี พิจารณาไม่ต่ออายุ หรือปรับอัตราเบี้ยประกันภัยเพิ่มขึ้น
- อายุไม่เกิน 25 ปี นับจากวันที่จดทะเบียนรถยนต์
- เงื่อนไขความคุ้มครอง และจำนวนเงินจำกัดความรับผิดต้องเป็นไปตาม Package ที่ BS กำหนด
- ไม่มีส่วนลดเบี้ยประกันภัย (Fleet, NCB, Others)
- ไม่รับโอนส่วนลดประวัติดีทุกกรณี
- กรมธรรม์แบบเต็มปี (ไม่มีระยะสั้น หรือระยะยาว)

Safe Guard

มีผลบังคับใช้จนถึง วันที่ 31 ธันวาคม 2561

Policy Type	WEF. 01.05.2017 - 31.12.2017	Safe Guard (Standard)				Safe Guard (Mini)				Safe Guard (Easy)				
	Policy Class	**MS8**				**MS7**				**MS6**				
	Commission	**ตาม Com. Table**				**ตาม Com. Table**				**10%**				
	ความคุ้มครอง ความรับผิดต่อบุคคลภายนอก ต่อชีวิต ร่างกาย อนามัย	500,000				500,000				400,000				บาท ต่อคน
		10,000,000				10,000,000				10,000,000				บาท ต่อครั้ง
	ทรัพย์สิน	1,000,000				1,000,000				1,000,000				บาท ต่อครั้ง
	ความคุ้มครองต่อภัยเพิ่มเติม													
	อุบัติเหตุส่วนบุคคล (5 ที่นั่ง)	100,000				50,000				50,000				บาท ต่อคน
	ค่ารักษาพยาบาล (5 ที่นั่ง)	100,000				50,000				50,000				บาท ต่อคน
	ประกันตัวผู้ขับขี่	300,000				300,000				200,000				บาท ต่อครั้ง
3+	ทุนประกันภัย ชนกับยานพาหนะทางบก*, น้ำท่วม	PLAN	Excess*	PLAN	No Excess	PLAN	Excess*	PLAN	No Excess	PLAN	Excess*	PLAN	No Excess	
	100,000	Safe 1	6,500	Safe 4	7,500	Safe 1	6,000	Safe 4	6,500	Safe 1	4,900	Safe 4	5,300	
	200,000	Safe 2	7,500	Safe 5	8,500	Safe 2	6,800	Safe 5	7,300					
	300,000	Safe 3	8,500	Safe 6	9,500	Safe 3	7,500	Safe 6	8,300					
2+	ทุนประกันภัย ชนกับยานพาหนะทางบก*, น้ำท่วม, สูญหาย (ทั้งคัน), ไฟไหม้	PLAN	Excess*	PLAN	No Excess	PLAN	Excess*	PLAN	No Excess	PLAN	Excess*	PLAN	No Excess	
	100,000	Safe 1	7,000	Safe 4	7,900	Safe 1	6,900	Safe 4	7,200					
	200,000	Safe 2	8,000	Safe 5	8,900	Safe 2	7,300	Safe 5	8,200					
	300,000	Safe 3	9,000	Safe 6	9,900	Safe 3	8,500	Safe 6	8,800					
เงื่อนไขการต่ออายุ ถ้า LR ต่ำกว่า 50%		ลดเบี้ยสูงสุด 500 บาท				ไม่มีส่วนลดเบี้ยปีต่ออายุ								
* Excess = 2,000 :		ผู้เอาประกันภัยต้องรับผิดชอบต่อความเสียหายส่วนแรก 2,000 บาท กรณีรถชนกับยานพาหนะทางบกและเป็นฝ่ายประมาทเท่านั้น												

อัตราเบี้ยและความคุ้มครองกรมธรรม์เฉพาะภัย (ประเภท 5) สำหรับรถใช้เพื่อการพาณิชย์

ความคุ้มครอง	ความรับผิดต่อบุคคลภายนอก		ความคุ้มครองตามเอกสารแนบท้าย		
วงเงินคุ้มครอง (บาท)	ชีวิต , ร่างกาย	ทรัพย์สิน	อุบัติเหตุส่วนบุคคล	ค่ารักษาพยาบาล	การประกันตัวผู้ขับขี่
ต่อคน	300,000		50,000	50,000	
ต่อครั้ง	10,000,000	1,000,000			300,000

ประเภทความคุ้มครอง	2+			3+		
⊙ ความเสียหายต่อรถยนต์จากการชนกับยานพาหนะทางบก	100,000	200,000	300,000	100,000	200,000	300,000
ความเสียหายส่วนแรกกรณีเป็นฝ่ายประมาท (ต่อครั้ง)	2,000	2,000	2,000	2,000	2,000	2,000
⊙ รถยนต์สูญหาย/ไฟไหม้	100,000	200,000	300,000	ไม่คุ้มครอง		

เบี้ยประกันภัย (ไม่รวม พ.ร.บ.)	2+			3+		
⊙ รถยนต์นั่งเพื่อการพาณิชย์ ไม่เกิน 7 ที่นั่ง (รหัส 120)	9,400	10,400	10,800	8,700	9,700	10,100
⊙ รถยนต์โดยสารเพื่อการพาณิชย์ ไม่เกิน 12 ที่นั่ง (รหัส 220)	13,640	16,306	18,612	9,700	10,700	11,100
⊙ รถยนต์บรรทุกเพื่อการพาณิชย์ ไม่เกิน 4 ตัน (รหัส 320)	10,400	11,700	13,100	9,700	10,700	11,100

เงื่อนไขการรับประกันภัย

1. เป็นแผนการประกันภัยสำหรับรถบุคคลหรือนิติบุคคลที่ใช้เพื่อการพาณิชย์ (ไม่ใช่รับจ้างสาธารณะหรือให้เช่า)
2. ทุนประกันภัยต้องไม่เกิน 100% ของมูลค่าตลาด และมูลค่ารถยนต์ขณะที่เอาประกันภัยต้องไม่ต่ำกว่าจำนวนเงินเอาประกันภัยตามแผนการประกันภัย
3. จำนวนที่นั่งสำหรับความคุ้มครองอุบัติเหตุส่วนบุคคลและค่ารักษาพยาบาล รหัส 120 / 220 / 320 จำนวนที่นั่งที่คุ้มครอง 7 / 12 / 3 ที่นั่งตามลำดับ
4. แผน 2+ คุ้มครองรถยนต์สูญหาย/ไฟไหม้ จำกัดอายุรถยนต์ไม่เกิน 15 ปี , แผน 3+ ไม่จำกัดอายุรถยนต์
5. รับประกันเฉพาะรถกลุ่ม 3-5 เท่านั้น , ยกเว้นรถ sports car และ imported car
6. อัตราเบี้ยนี้เป็นอัตราเบี้ยพิเศษคงที่ สำหรับซ่อมอู่มาตรฐาน และไม่ได้ขยายความคุ้มครองอุปกรณ์เสริมอื่นๆ ที่ไม่ใช่มาตรฐานโรงงาน
7. อัตราเบี้ยนี้เริ่มมีผลตั้งแต่วันที่ 1 กุมภาพันธ์ 2555 เป็นต้นไป

ภาคผนวก ข

ประวัติอย่างย่อของผู้จัดทำ