

การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่ม
การเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่ง ในประเทศอินเดีย
EFFICIENCY COMPARISON OF DATA MINING CLASSIFICATION
METHODS FOR CHRONIC KIDNEY DISEASE : A CASE STUDY
OF A HOSPITAL IN INDIA



สุรวัชร ศรีเปารยะ
Surawat Sripaoraya

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2559

KMITL-2016-SC-M-050-010

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่ม
การเป็นโรคไตเรื้อรัง : กรณีศึกษาโรงพยาบาลแห่งหนึ่ง ในประเทศอินเดีย
EFFICIENCY COMPARISON OF DATA MINING CLASSIFICATION
METHODS FOR CHRONIC KIDNEY DISEASE : A CASE STUDY
OF A HOSPITAL IN INDIA



T145910

สุรวัชร ศรีเปารยะ
Surawat Sripaoraya

เลขหมู่.....
เลขทะเบียน 145910
วันเดือนปี 27 มี.ค. 2560

b.....
i.....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์
ภาควิชาสถิติ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2559

KMITL-2016-SC-M-050-010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

EFFICIENCY COMPARISON OF DATA MINING CLASSIFICATION
METHODS FOR CHRONIC KIDNEY DISEASE : A CASE STUDY
OF A HOSPITAL IN INDIA



A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE
DEGREE OF MASTER IN STATISTICS

DEPARTMENT OF STATISTICS

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2016

KMITL-2016-SC-M-050-010

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2016

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ “การเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่ม การเป็นโรคไตเรื้อรัง : กรณีศึกษา
โรงพยาบาลแห่งหนึ่ง ในประเทศอินเดีย”
EFFICIENCY COMPARISON OF CLASSIFICATION METHODS FOR CHRONIC
KIDNEY DISEASE : CASE STUDY OF A HOSPITAL IN INDIA

ชื่อนักศึกษา นายสุรวัชร ศรีเปารยะ
รหัสประจำตัว 57605105
ปริญญา วิทยาศาสตร์มหาบัณฑิต (สาขาวิชาสถิติประยุกต์)
ภาควิชา สถิติ
อาจารย์ที่ปรึกษาวิทยานิพนธ์ รองศาสตราจารย์ สายชล สีนสมบูรณ์ทอง

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
ดร.บุญญสิทธิ์ วรจันทร์ ประธานกรรมการ ดร.ชานินทร์ ศรีสุวรรณภา อาจารย์บัณฑิตประจำ (ในสาขาวิชาที่เกี่ยวข้อง) ผศ.ดร.จุฬารัตน์ สีนสมบูรณ์ทอง ผู้ทรงคุณวุฒิจากภายนอกสถาบันฯ รศ.สายชล สีนสมบูรณ์ทอง อาจารย์ที่ปรึกษาวิทยานิพนธ์	

วัน/ เดือน/ ปี ที่สอบ 21 ธันวาคม พ.ศ.2559 เวลา 16.00-19.00 น.
สถานที่สอบ ณ ห้อง 115 อาคารจุฬารัตน์วลัยลักษณ์ 1

คณะวิทยาศาสตร์รับรองแล้ว

(รองศาสตราจารย์ ดร.ศุภณี ธนบุรีพัฒน์)
คณบดีคณะวิทยาศาสตร์

วันที่ 22 เดือน ๑๓ พ.ศ. ๕๙

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อผู้ใช้เห็นเข้าใช้ประโยชน์จากการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	EFFICIENCY COMPARISON OF DATA MINING CLASSIFICATION METHODS FOR CHRONIC KIDNEY DISEASE : A CASE STUDY OF A HOSPITAL IN INDIA
Students	Surawat Sriporaya
Student ID	57606105
Degree	Master of Science
Department	Statistics
Academic Year	2016
Thesis Advisor	Assoc.Prof. Saichon Sinsomboonthong

ABSTRACT

The objective of this research was to compare the efficiency of 7 data mining classification methods for chronic kidney disease from Apollo hospital, India that composed of K-Nearest Neighbor, Decision Tree, Artificial Neural Network, Support Vector Machine, Rule-Based, Logistic Regression and Naïve Bayes. This dataset was divided randomly into two groups—a training dataset and a testing dataset—ratio of 70:30 respectively. The efficiency measures were accuracy and mean square error. Based on these measures and the Apollo hospital dataset, the best classification method was the Decision Tree method that achieved an accuracy of 100% and a mean square error of 0.0059.

KEYWORDS: Logistic Regression, K-Nearest Neighbor, Artificial Neural Network, Support Vector Machine, Rule-Based, Decision Tree, Naïve Bayes

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีและมีความถูกต้องในเนื้อหาเนื่องด้วยได้รับความอนุเคราะห์จาก รศ.สายชล สินสมบุรณ์ทอง อาจารย์ที่ปรึกษา ที่ได้ให้ความรู้ ให้คำแนะนำ ให้ความช่วยเหลือ และตรวจสอบในการทำวิทยานิพนธ์จนสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบพระคุณ ดร.บุญยสิทธิ์ วรจันทร์ ดร.ชานินทร์ ศรีสุวรรณณา และ ผศ.ดร.จุฑาภรณ์ สินสมบุรณ์ทอง ที่ได้คำปรึกษาเกี่ยวกับงานวิจัยฉบับนี้ทั้งหมด สำหรับแนวคิด วิธีการ และข้อมูลทุกอย่างที่เป็นประโยชน์ในการวิเคราะห์วิธีการจำแนกกลุ่ม

ขอขอบพระคุณอาจารย์สาขาวิชาสถิติประยุกต์ทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้ พร้อมทั้งให้คำแนะนำ และช่วยเหลือในเรื่องต่างๆ มาโดยตลอด

ขอขอบพระคุณ คุณอัจฉรา แผ้วบาง และเจ้าหน้าที่สาขาวิชาสถิติประยุกต์ทุกท่าน ที่ให้ความอนุเคราะห์จัดหาอุปกรณ์ในการทำวิทยานิพนธ์ฉบับนี้

สุดท้ายนี้ขอขอบพระคุณบิดา-มารดา ของผู้จัดทำซึ่งสนับสนุนในด้านกำลังทรัพย์ และให้กำลังใจเสมอมาและขอบคุณเพื่อนๆ ทุกคน ที่ได้คำปรึกษา ช่วยเหลือในการทำงานมาโดยตลอด ทำให้วิทยานิพนธ์ฉบับนี้สำเร็จไปได้ด้วยดี

นาย สุรวีร

ศรีเปารยะ

สารบัญ

เรื่อง	หน้า
บทคัดย่อภาษาไทย	I
บทคัดย่อภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	VI
สารบัญรูป	VII
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญของงานวิจัย	1
1.2 วัตถุประสงค์	2
1.3 ขอบเขตของการศึกษา	2
1.4 นิยามศัพท์เฉพาะ	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	
2.1 การทำเหมืองข้อมูล	5
2.2 โรคไตเรื้อรัง	6
2.3 วิธีการแบ่งประเภทข้อมูลของวิธีการจำแนกกลุ่ม	14
2.4 วิธีความใกล้เคียงกันมากที่สุด	14
2.5 วิธีต้นไม้ตัดสินใจ	15
2.6 วิธีโครงข่ายประสาทเทียม	18
2.7 วิธีซัพพอร์ตเวกเตอร์แมชชีน	22
2.8 วิธีฐานกฎ	26
2.9 วิธีการถดถอยลอจิสติก	29
2.10 วิธีนาอ็ฟ เบย์	31
2.11 การเปรียบเทียบประสิทธิภาพของวิธีการจำแนกกลุ่ม	32
2.12 งานวิจัยที่เกี่ยวข้อง	33

สารบัญ (ต่อ)

	หน้า
บทที่ 3 วิธีการดำเนินงานวิจัย	
3.1 การเก็บรวบรวมข้อมูล	36
3.2 เครื่องมือที่ใช้ในการวิจัย	38
3.3 ขั้นตอนในการดำเนินงานวิจัย	38
3.4 วิธีการวิเคราะห์ข้อมูล	39
บทที่ 4 ผลการวิเคราะห์ข้อมูล	
4.1 ผลการแบ่งข้อมูลด้วยโปรแกรม SPSS	45
4.2 ผลการวิเคราะห์ด้วยวิธีความใกล้เคียงกันมากที่สุด	47
4.3 ผลการวิเคราะห์ด้วยวิธีต้นไม้ตัดสินใจ	49
4.4 ผลการวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม	51
4.5 ผลการวิเคราะห์ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน	60
4.6 ผลการวิเคราะห์ด้วยวิธีฐานกฎ	62
4.7 ผลการวิเคราะห์ด้วยวิธีการถดถอยลอจิสติก	63
4.8 ผลการวิเคราะห์ด้วยวิธีนาอ็ฟ เบย์	65
บทที่ 5 สรุปผล อภิปราย และข้อเสนอแนะ	
5.1 สรุปผลการวิจัย	67
5.2 อภิปรายผลการวิจัย	67
5.3 ข้อเสนอแนะ	69
บรรณานุกรม	70
ภาคผนวก ก	73
ภาคผนวก ข	100
ประวัติผู้เขียน	127

สารบัญตาราง

ตารางที่		หน้า
3-1	ตัวแปรอิสระและตัวแปรตามของข้อมูลผู้ป่วยโรคไตเรื้อรัง	36
4-1	การเปรียบเทียบค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของข้อมูลผู้ป่วยโรคไตเรื้อรัง	66
ก-1	คุณลักษณะและรายละเอียดของข้อมูลผู้ป่วยโรคไตเรื้อรัง	79
ก-2	ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ	80
ก-3	ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ	94



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
2-1 ตัวอย่างของ ความใกล้เคียงกันมากที่สุด	15
2-2 ส่วนประกอบของต้นไม้ตัดสินใจ	16
2-3 โครงข่ายของเซลล์ประสาท	18
2-4 ลักษณะโครงข่ายประสาทเทียมแบบส่งสัญญาณไปข้างหน้า	20
2-5 ลักษณะโครงข่ายประสาทเทียมแบบมีการย้อนกลับ	20
2-6 โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น	21
2-7 การขยายตัวของเส้นขอบ	22
2-8 เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง	23
2-9 รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้	25
2-10 รูปปัญหาต้นไม้ย่อยที่มีการทำซ้ำ	27
2-11 ปัญหาต้นไม้ย่อยที่มีการทำซ้ำในการมองแบบตัวอย่างคู่	27
2-12 อิทธิพลของกฎ	28
2-13 ฟังก์ชันลอจิสติก	29
2-14 เมทริกซ์ความสับสน	32
4-1 การเลือกคำสั่ง Select Cases	45
4-2 การเลือกคำสั่ง Random Sample	46
4-3 ผลจากการสุ่มตัวอย่าง	46
4-4 ผลการเปรียบเทียบค่า k ของวิธีความใกล้เคียงกันมากที่สุด	47
4-5 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด	48
4-6 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีต้นไม้ตัดสินใจ	49
4-7 แบบจำลองต้นไม้ตัดสินใจเพื่อใช้ในการทำนาย	50
4-8 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีโครงข่ายประสาทเทียม	51
4-9 ผลค่าถ่วงน้ำหนักบางค่าของโครงข่ายประสาทเทียม	52
4-10 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 0	52
4-11 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 1	53
4-12 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 2	53
4-13 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 3	54

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่	เนื้อหา	หน้า
4-14	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 4	54
4-15	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 5	55
4-16	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 6	55
4-17	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 7	56
4-18	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 8	56
4-19	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 9	57
4-20	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 10	57
4-21	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 11	58
4-22	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 12	58
4-23	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 13	59
4-24	ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 14	59
4-25	ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน	60
4-26	ผลสมการซัพพอร์ตเวกเตอร์แมชชีนของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ	61
4-27	ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีฐานกฎ	62
4-28	ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีการถดถอยลอจิสติก	63
4-29	ผลสมการถดถอยลอจิสติกของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ	64
4-30	ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรัง สำหรับการทดสอบตัวแบบด้วยวิธีนาอีฟ เบย์	65
4-31	การเปรียบเทียบค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ของข้อมูลผู้ป่วยโรคไตเรื้อรัง	66
ก-1	โปรแกรม Weka	74
ก-2	หน้าจอแสดงโปรแกรม Weka	74
ก-3	ไฟล์การสร้างตัวแบบ	75
ก-4	หน้าจอแสดงวิธีการจำแนกกลุ่ม	75
ก-5	ขั้นตอนการสร้างตัวแบบ	76
ก-6	หน้าจอแสดงผลการสร้างตัวแบบ	76
ก-7	ขั้นตอนการทดสอบตัวแบบ	77

สารบัญรูป (ต่อ)

รูปที่		หน้า
ก-8	ขั้นตอนการเลือก Classifier evaluation options	77
ก-9	หน้าจอแสดงผลการทดสอบตัวแบบ	78
ข-1	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)	101
ข-2	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)	101
ข-3	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=2)	102
ข-4	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=2)	102
ข-5	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=3)	103
ข-6	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=3)	103
ข-7	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)	104
ข-8	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีต้นไม้ตัดสินใจ	109
ข-9	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีต้นไม้ตัดสินใจ	109
ข-10	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีต้นไม้ตัดสินใจ	110
ข-11	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีโครงข่ายประสาทเทียม	112
ข-12	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีโครงข่ายประสาทเทียม	112
ข-13	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีโครงข่ายประสาทเทียม	113
ข-14	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน	115
ข-15	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน	115
ข-16	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน	116
ข-17	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีฐานกฎ	118
ข-18	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีฐานกฎ	118
ข-19	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีฐานกฎ	119
ข-20	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีการถดถอยลอจิสติก	121
ข-21	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีการถดถอยลอจิสติก	121
ข-22	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีการถดถอยลอจิสติก	122
ข-23	ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีนาอ็ฟ เบย์	124
ข-24	ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีนาอ็ฟ เบย์	124
ข-25	ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีนาอ็ฟ เบย์	125

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันในยุคของข้อมูลข่าวสาร องค์กรส่วนใหญ่มีข้อมูลที่ต้องจัดเก็บอยู่เป็นจำนวนมากมาย เช่น ระบบร้านค้าปลีก จะเก็บข้อมูลพนักงานในองค์กร ข้อมูลการซื้อขาย ข้อมูลสินค้าและข้อมูลลูกค้า เป็นต้น จะเห็นได้ว่ายิ่งองค์กรหรือรูปแบบธุรกิจมีขนาดใหญ่เท่าไร ย่อมทำให้การเก็บสะสมข้อมูลสำหรับองค์กรต่างๆ มีจำนวนมากขึ้น การเก็บข้อมูลจำนวนมากเหล่านี้ลงในฐานข้อมูลเป็นวิธีที่นิยมใช้ในหลายองค์กร แต่ระบบการจัดการฐานข้อมูลทั่วไปไม่สามารถจัดการกับข้อมูลเหล่านี้ได้อย่างมีประสิทธิภาพ เนื่องจากใช้เวลานานในการดึงข้อมูลที่มีความสำคัญออกมาวิเคราะห์ ดังนั้นจึงได้เกิดเทคโนโลยีในการวิเคราะห์ข้อมูลที่มีความสำคัญออกมาจากแหล่งเก็บข้อมูลขนาดใหญ่ เรียกเทคโนโลยีนี้ว่า การทำเหมืองข้อมูล หรือ การขุดค้นข้อมูล (Data Mining)

ข้อมูลสารสนเทศถือว่าเป็นข้อมูลที่สำคัญในการนำมาประกอบการตัดสินใจ จัดทำงานวิจัย จัดทำแผนนโยบายและแผนยุทธศาสตร์ของหน่วยงานภาครัฐและภาคเอกชน สำหรับข้อมูลสารสนเทศที่ดีจะต้องมีความน่าเชื่อถือ มีความเป็นปัจจุบัน ทันสมัย ทันเวลา ทันต่อเหตุการณ์และสามารถนำมาใช้งานได้อย่างมีประสิทธิภาพ เพื่อให้เกิดประโยชน์สูงสุดต่อหน่วยงาน แต่การนำข้อมูลมาใช้ในบางครั้งอาจไม่สามารถใช้งานได้อย่างเต็มประสิทธิภาพและไม่ตรงตามความต้องการของหน่วยงานหรือองค์กร เนื่องจากอาจมีการนำข้อมูลไปวิเคราะห์โดยการจำแนกผิดวิธี ทำให้ความถูกต้องและผลลัพธ์ที่ได้จากการประมวลผลมีค่าลดลงหรืออาจจะได้ค่าผลลัพธ์ที่คลาดเคลื่อน ไม่ตรงกับข้อเท็จจริง สำหรับการจำแนกเหล่านี้สามารถทำได้หลายวิธี ซึ่งแต่ละวิธีจะมีความเหมาะสมกับข้อมูลแตกต่างกันไป ดังนั้นผู้ใช้งานหรือผู้วิจัยต้องเลือกวิธีจำแนกกลุ่มที่เหมาะสมกับข้อมูลโรคไตเรื้อรัง

ในช่วงที่ผ่านมา มีผู้วิจัยได้ศึกษาวิธีการจำแนกกลุ่มด้วยวิธีการต่างๆ เพื่อหารูปแบบที่มีความเหมาะสมและมีประสิทธิภาพ ให้ค่าผลลัพธ์ใกล้เคียงค่าจริงมากที่สุด เช่น การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มรูปแบบการเรียนรู้ VARK ด้วยวิธีนาอิว เบย์ (Naive Bayes) วิธีต้นไม้ตัดสินใจ (Decision Tree) และวิธีฐานกฎ (Rules-Based) ผลการศึกษาพบว่าวิธีต้นไม้ตัดสินใจให้ประสิทธิภาพสูงสุด (อรนุช และมนชัย, 2557) หรือการจำแนกบุคคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้ โดยข้อมูลที่ใช้ในการทดลองได้จากหน่วยงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง ด้วยอัลกอริทึม 3 แบบ ได้แก่ Decision Tree (J48), OneR และ Naive Bayes จากการทดลองพบว่าอัลกอริทึมที่ให้ค่าความถูกต้องในการทำนายสูงสุด คือ J48 (บุษกร, 2551) หรือการเปรียบเทียบปัจจัยโรคประจำตัวผู้สูงอายุโดยใช้อัลกอริทึมการจำแนกกลุ่ม J48 และ Naive Bayes

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดสอบพบว่าวิธีต้นไม้ตัดสินใจโดยใช้อัลกอริทึม J48 ให้ค่าความถูกต้องสูงกว่าวิธีนาอิว เบย์ (เบญจภาคี, 2558) และการทดสอบประสิทธิภาพการพยากรณ์สถานะการชำระหนี้ลูกหนี้ธนาคารด้วยวิธีต้นไม้ตัดสินใจ (Decision Tree) วิธีการถดถอยลอจิสติก (Logistic Regression) และวิธีนาอิว เบย์ (Naive Bayes) ผลการทดสอบพบว่าวิธีต้นไม้ตัดสินใจให้ค่าความแม่นยำสูงที่สุด รองลงมาคือวิธีการถดถอยลอจิสติกและวิธีนาอิว เบย์ ตามลำดับ (ชณัฐตาภรณ์, 2557)

จากการศึกษาพบว่าผู้วิจัยได้ให้ความสำคัญกับปัญหาที่ต้องการวิธีการจำแนกกลุ่มข้อมูลที่ต้องการ แม่นยำ และรวดเร็ว โดยผู้วิจัยมีความสนใจที่จะวิจัยเกี่ยวกับข้อมูลทางด้านการแพทย์ ซึ่งจะนำข้อมูลผู้ป่วยโรคไตเรื้อรังมาทำการวิเคราะห์ในงานวิจัยครั้งนี้ เพื่อเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มด้วยวิธีทางด้านสถิติ 7 วิธี ได้แก่ วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) วิธีต้นไม้ตัดสินใจ (Decision Tree) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) วิธีฐานกฎ (Rule-Based) วิธีการถดถอยลอจิสติก (Logistic Regression) และวิธีนาอิว เบย์ (Naive Bayes)

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มของผู้ป่วยโรคไตเรื้อรัง 7 วิธี คือ วิธีความใกล้เคียงกันมากที่สุด วิธีต้นไม้ตัดสินใจ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน วิธีฐานกฎ วิธีการถดถอยลอจิสติก และวิธีนาอิว เบย์

1.3 ขอบเขตของการศึกษา

ในการทำวิจัยได้ใช้ชุดข้อมูลผู้ป่วยโรคไตเรื้อรังของโรงพยาบาลสอพลโล ประเทศอินเดีย โดยมีข้อมูลผู้ป่วยโรคไตเรื้อรังจำนวน 400 ระเบียบ แบ่งข้อมูลอย่างสุ่มออกเป็น 2 ส่วน ในอัตราส่วนละ 70:30 ในข้อมูลส่วนที่ 1 ข้อมูลเรียนรู้ (Training data) นำไปสร้างตัวแบบ (Model) 70 เปอร์เซ็นต์ หรือคิดเป็น 280 ระเบียบ และข้อมูลส่วนที่ 2 ข้อมูลการทดสอบ (Testing data) นำไปทดสอบตัวแบบ 30 เปอร์เซ็นต์ หรือคิดเป็น 120 ระเบียบ ซึ่งวิธีที่ใช้ในการทดสอบในครั้งนี้มี 7 วิธี

(http://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease)

1. วิธีความใกล้เคียงกันมากที่สุด ใช้อัลกอริทึมชนิด IBk
2. วิธีต้นไม้ตัดสินใจ ใช้อัลกอริทึมชนิด J48
3. วิธีโครงข่ายประสาทเทียม ใช้อัลกอริทึมชนิด Multilayer Perceptron
4. วิธีซัพพอร์ตเวกเตอร์แมชชีน ใช้อัลกอริทึมชนิด Polynomial Kernel
5. วิธีฐานกฎ ใช้อัลกอริทึมชนิด Decision Table
6. วิธีการถดถอยลอจิสติก
7. วิธีนาอิว เบย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และผู้วิจัยได้นำวิธีเหล่านี้มาประยุกต์กับงานวิจัย โดยนำมาเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มในการทำวิทยานิพนธ์ครั้งนี้ ผู้วิจัยได้สนใจศึกษาวิธีการจำแนกกลุ่ม (Classification) เพื่อเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มโดยพิจารณาจากค่าความถูกต้อง (Accuracy) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE)

1.4 นิยามศัพท์เฉพาะ

การทำเหมืองข้อมูล (Data Mining) หมายถึง กระบวนการนำข้อมูลที่มีอยู่ในฐานข้อมูลขนาดใหญ่มาศึกษา วิเคราะห์ ทำความเข้าใจ และนำผลลัพธ์ที่ได้จากการศึกษามาใช้ในการตัดสินใจทางธุรกิจ (Connolly, T. M. and Begg, C. E. 2002)

การจำแนกกลุ่ม (Classification) หมายถึง กระบวนการสร้างโมเดลสำหรับการจัดกลุ่มหรือประเภทของข้อมูลโดยการวิเคราะห์ข้อมูลจากระบบเพื่อเรียนรู้ โดยแต่ละแถวประกอบด้วยข้อมูลเชิงคุณลักษณะที่มีลักษณะเป็นข้อมูลเชิงตัวเลขที่มีค่าต่อเนื่อง (พรคิต อ้นขาว และพยุง มีสัจ, 2558)

ความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) หมายถึง กระบวนการตัดสินใจของกลุ่มสำหรับแทนเงื่อนไขหรือกรณีใหม่ โดยการตรวจสอบจำนวนบางจำนวนหรือเงื่อนไขที่เหมือนกันหรือใกล้เคียงกันมากที่สุด ใช้เวลาในการคำนวณสูงเพราะการคำนวณเป็นการเพิ่มขึ้นแบบแฟคทอเรียล (Factorial) ตามจุดทั้งหมด (วาทีนี น้อยเพียร และคณะ, 2553)

ต้นไม้ตัดสินใจ (Decision Tree) หมายถึง กระบวนการในการจัดแบ่งข้อมูล โดยจะมีลักษณะการทำงานเหมือนโครงสร้างต้นไม้ โดยที่แต่ละโหนด (Node) จะแสดงคุณลักษณะ (Attribute) ที่ใช้ทดสอบข้อมูล แต่ละกิ่งจะแสดงผลในการทดสอบ และโหนดใบ (Leaf Node) จะแสดงกลุ่ม (Class) ที่กำหนดไว้ ซึ่งต้นไม้ตัดสินใจนี้ง่ายต่อการเข้าใจและเป็นวิธีที่ค่อนข้างแพร่หลาย (ภัทร์พงศ์ พงศ์ภัทรกานต์, 2552)

โครงข่ายประสาทเทียม (Artificial Neural Network) หมายถึง ตัวแบบทางคณิตศาสตร์สำหรับประมวลผลสารสนเทศ ได้มาจากการศึกษาข่ายงานไฟฟ้าชีวภาพ (Bioelectric Network) เกิดจากการเชื่อมต่อระหว่างเซลล์ประสาท เป็นการเลียนแบบวิธีการทำงานของสมองมนุษย์ ซึ่งเป็นรูปแบบการคำนวณที่ค่อนข้างซับซ้อน โดยนำข้อมูลต่าง ๆ ไปใช้ในการวิเคราะห์ ตีความหรือคาดคะเน (วาทีนี น้อยเพียร และคณะ, 2553)

ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) หมายถึง สมการที่ใช้ในการจำแนกค่าคุณลักษณะของ 2 กลุ่มที่วางตัวอยู่ในพื้นที่คุณลักษณะ (Feature Space) ออกจากกันโดยจะสร้างเส้นแบ่งที่เป็นเส้นตรงขึ้นมา เพื่อให้ทราบว่าเส้นตรงที่แบ่ง 2 กลุ่มออกจากกันนั้นเส้นตรงใดดีที่สุด (ภัทร์พงศ์ พงศ์ภัทรกานต์, 2552)

ฐานกฎ (Rule-Based) หมายถึง การใช้ชุดลำดับของกฎมาสร้างรูปแบบการแยกประเภทข้อมูล โดยส่วนใหญ่แล้วจะใช้กฎที่เป็น ถ้า...แล้ว ซึ่งเป็นกฎอย่างง่าย (อรนุช พันโท และมนชัย เอกสารนี้เป็นเพียงตัวอย่าง, 2557) สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การถดถอยลอจิสติก (Logistic Regression) หมายถึง การถดถอยเชิงเส้นใช้เพื่อประมาณค่าความสัมพันธ์ระหว่างตัวแปรตามซึ่งเป็นตัวแปรเชิงคุณภาพ และตัวแปรอิสระซึ่งเป็นได้ทั้งตัวแปรเชิงปริมาณหรือคุณภาพ และจะนำสมการการถดถอยลอจิสติกที่ได้ไปประมาณหรือพยากรณ์ค่าของตัวแปรตามเมื่อกำหนดค่าของตัวแปรอิสระ (สายชล สินสมบูรณ์ทอง, 2559)

นาอิว เบย์ (Naïve Bayes) หมายถึง เครื่องจักรเรียนรู้ที่อาศัยหลักการความน่าจะเป็น (Probability) ตามทฤษฎีนาอิว เบย์ ซึ่งมีอัลกอริทึมที่ไม่ซับซ้อน เป็นอัลกอริทึมในการจำแนกข้อมูล โดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ เป็นการจำแนกข้อมูลโดยใช้ความน่าจะเป็นและคำนวณการแจกแจงความน่าจะเป็นตามสมมติฐานที่ตั้งไว้ให้กับข้อมูล (วรพจน์ สุเมธาวัฒนา และคณะ, 2557)

อัลกอริทึม (Algorithm) คือ กระบวนการแก้ปัญหาที่สามารถเข้าใจได้ มีลำดับหรือวิธีการในการแก้ไขปัญหาใดปัญหาหนึ่งอย่างเป็นขั้นตอนและชัดเจน (สมชาย ประสิทธิ์จตุระกุล, 2545)

1.5 ประโยชน์ที่คาดว่าจะได้รับ

สามารถนำผลการเปรียบเทียบประสิทธิภาพที่ได้ ใช้เป็นแนวทางในการเลือกวิธีการจำแนกกลุ่มที่เหมาะสมที่สุดสำหรับข้อมูลโรคไตเรื้อรัง

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้ ผู้วิจัยได้นำเสนอเนื้อหาที่เน้นถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องโดยมีรายละเอียดของเนื้อหาประกอบด้วยหัวข้อย่อย 12 หัวข้อ ดังนี้

- 2.1 การทำเหมืองข้อมูล (Data Mining)
- 2.2 โรคไตเรื้อรัง (Kidney Disease)
- 2.3 วิธีการแบ่งประเภทข้อมูลของวิธีการจำแนกกลุ่ม (Classification)
- 2.4 วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor)
- 2.5 วิธีต้นไม้ตัดสินใจ (Decision Tree)
- 2.6 วิธีโครงข่ายประสาทเทียม (Artificial Neural Network)
- 2.7 วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine)
- 2.8 วิธีฐานกฎ (Rule-Based)
- 2.9 วิธีการถดถอยลอจิสติก (Logistic Regression)
- 2.10 วิธีนาอิว เบย์ (Naïve Bayes)
- 2.11 การเปรียบเทียบประสิทธิภาพในการทำงานนายผลการจำแนก
- 2.12 งานวิจัยที่เกี่ยวข้อง

2.1 การทำเหมืองข้อมูล (Data Mining) (รุจิรา ธรรมสมบัติ, 2554)

การทำเหมืองข้อมูลคือการวิเคราะห์ข้อมูลเพื่อแยกประเภท จำแนกรูปแบบและความสัมพันธ์ของข้อมูลจากฐานข้อมูลที่มีขนาดใหญ่หรือคลังข้อมูล โดยมีวิธีต่าง ๆ หลายวิธี ซึ่งรูปแบบการทำเหมืองข้อมูลนั้นได้รวบรวมความรู้จากหลายแขนงเข้าไว้ด้วยกันซึ่งประกอบด้วยระบบการเรียนรู้ของเครื่องจักร (Machine Learning) ร่วมกับวิทยาศาสตร์สารสนเทศ (Information Science) สถิติ (Statistic) และระบบฐานข้อมูล (Database System) โดยทั่วไปแล้ววิธีที่นำมาใช้ส่วนใหญ่มี 5 ประเภท

2.1.1 วิธีการจำแนกกลุ่ม (Classification)

เป็นวิธีในการจำแนกกลุ่มข้อมูลด้วยคุณลักษณะต่างๆ ที่ได้มีการกำหนดไว้แล้ว วิธีนี้เหมาะกับการสร้างตัวแบบเพื่อการพยากรณ์ค่าข้อมูล (Predictive Modeling) ในอนาคตจากการที่ได้จำแนกกลุ่มข้อมูลตัวอย่างไว้แล้ว ซึ่งในลักษณะดังกล่าวเรียกว่า การเรียนรู้แบบมีผู้สอน (Supervised Learning) วิธีการจำแนกกลุ่มเป็นกระบวนการสร้างตัวแบบเพื่อจัดข้อมูลให้อยู่ในกลุ่มที่กำหนด ตัวอย่างเช่น การแบ่งประเภทลูกค้าว่าเชื่อถือได้หรือไม่ ซึ่งเป็นการสร้างตัวแบบโดยการเรียนรู้จากข้อมูลที่ได้กำหนดไว้เรียบร้อยแล้ว

เอกสารนี้เป็นเอกสารที่รวบรวมไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 วิธีการค้นหาความสัมพันธ์ (Association Rule Discovery)

เป็นวิธีที่ใช้ในการค้นหาความสัมพันธ์ของฐานข้อมูลที่มีขนาดใหญ่ เพื่อที่จะทำการวิเคราะห์ข้อมูลและหาสิ่งที่ซ่อนอยู่ในข้อมูลนั้น เช่น การวิเคราะห์ข้อมูลการซื้อขายในซูเปอร์มาร์เก็ต เพื่อทำการวางแผนการส่งเสริมการขาย (Promotion) และเตรียมการวางแผนการเรียงชั้นวางสินค้า (Shelf) เช่น การวางน้ำอัดลมกับข้าวโพดคั่วไว้ใกล้กัน

2.1.3 วิธีการจัดกลุ่ม (Clustering)

เป็นวิธีการลดขนาดของข้อมูลด้วยการรวมกลุ่มตัวแปรที่มีลักษณะเดียวกันไว้ด้วยกัน ทำให้สามารถค้นหาข้อมูลที่ถูกกลืนหายไปได้ วิธีนี้มักถูกใช้เป็นขั้นตอนเบื้องต้นในการทำเหมืองข้อมูล และเหมาะกับข้อมูลที่ยังไม่มีกลุ่มอย่างชัดเจน จึงทำการรวมกลุ่มเพื่อหากลุ่มต่างๆ ของข้อมูล โดยจำนวนกลุ่มของข้อมูลแทนด้วย k ซึ่งผู้ที่ใช้วิธีนี้จะเป็นผู้กำหนดจำนวนกลุ่ม วิธีนี้อาจเรียกว่าการจัดกลุ่มแบบเฉลี่ย k กลุ่ม (K-mean Clustering)

2.1.4 วิธีการหาค่าที่แตกต่างจากค่ามาตรฐาน (Deviation Detection)

เป็นวิธีในการหาค่าที่แตกต่างไปจากค่ามาตรฐาน หรือค่าที่คาดคิดว่าต่างไปมากน้อยเพียงใด โดยทั่วไปมักใช้วิธีทางสถิติหรือการแสดงให้เห็นภาพ สำหรับวิธีนี้ใช้ในการตรวจสอบลายเซ็น หรือปลอมบัตรเครดิต เป็นต้น

2.1.5 วิธีการวิเคราะห์ลำดับ (Sequential Analysis)

เป็นวิธีในการวิเคราะห์ลำดับเพื่อค้นหารูปแบบของการปรากฏของข้อมูล ซึ่งปรากฏในรายการที่แยกออกมา เช่น ถ้าผู้ซื้อซื้อสินค้า A แล้วเขาจะซื้อสินค้า B ในภายหลัง วิธีนี้จะแตกต่างจากวิธีการค้นหาความสัมพันธ์ เพราะคำนึงถึงลำดับการซื้อด้วย

2.2 โรคไตเรื้อรัง (Chronic Kidney Disease) (สุขฤทัย ฉิมชาติ, 2556)

2.2.1 ความหมาย

โรคไตเรื้อรัง หมายถึง ภาวะที่ไตถูกทำลายเป็นเวลา 3 เดือน หรือมากกว่า 3 เดือน มีความผิดปกติทางพยาธิสภาพ หรือมีตัวบ่งชี้ว่าไตถูกทำลายจากความผิดปกติของเลือดหรือปัสสาวะ หรือการตรวจทางรังสีหรืออัตราการกรองของไตลดลงน้อยกว่า 60 มิลลิลิตร/นาที/พื้นผิวร่างกาย 1.73 ตารางเมตร เป็นเวลา 3 เดือน หรือมากกว่า 3 เดือน

2.2.2 สาเหตุของโรคไตเรื้อรัง

1) สาเหตุนอกไต ได้แก่

1.1) เบาหวาน พบว่ามีผู้ป่วยที่เป็นเบาหวานชนิดที่ 1 ที่พึ่งอินซูลิน 20-50% ที่ทำให้เกิดไตวายเรื้อรังระยะสุดท้ายภายในเวลา 20-30 ปีที่เริ่มรักษาด้วยการให้อินซูลินและเบาหวานยังทำให้เกิดโรคไตเรื้อรังได้ถึงร้อยละ 30-40 และทำให้เกิดไตวายเรื้อรังระยะสุดท้ายได้ถึงร้อยละ 45 นอกจากนั้น

เอกสารนี้เป็นเบาหวานยังเป็นปัจจัยเสี่ยงที่ทำให้เกิดโรคหัวใจและหลอดเลือด ความดันโลหิตสูงและไขมันในเลือด ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สูงได้ เบาหวานทำให้มีความผิดปกติของหลอดเลือดที่ไกลเมอรูลัสหรือหลอดเลือดไต ทำให้หลอดเลือดแข็ง เพิ่มแรงต้านของหลอดเลือดที่ไตและระบบความดันโลหิตสูงขึ้น ไตได้รับเลือดน้อยลงและขาดเลือด จึงทำให้เกิดไตล้มเหลวตามมา

1.2) ความดันโลหิตสูง พบว่าความดันโลหิตสูงเป็นสาเหตุที่ทำให้เกิดโรคไตเรื้อรังได้ถึงร้อยละ 28 เนื่องจากไตจึงจำเป็นต้องได้รับเลือดมาเลี้ยงเป็นจำนวนมากจากการบีบตัวของหัวใจซึ่งมีผลต่ออัตราการกรองและการทำหน้าที่ของไต ความดันโลหิตสูงจึงทำให้เลือดมาเลี้ยงที่ไตลดลง ทำให้การทำหน้าที่ของไตผิดปกติเช่นกัน ความดันโลหิตสูงเกิดเนื่องจากหลอดเลือดแดงที่ไตตีบแข็ง หรือขาดเลือด ทำให้เลือดมาเลี้ยงที่ไตลดลงและกระตุ้นระบบเรนินแองจิโอเทนซิน อัลโดสเทอโรน ทำให้เพิ่มความดันโลหิต นอกจากนี้ ความดันโลหิตสูงยังเกี่ยวข้องกับโรคของเนื้อไต เช่น Glomerulonephritis, Polycystic Disease และ Pyelonephritis เป็นต้น ทำให้ไตขับน้ำและเกลือได้ลดลง มีการคั่งของน้ำและเกลือเพิ่มขึ้น

1.3) ความดันโลหิตต่ำ ภาวะช็อกจากหัวใจและหลอดเลือด หรือความดันโลหิตต่ำ มีผลต่อการทำหน้าที่ของไต ทำให้หลอดเลือดที่ไตหดตัว เลือดไปเลี้ยงที่ไตลดลง

1.4) Rhabdomyolysis เกิดจากการบาดเจ็บของเนื้อเยื่อกระดูกและกล้ามเนื้อ ทำให้หลัง Myoglobin และ Intracellular Substances เข้าไปในเลือด ซึ่งเกิดจากหลังได้รับการบาดเจ็บที่รุนแรง การออกกำลังกายที่หนัก มีการชก Heat Stroke ภาวะที่ไม่รู้สึกตัวที่นานๆ การได้รับยาเกินขนาดและผลข้างเคียงจากการใช้ยา Statin ทำให้มีไข้ คลื่นไส้ อาเจียน กล้ามเนื้ออ่อนแรง ปวดกล้ามเนื้อและบวม โพแทสเซียมสูง ฟอสเฟตสูง กรดยูริกในเลือดสูง ระดับของ Creatinin Kinase สูงขึ้น

1.5) โรคระบบหัวใจและหลอดเลือด มีผลต่อปริมาณเลือดที่ออกจากหัวใจ และระบบไหลเวียนเลือด ซึ่งมีผลต่อการทำหน้าที่ของไต ทำให้ไตลดการขับน้ำและโซเดียม มีการคั่งของน้ำในหลอดเลือด ทำให้เกิดอาการบวม

1.6) โรคของหลอดเลือดส่วนปลาย ได้แก่ การเกิดลิ่ม เลือดอุดตันในหลอดเลือด (Thromboembolic) ภาวะ Disseminated Intravascular Coagulopathy มีผลต่อระบบการไหลเวียนของเลือดที่ไต เป็นสาเหตุให้ไตขาดเลือด

1.7) การติดเชื้อในกระแสเลือดอาจมีผลต่อการทำหน้าที่ของไตที่มีผลต่อระบบไหลเวียนเลือด ทำให้ความดันโลหิตต่ำและจะกระตุ้นระบบภูมิคุ้มกันของร่างกาย ทำให้เกิด Glomerulonephritis

1.8) การตั้งครรภ์มีผลต่อการทำหน้าที่ของไต การตั้งครรภ์ในไตรมาสแรกทำให้ไตมีขนาดใหญ่ขึ้น ซึ่งอาจจะคงอยู่ 9-12 สัปดาห์ ทำให้อัตราการกรองของไตเพิ่มขึ้น 30-50% ระหว่างตั้งครรภ์ ทำให้ Creatinine Clearance เพิ่มขึ้น การขับกรดยูริกลดลง การตั้งครรภ์อาจทำให้โปรตีนในปัสสาวะเพิ่มขึ้น ปัสสาวะมากขึ้น และปัสสาวะบ่อยในตอนกลางคืน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) สารที่มีพิษต่อไตจะทำลายเซลล์ของไต ทำให้ไตได้รับบาดเจ็บ เกิด Acute Tubular Necrosis สารที่มีพิษต่อไต ได้แก่ ยาปฏิชีวนะ Aminoglycosides, Tetracyclines, Amphotericin B, Cephalosporin, Sulfonamide โลหะหนัก เช่น ตะกั่วปรอท สารหนู ทองแดง แคดเมียม ทองลิเทียม พิษต่างๆ เช่น เห็ดพิษ แมลงกัดต่อย สมุนไพรที่เป็นพิษ พิษจากงู ยาชา สารทึบแสง ยาแก้ปวด เช่น Salicylates, Acetaminophen, Phenacetin และ NSAID เป็นต้น

3) โรคที่เกิดจากไตเอง

3.1) นิ่ว ทำให้มีการเคลื่อนที่มาอุดตันไตในระบบทางเดินปัสสาวะและมีการทำลายเนื้อไต

3.2) การอักเสบที่กรวยไต ทำให้มีการตอบสนองต่อการอักเสบ ทำให้เม็ดเลือดขาวเพิ่มขึ้น กระบวนการอักเสบทำให้เกิดการบวมของเนื้อเยื่อ เมื่อการอักเสบได้รับการรักษาก็จะทำให้เกิด Fibrosis ทำให้มีการดุดกลับ และการขับสิ่งต่างๆ เปลี่ยนแปลงไป ทำให้การทำหน้าที่ของไตลดลง

3.3) ภาวะไตบวมน้ำ ทำให้มีการขยายของกรวยไต และ Calices ทำให้มีการอุดตันของปัสสาวะ การสะสมของน้ำปัสสาวะ ทำให้เกิดแรงดันในกรวยไตเพิ่มขึ้น และเป็นสาเหตุให้กรวยไตถูกทำลาย

3.4) มะเร็งของไต เนื้องอกที่โตขึ้นอย่างรวดเร็วทำให้เกิดการอุดตันของระบบทางเดินปัสสาวะและทำให้เกิดไตบวมน้ำตามมา

3.5) โรคที่เกิดการอักเสบของเนื้อเยื่อระหว่างหลอดฝอยไต (Tubulointerstitial Disease) ทำให้เกิดการอุดตันที่หลอดฝอยไตและหลอดเลือดจากไตถูกทำลาย ซึ่งสาเหตุมาจากการอักเสบของกรวยไตอย่างเรื้อรัง

3.6) การอักเสบของโกลเมอรูลัส (Glomerulonephritis) มีการทำลายโกลเมอรูลัสซึ่งสาเหตุมาจากความผิดปกติของระบบภูมิคุ้มกันของร่างกาย เมื่อโกลเมอรูลัสถูกทำลาย ผู้ป่วยจะมีอาการปัสสาวะเป็นเลือด และมีโปรตีนออกมากับปัสสาวะ ตรวจเลือดพบอัลบูมิน (Albumin) ต่ำมีอาการบวมทั่วร่างกาย มีความผิดปกติของระบบเรนิน แองจิโอเทนซิน อัลโดสเตอโรน (Renin-angiotensin-aldosterone System) ทำให้ความดันโลหิตสูง

2.2.3 การรักษาโรคไตเรื้อรัง

การรักษาโรคไตเรื้อรังขึ้นอยู่กับระยะการดำเนินของโรคและระดับความรุนแรงของโรคซึ่งแนวทางการรักษามีอยู่ 2 ประการคือการรักษาเพื่อชะลอการเสื่อมของไตและการบำบัดทดแทนไต

1) การรักษาแบบประคับประคอง (Conservative Treatment) เป็นการรักษาด้วยยา การจัดการกับอาหาร น้ำและการหลีกเลี่ยงปัจจัยกระตุ้นที่ทำให้การเสื่อมของไตมากขึ้น เพื่อชะลอการเสื่อมของไตและการคงไว้ ซึ่งการทำหน้าที่ของไตที่เหลืออยู่ ได้แก่

1.1) การควบคุมอาหารในผู้ป่วยไตเรื้อรังไม่ว่าจะอยู่ในระยะใดก็ตาม จำเป็นต้องมีการ

ควบคุมและจำกัดสารอาหารเนื่องจากไตเสื่อมหน้าที่ในการขับของเสียออกจากร่างกายโดยเฉพาะสาร
 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ยูเรีย จุดมุ่งหมายของการรักษานี้เพื่อป้องกันการสลายตัวของโปรตีน โดยการให้สารอาหารทดแทน และพลังงานที่เพียงพอต่อร่างกายดังนี้

1.1.1) การควบคุมอาหารโปรตีน จะช่วยชะลอการเสื่อมของไตโดยปริมาณโปรตีนที่ควรได้รับในแต่ละวันขึ้นอยู่กับการทำงานของไต ถ้าค่าครีเอตินินในเลือดน้อยกว่า 2.5 มิลลิกรัมต่อเดซิลิตร ควรได้รับอาหารโปรตีนต่ำคือประมาณ 0.6-0.75 กรัมต่อกิโลกรัมของน้ำหนักตัว ถ้าค่าครีเอตินินในเลือดมากกว่า 2.5 มิลลิกรัมต่อเดซิลิตร ควรได้รับอาหารโปรตีนต่ำมากคือประมาณ 0.4 กรัมต่อกิโลกรัมของน้ำหนักตัว รวมทั้งการรับประทานกรดอะมิโนที่จำเป็นหรือให้กรดคีโตน (Keto Acid) เสริมวันละประมาณ 10 กรัม ในระยะแรกของไต การจำกัดโปรตีนไม่ค่อยจำเป็นนัก เนื่องจากไตยังสามารถทำหน้าที่ได้ แต่ในระยะที่ 2, 3 และ 4 เป็นระยะที่ไตมีความสามารถในการขับของเสียลดลงหรือเสียไป จึงจำเป็นต้องมีการจำกัดจำนวนโปรตีนและควรรับประทานโปรตีนที่มีคุณภาพที่มีกรดอะมิโนที่จำเป็น ได้แก่ เนื้อสัตว์ (หมู ไก่ วัว เป็ด ปลา) นม และไข่ วันละประมาณ 0.6 กรัมต่อกิโลกรัมของน้ำหนักตัว

1.1.2) การควบคุมโซเดียม ขึ้นอยู่กับสมรรถภาพของไตในการขับโซเดียมได้มากน้อยเพียงใด เนื่องจากปริมาณโซเดียมมีความสัมพันธ์กับปริมาตรของสารอาหารที่อยู่ภายนอกเซลล์ ดังนั้นจึงควรจำกัดจำนวนสารเหลวที่อยู่ภายนอกเซลล์ให้พอเหมาะ เพราะถ้าหากขาดน้ำแม้เพียงเล็กน้อยก็จะก่อให้เกิดภาวะยูรีเมียได้เร็ว แต่ถ้าผู้ป่วยมีภาวะน้ำเกินก็จะเกิดภาวะความดันโลหิตสูงหัวใจวายได้ ถ้าผู้ป่วยอยู่ในระยะที่ไตสามารถขับปัสสาวะออกมาได้มาก ก็ไม่จำเป็นต้องจำกัดโซเดียม เนื่องจากระยะนี้ผู้ป่วยมีปัสสาวะมากอาจจะขาดน้ำได้ แต่ในระยะไตวายเรื้อรังระยะสุดท้าย ไตไม่สามารถขับปัสสาวะออกมาได้ ทำให้ร่างกายบวม จำเป็นต้องจำกัดโซเดียมเหลือวันละ 2-4 กรัม ผู้ป่วยควรหลีกเลี่ยงอาหารกระป๋อง ของหมักดอง อาหารทะเลอาหารที่ถูกปรุงแต่ง (หมูแดง หมูเค็ม หมูหยอง ไข่พะโล้) ซอส น้ำปลา ซีอิ๊ว เป็นต้น นอกจากนี้การควบคุมปริมาณของเกลือโซเดียมให้ต่ำกว่า 2 กรัมจะช่วย เสริมฤทธิ์ของยาลดความดันโลหิต ทำให้ควบคุมความดันโลหิตได้ดีขึ้น

1.1.3) การควบคุมปริมาณโคเลสเตอรอล ควรควบคุมปริมาณโคเลสเตอรอลในอาหารไม่ให้เกินวันละ 300 มิลลิกรัม โดยจำกัดอาหารที่มีโคเลสเตอรอลสูง ได้แก่ ไข่แดง สมอตับสัตว์ ไขมันสัตว์และกะทิ เป็นต้น เนื่องจากทำให้โคเลสเตอรอลในเลือดสูงขึ้น ซึ่งเป็นอันตรายต่อหัวใจ หลอดเลือดต่างๆ และอาจเร่งให้ไตเสื่อมเร็วขึ้น

1.1.4) การควบคุมโพแทสเซียม ในผู้ป่วยที่ไตยังขับโพแทสเซียมได้อย่างเพียงพอ และระดับโพแทสเซียมในเลือดไม่สูงมากเกินไป ผู้ป่วยไม่จำเป็นต้องจำกัดหรือควบคุมอาหารประเภทนี้ แต่เมื่ออัตราการกรองของไตต่ำมากและมีปัสสาวะออกน้อยกว่า 500 มิลลิลิตรต่อวัน จึงควรควบคุมอาหารที่มีโพแทสเซียมสูง เช่น ก๋วย ล้ม องุ่น และพืชจำพวกผัก

1.1.5) การควบคุมฟอสเฟต อาหารที่มีฟอสเฟตสูงทำให้ไตเสื่อมเร็วขึ้นและทำให้ผู้ป่วยมีระดับฮอร์โมนพาราไทรอยด์สูงขึ้น จะเป็นผลเสียในระยะยาว จึงควรหลีกเลี่ยงอาหารที่มีเอกสารนี้เป็ฟอสเฟตสูง เช่น เนื้อสัตว์ ไข่แดง นมสดและเมล็ดพืช โดยผู้ป่วยควรรับประทานฟอสเฟตน้อยกว่าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

600 มิลลิกรัมต่อวันและควรควบคุมฟอสเฟตในเลือดให้อยู่ในระดับปกติซึ่งต้องใช้ยาจับฟอสเฟต (Phosphate Binder)

1.2) การจำกัดน้ำดื่ม ผู้ป่วยไตเรื้อรังระยะต่างๆ มีการจำกัดน้ำไม่เท่ากัน ในผู้ป่วยที่ไตยังสามารถขับปัสสาวะออกมาได้มาก ก็สามารถดื่มน้ำได้ตามปกติ ไม่จำเป็นต้องจำกัดน้ำดื่ม โดยทั่วไปควรดื่มน้ำประมาณ 2 ลิตรต่อวัน แต่ถ้าผู้ป่วยมีความรุนแรงของโรคมมากขึ้น โดยมีค่าครีเอตินินในเลือดอยู่ในช่วง 8-10 มิลลิกรัมต่อเดซิลิตร หรือมีอาการบวมตามตัว แล้วปัสสาวะออกลดลง ต้องจำกัดปริมาณน้ำดื่ม โดยถือตามปริมาณปัสสาวะที่ออกใน 24 ชั่วโมง ของผู้ป่วยเป็นเกณฑ์หรือโดยการชั่งน้ำหนักทุกวันในการประมาณการปรับปริมาณน้ำดื่ม คือถ้าน้ำหนักตัวเพิ่มขึ้นควรลดปริมาณน้ำดื่ม

1.3) การใช้ยาเพื่อชะลอการเสื่อมของไต ยาที่ใช้ได้แก่

1.3.1) ยาจับฟอสเฟต (Phosphate Binder) ใช้เพื่อควบคุมฟอสเฟตในเลือดให้อยู่ในระดับปกติเพื่อช่วยชะลอการเสื่อมของไตร่วมกับการควบคุมอาหารที่มีฟอสเฟต ยาในกลุ่มนี้ได้แก่ เกลือแคลเซียม เกลือแมกนีเซียม เกลืออลูมิเนียม โดยออกฤทธิ์จับฟอสเฟต ผลข้างเคียงคืออาจเกิดภาวะแคลเซียมในเลือดสูงเนื่องจากถ้าใส่ชุดแคลเซียมเพิ่มขึ้น จึงควรรับประทานยาพร้อมอาหาร ในกรณีที่ฟอสเฟตในเลือดสูงมากควรเลือกใช้เกลืออลูมิเนียมแทนในช่วงแรกก่อน

1.3.2) ยาลดความดันโลหิต ยาที่ใช้ในการลดความดันโลหิตคือ Angiotensin Converting Enzyme Inhibitors (ACE-inhibitor) หรือ Angiotensin Receptor Blocker ผู้ป่วยควรควบคุมความดันโลหิตให้เท่ากับหรือต่ำกว่า 130/80 มิลลิเมตรปรอท

1.3.3) ยาขับปัสสาวะออกฤทธิ์มีการขับเกลือแร่และน้ำที่คั่งในภาวะไตวาย ส่งผลให้ลดอาการบวมและลดความดันโลหิต ยาขับปัสสาวะที่ใช้ได้แก่ ยาในกลุ่ม Loop Diuretic เช่น Lasix

1.3.4) ยารักษาภาวะซีดหรือโลหิตจาง ยาที่ใช้ได้แก่ ฮอร์โมนอีริโธรโปอีติน เช่น รีคอร์มอน อีเพ็กซ์ ฮีแมกซ์ ซึ่งก่อนเริ่มให้ตรวจหาปริมาณธาตุเหล็กในร่างกาย หากมีภาวะขาดธาตุเหล็กด้วย ก็ให้ธาตุเหล็กทดแทนทางปาก เช่น เฟอร์รัสซัลเฟต เฟอร์โรสเลต

1.3.5) การให้ต่างทดแทน ภาวะความเป็นกรดมีผลเสียต่อร่างกายหลายอย่าง เช่น ทำให้กระดูกพรุน มีการสลายตัวของโปรตีนจากกล้ามเนื้อมากขึ้น ส่งผลให้เกิดระดับของเสียในเลือดเพิ่มสูงขึ้น จึงให้ต่างทดแทน ได้แก่ โซเดียม

1.3.6) วิตามินรวมและกรดโฟลิกอาจช่วยลดการเกิดภาวะหลอดเลือดแดงแข็งได้ นอกจากนี้ผู้ป่วยอาจมีภาวะขาดธาตุสังกะสี ซึ่งมีความสำคัญต่อระบบภูมิคุ้มกันและความอยากอาหารของผู้ป่วย

2) การรักษาโดยการบำบัดทดแทนไต (Renal Replacement Therapy) ในการบำบัดทดแทนไต ต้องพิจารณาอัตราการเสื่อมของโรค ภาวะแทรกซ้อนอื่นที่มีผลต่อการดำเนินของโรคและการตอบสนองต่อการรักษาโรคร่วมอื่นๆ ส่วนผู้ป่วยต้องพิจารณาสภาพจิตใจ เศรษฐกิจ ความร่วมมือในการรักษา ซึ่งผู้ป่วยบางรายอาจปฏิเสธ แม้มีข้อบ่งชี้ชัดเจนในการทำบำบัดทดแทนไต ขณะที่บาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายต้องเริ่มทำตั้งแต่เนิ่นๆ เพื่อหลีกเลี่ยงภาวะแทรกซ้อนของโรคไตเรื้อรัง ซึ่งการรักษาโดยการบำบัดทดแทนไตประกอบด้วย

2.1) การฟอกเลือดด้วยเครื่องไตเทียม (Hemodialysis) เป็นการกรองของเสียจากเลือดโดยให้เลือดจากหลอดเลือดผู้ป่วยผ่านเข้าหลอดฝอยซึ่งมีจำนวนมากในตัวกรองเลือด (Dialyser) ของเสียในเลือดจะซึมผ่านผนังหลอดเลือดฝอยออกมาในน้ำยาที่อาบอยู่รอบนอกของหลอดฝอยในตัวกรองเลือดโดยอาศัยหลักการที่ว่า สารละลายจะซึมผ่านผนังจากด้านที่มีความเข้มข้นสูงมาสู่ด้านที่มีความเข้มข้นต่ำ เมื่อเลือดผ่านตัวกรองเลือดแล้วก็จะไหลกลับเข้าร่างกายทางหลอดเลือดอีกทางหนึ่ง โดยทั่วไปใช้เวลาในการฟอกเลือดครั้งละ 4-5 ชั่วโมง สัปดาห์ละ 2-3 ครั้ง

2.2) การล้างไตทางหน้าท้องแบบถาวร (Continuous Ambulatory Peritoneal Dialysis) เป็นการขจัดของเสียออกทางช่องท้องอย่างต่อเนื่องโดยใช้ผนังเยื่อช่องท้องของผู้ป่วย (Peritoneal Membrane) เป็นตัวกรองน้ำ ของเสียและเกลือแร่ โดยปริมาณของเสียในเลือดจะซึมผ่านผนังหลอดเลือดฝอยที่มีจำนวนมากที่เยื่อช่องท้องออกมาในน้ำยาที่ใส่ไว้ในช่องท้องและถูกปล่อยออกมาจากช่องท้อง วิธีการนี้ผู้ป่วยสามารถทำตัวเอง ภายหลังจากได้รับการฝึกสอนจากแพทย์และพยาบาลเป็นเวลา 1-4 สัปดาห์ ขณะที่ทำการรักษาผู้ป่วยสามารถทำงานและปฏิบัติภารกิจได้ตามปกติ

2.3) การผ่าตัดปลูกถ่ายไต (Renal Transplantation) หรือที่เรียกกันทั่วไปว่าการเปลี่ยนไต เป็นวิธีการรักษาโรคไตเรื้อรังระยะสุดท้ายที่เป็นที่ยอมรับและได้ผลดีที่สุดในปัจจุบัน ผู้ป่วยจะมีคุณภาพชีวิตเหมือนปกติ เพียงแต่ต้องรับประทานยากดภูมิคุ้มกันภายหลังจากผ่าตัด แต่การรักษาด้วยวิธีนี้ก็ยังมีอัตราเสี่ยงสูงและผลของการรักษาขึ้นอยู่กับสภาพร่างกายของผู้ป่วยก่อนได้รับการปลูกถ่ายไตด้วย ดังนั้นการคัดเลือกผู้ป่วยที่เหมาะสมและการเตรียมผู้ป่วยให้พร้อมทั้งทางด้านร่างกายและจิตใจจึงมีความสำคัญเป็นอย่างยิ่ง

2.2.4 แนวทางการพยาบาลเพื่อชะลอการเสื่อมของไต

การดูแลเพื่อชะลอการเสื่อมของไตเป็นสิ่งสำคัญ เพื่อที่จะไม่ทำให้ผู้ป่วยเข้าสู่ไตเรื้อรังระยะสุดท้ายเร็วขึ้น และต้องเข้ารับการบำบัดทดแทนไต โดยมีแนวทางการพยาบาลเพื่อชะลอการเสื่อมของไตดังนี้

1) การปรับเปลี่ยนวิถีชีวิต (Lifestyle Modification) ได้แก่

1.1) การควบคุมน้ำหนัก ไม่ให้มากหรือน้อยกว่าน้ำหนักตามเกณฑ์ ผู้ที่มีน้ำหนักตัวเกินควรลดปริมาณอาหารต่อมื้อให้น้อยลง อาจเหลือเพียงครึ่งหนึ่งของที่เคยรับประทาน ห้ามลดน้ำหนักด้วยการลดอาหารอย่างรวดเร็วหรืออดอาหาร หลีกเลี่ยงอาหารที่ทำให้น้ำหนักตัวเพิ่มขึ้นอย่างรวดเร็ว เช่น น้ำตาล ขนมหวาน ไอศกรีม และขนมเค้ก

1.2) กรณีไม่สบายแม้เพียงเล็กน้อย เช่น อาการเจ็บคอ ควรรีบดูแลให้หาย เนื่องจากไตอาจ

อีกเสบตามมาได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3) ควรสร้างอุปนิสัยในการขับถ่ายที่ดี เพราะอุปนิสัยในการขับถ่ายที่ไม่ถูกต้องจะนำไปสู่การสูญเสียสมรรถภาพในการควบคุมระบบประสาท กระเพาะปัสสาวะ ทำให้การเก็บและขับถ่ายปัสสาวะผิดปกติไปด้วย อุปนิสัยที่ดีในการขับถ่ายสามารถฝึกโดยการดื่มน้ำสะอาดวันละประมาณ 6-8 แก้ว (1500-2000 ซีซี/วัน) ใช้เวลาเต็มที่เมื่อถ่ายปัสสาวะ ไม่รีบร้อน และแน่ใจว่าปัสสาวะออกหมด

1.4) ระวังอย่าให้ท้องเสีย เพราะถ้าสูญเสียน้ำเป็นจำนวนมากและรวดเร็วจะเกิดไตวายเฉียบพลัน ต้องทำการล้างไตฉุกเฉิน และพบว่าไตมักจะไม่ฟื้นกลับมา

1.5) ปรับอุปนิสัยการรับประทานอาหาร ให้รับประทานอาหารที่มีประโยชน์ต่อสุขภาพและในปริมาณที่เหมาะสม ควรลดการบริโภคสารอาหารกลุ่มไขมันอิ่มตัว (Saturated Fat) ได้แก่ น้ำมันหมู กะทิ นม เนย เนื้อมะพร้าว น้ำมันมะพร้าว น้ำมันปาล์ม เนยเทียม และควรลดการบริโภคเกลือลง เนื่องจากผู้ป่วยโรคไตเรื้อรังมีจำนวนหน่วยไต (Nephron) น้อยกว่าคนปกติ ทำให้เกิดการคั่งของเกลือได้ง่าย

1.6) ออกกำลังกายอย่างเหมาะสม เป็นประจำอย่างสม่ำเสมอ การออกกำลังกายควรเป็นแบบ Aerobic Exercise เช่น โยคะ โทเก็ท และเดินแอโรบิค เป็นต้น

1.7) เลิกสูบบุหรี่ เนื่องจากการสูบบุหรี่มีความสัมพันธ์โดยตรงต่อโรคหลอดเลือดและหัวใจ ในผู้ป่วยเบาหวานที่สูบบุหรี่จะเสี่ยงภาวะไตเสื่อมมากกว่าผู้ป่วยเบาหวานที่ไม่สูบบุหรี่ถึง 1-6 เท่า

1.8) เลิกดื่มแอลกอฮอล์ เนื่องจากแอลกอฮอล์ให้พลังงานสูง ซึ่งจะทำให้น้ำหนักตัวเพิ่มขึ้น ความดันโลหิตสูงขึ้น ไม่เป็นผลดีต่อหัวใจ แอลกอฮอล์ยังทำให้ปัสสาวะมากขึ้น ทำให้สมองควบคุมกระเพาะปัสสาวะได้ลำบาก

1.9) งดรับประทานยาที่ไม่จำเป็น เช่น ยาสมุนไพร ยาจีน ยาลูกกลอน เนื่องจากอาจมีส่วนผสมของสเตอรอยด์ (Steroid) และเกลือผสมอยู่มาก ทำให้เกิดการสะสมของน้ำและเกลือแร่ เพิ่มระดับความดันโลหิตและเพิ่มระดับน้ำตาลในเลือด

1.10) หลีกเลี่ยงสารโลหะหนัก เช่น ตะกั่ว ทองแดง ปรอท ซึ่งมักใช้ในอุตสาหกรรม แบตเตอรี่รถยนต์ เหล็กกล้า สีทาบ้าน น้ำมันเบนซิน เมื่อตกค้างในร่างกายจะกลายเป็นพิษทำลายไต

1.11) พักผ่อนให้เพียงพอ ควรพักผ่อนอย่างน้อย 6-8 ชั่วโมงต่อวัน หลีกเลี่ยงความเครียด ทำจิตใจให้เบิกบานอยู่เสมอ

2) ควบคุมความดันโลหิต ผู้ป่วยโรคไตเรื้อรังควรควบคุมความดันโลหิตให้อยู่ในระดับต่ำกว่า 130/80 มิลลิเมตรปรอท ผู้ป่วยควรได้รับยาลดความดันโลหิตเพื่อชะลอการเสื่อมของไต ได้แก่ ยากลุ่ม Angiotensin-Converting Enzyme Inhibitor (ACE-I) หรือยา Angiotensin Receptor Blockers (ARB)

3) ควบคุมระดับน้ำตาลในเลือด ผู้ป่วยไตและเป็นเบาหวานควรควบคุมระดับน้ำตาลในเลือดให้อยู่ในเกณฑ์อย่างเคร่งครัด โดยสมาคมโรคเบาหวานแห่งสหรัฐอเมริกากำหนดให้ควบคุมระดับน้ำตาลในเลือดเท่ากับ 90-130 มิลลิกรัมต่อเดซิลิตร ส่วนสมาคมต่อมไร้ท่อแห่งสหรัฐอเมริกาแนะนำไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้คุมให้น้อยกว่า 110 มิลลิกรัมต่อเดซิลิตร เพื่อป้องกันภาวะแทรกซ้อนทางหลอดเลือดแดงเล็กและใหญ่ และระดับ HbA1C สมาคมโรคเบาหวานแห่งสหรัฐอเมริกาแนะนำให้ควบคุมให้น้อยกว่า 7 เปอร์เซ็นต์ และสมาคมต่อมไร้ท่อแห่งสหรัฐอเมริกาแนะนำให้ควบคุมให้น้อยกว่าหรือเท่ากับ 6.5 เปอร์เซ็นต์

4) จำกัดอาหารโปรตีน ผู้ป่วยโรคไตเรื้อรังระยะที่ 3 เป็นต้นไป ควรได้รับอาหารโปรตีนไม่เกิน 0.8 กรัมต่อน้ำหนักตัว 1 กิโลกรัม เพื่อลดภาวะเลือดเป็นกรด แต่ต้องระมัดระวังการควบคุมมากเกินไป อาจทำให้เกิดภาวะขาดสารอาหารได้

5) ไขมันในเลือด เนื่องจากผู้ป่วยโรคไตเรื้อรังมีความเสี่ยงต่อการเกิดโรคหัวใจและหลอดเลือด ระดับไขมันในเลือดที่ต้องควบคุมคือ LDL-cholesterol ควรต่ำกว่า 100 มิลลิกรัมต่อเดซิลิตร

6) หลีกเลี่ยงยาที่มีผลต่อไต ได้แก่ ยาบรรเทาอาการปวดและยาลดการอักเสบในกลุ่ม NSAID ยาปฏิชีวนะกลุ่ม Aminoglycoside และยาแก้ปวด Paracetamol นอกจากนี้มีผลต่อดับ ถ้าได้รับเกินขนาด อาจทำให้มีการเสื่อมของไตเร็วกว่าปกติ

7) ควบคุมระดับความสมดุลของกรดต่างในร่างกาย ผู้ป่วยอาจได้รับยา Sodamint ผลไบคาร์บอเนตในเลือดไม่ควรต่ำกว่า 22 มิลลิอีควิวเลนซ์ต่อลิตร

8) ควบคุมอาหารเค็ม ความต้องการเกลือโซเดียมในผู้ป่วยโรคไตเรื้อรังจะมากน้อยเพียงใดขึ้นอยู่กับความสามารถของไตที่เหลืออยู่ ผู้ป่วยควรรับประทานอาหารที่ใช้เกลืออนอมอาหารงดการใช้เครื่องปรุงรสที่มีเกลือผสม งดใช้สารเคมีที่มีโซเดียมสูง หลีกเลี่ยงอาหารธรรมชาติที่มีโซเดียมสูง เช่น อาหารทะเล หอย กุ้ง และปลาหมึก

9) ฝ้าระวังระดับสารฟอสฟอรัสในเลือดให้ต่ำกว่า 5 มิลลิกรัมต่อเดซิลิตร และรักษาระดับแคลเซียมและฟอสเฟตให้ต่ำกว่า 55 มิลลิกรัมต่อเดซิลิตร

10) ฝ้าระวังค่าฮีโมโกลบินในเลือดให้อยู่ระหว่าง 11-12 กรัมต่อเดซิลิตร

11) ฝ้าระวังการดำเนินของโรคจากผลเลือดทางห้องปฏิบัติการ ได้แก่ ค่าของเสียในเลือดปียูเร็น (Blood Urea Nitrogen) และค่าครีเอตินิน (Creatinine) การส่งต่อผู้ป่วยไปพบแพทย์เพื่อประเมินผู้ป่วย เมื่อผู้ป่วยมีค่า GFR ต่ำกว่า 30 มิลลิลิตรต่อนาทีต่อพื้นที่ผิวกาย 1.73 ตารางเมตรหรือเร็วกว่านั้น ถ้าผู้ป่วยมีปัจจัยเสี่ยงที่ทำให้ไตเสื่อมอย่างรวดเร็ว

2.3 วิธีการแบ่งประเภทข้อมูลของวิธีการจำแนกกลุ่ม (Classification)

การแบ่งประเภทข้อมูลคือกระบวนการสร้างตัวแบบเพื่อจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนด เป็นการสร้างตัวแบบการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างของข้อมูลที่ได้กำหนดไว้ล่วงหน้า และสามารถพยากรณ์กลุ่มของข้อมูลที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ ตัวแบบที่ได้อาจอยู่ในรูปแบบต้นไม้ตัดสินใจ (Decision Tree) หรือโครงข่ายประสาทเทียม (Artificial Neural Network)

ในการจัดหมวดหมู่จำเป็นต้องแบ่งข้อมูลออกเป็น 2 ส่วน ส่วนแรกคือข้อมูลสำหรับการเรียนรู้ (Training Data) เพื่อให้ข้อมูลเรียนรู้และสร้างตัวแบบ (Model Construction) และส่วนที่สองคือข้อมูลสำหรับการทดสอบ (Testing Data) เพื่อประเมินความถูกต้องของตัวแบบ (Model Evaluation) อีกทั้งใช้ชุดข้อมูลที่ไม่เคยเห็นมาก่อน (Unseen Data) เพื่อทำการกำหนดกลุ่มให้กับข้อมูลใหม่ที่ได้มาหรือทำนายค่าออกมาตามที่ต้องการ เช่น การจัดหมวดหมู่ของผู้ยื่นบัตรเครดิต (Credits) เป็นระดับต่ำ ระดับกลางและระดับสูง ของความเสี่ยงที่จะได้รับหรือการอนุมัติบุคคลเข้ารับทำงานในลักษณะงานต่างๆ (พยุบ พาณิชยกุล, 2548)

2.4 วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor)

วิธีความใกล้เคียงกันมากที่สุดเป็นวิธีการที่ได้รับความนิยมในการใช้งานอย่างมาก สาเหตุเนื่องจากเป็นวิธีการที่ง่ายและมีประสิทธิภาพซึ่งสามารถนำไปประยุกต์ใช้กับงานได้อย่างหลากหลาย เช่น งานทางด้าน การจำแนกกลุ่ม (Classification) รวมถึงงานทางด้าน การแทนที่ข้อมูลที่สูญหาย (Missing Values Imputation) ซึ่งมีขั้นตอนการดำเนินการดังนี้ (Troyanskaya, O. et al., 2001)

- 1) กำหนดค่า k เพื่อใช้พิจารณาสมาชิกที่อยู่ใกล้เคียงกันมากที่สุด เช่น $k = 3$ คือ จะพิจารณาเฉพาะข้อมูล 3 ตัวแรกที่อยู่ใกล้กับจุดที่ต้องการจะทำนาย
- 2) คำนวณหาระยะห่างระหว่างข้อมูลตัวอย่างที่สนใจกับข้อมูลอื่น ๆ ทุกตัวด้วยวิธีระยะห่างยูคลิดีเนียน (Euclidian distance) จากสมการ (2-1)

$$\text{dist}(x_i, x_j) = \sqrt{\sum_{k=1}^n (x_{i,k} - x_{j,k})^2} \quad (2-1)$$

โดยที่ $\text{dist}(x_i, x_j)$ คือ ระยะห่างระหว่างตัวอย่าง x_i กับตัวอย่าง x_j

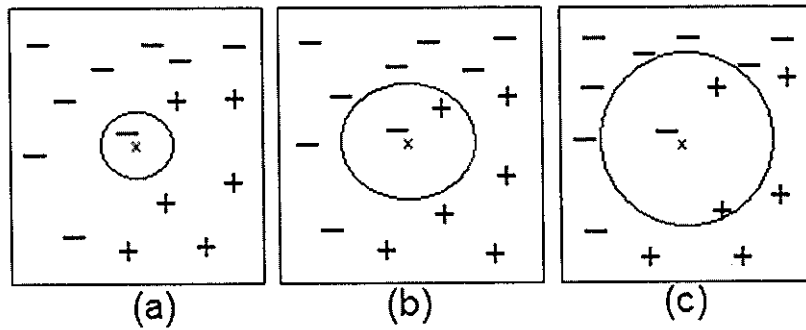
n คือ จำนวนคุณสมบัติทั้งหมดของตัวอย่าง

$x_{i,k}$ คือ คุณสมบัติที่ k ของตัวอย่าง x_i

$x_{j,k}$ คือ คุณสมบัติที่ k ของตัวอย่าง x_j

- 3) เลือกค่าข้อมูลที่มีค่าระยะห่างน้อยที่สุด k ตัว เพื่อนำมาพิจารณาหาค่าตอบ

ดังรูปที่ 2-1



- (a) ความใกล้เคียงกันมากที่สุดโดยพิจารณาจากข้อมูล 1 ตัว
 (b) ความใกล้เคียงกันมากที่สุดโดยพิจารณาจากข้อมูล 2 ตัว
 (c) ความใกล้เคียงกันมากที่สุดโดยพิจารณาจากข้อมูล 3 ตัว

รูปที่ 2-1 ตัวอย่างของความใกล้เคียงกันมากที่สุด

2.5 วิธีต้นไม้ตัดสินใจ (Decision Tree) (รุจิรา ธรรมสมบัติ, 2554)

วิธีต้นไม้ตัดสินใจเป็นตัวแทนทางคณิตศาสตร์เพื่อหาทางเลือกที่ดีที่สุด โดยการนำข้อมูลมาสร้างตัวแบบการพยากรณ์ในรูปแบบของโครงสร้างต้นไม้ ซึ่งมีการเรียนรู้ข้อมูลแบบมีผู้สอน (Supervised Learning) สามารถสร้างตัวแบบการจัดกลุ่ม (Clustering) ได้จากกลุ่มตัวอย่างของข้อมูลฝึกหัด (Training Data) ได้โดยอัตโนมัติและสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดกลุ่มได้อีกด้วย

โดยปกติมักประกอบด้วยกฎในรูปแบบ “ถ้า เงื่อนไข แล้ว ผลลัพธ์” เช่น

ถ้ารายได้สูงและยังไม่ได้แต่งงาน แล้วฐานะยากจน

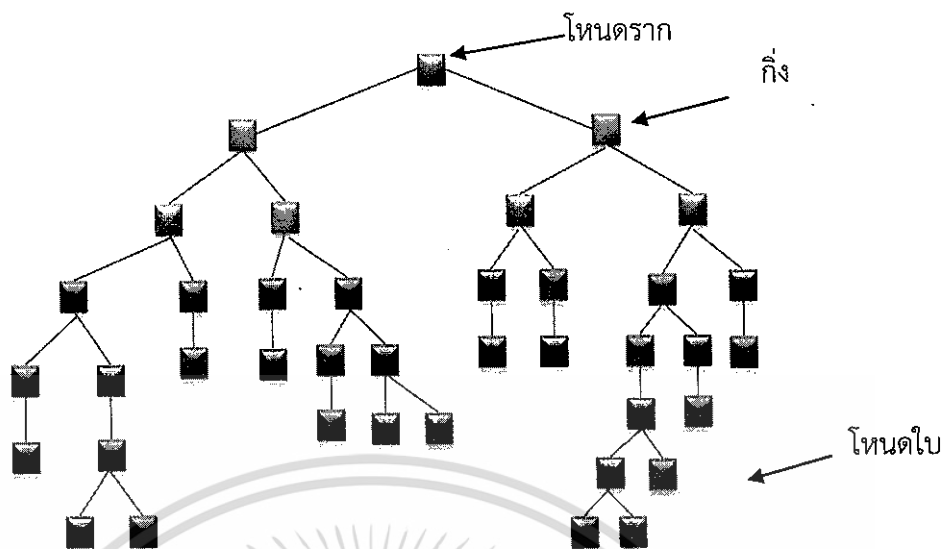
“If Income = High and Married = No THEN Risk = Poor”

ถ้ารายได้สูงและแต่งงาน แล้วฐานะร่ำรวย

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

2.5.1 ส่วนประกอบของต้นไม้ตัดสินใจ

- 1) โหนด (Node) คือ คุณสมบัติต่างๆ เป็นจุดที่แยกข้อมูลว่าจะให้ไปในทิศทางใด ซึ่งโหนดที่อยู่สูงสุดเรียกว่า โหนดราก (Root Node)
- 2) กิ่ง (Branch) คือ คุณสมบัติของโหนดที่แตกออกมา โดยจำนวนของกิ่งจะเท่ากับ คุณสมบัติของโหนด
- 3) ใบ (Leaf) คือ กลุ่มของผลลัพธ์ในการแยกแยะข้อมูล ซึ่งโหนดที่อยู่ล่างสุดเรียกว่า โหนดใบ (Leaf Node) โดยสามารถแสดงส่วนประกอบของต้นไม้ตัดสินใจ ดังรูปที่ 2-2



รูปที่ 2-2 ส่วนประกอบของต้นไม้ตัดสินใจ

2.5.2 การสร้างต้นไม้ตัดสินใจ

หลักการพื้นฐานของการสร้างต้นไม้ตัดสินใจเป็นการสร้างในลักษณะจากบนลงล่าง (Top-Down) คือเริ่มจากการสร้างรากของต้นไม้ก่อนแล้วจึงแตกกิ่งไปจนถึงใบ โดยแสดงขั้นตอนการสร้างต้นไม้ตัดสินใจได้ดังนี้

- 1) ต้นไม้เริ่มต้นโดยมีโหนดเพียงโหนดเดียวแสดงถึงชุดข้อมูลฝึก (Training Data)
- 2) ถ้าข้อมูลทั้งหมดอยู่ในกลุ่มเดียวกันแล้ว ให้โหนดนั้นเป็นใบและตั้งชื่อแยกตามกลุ่มของข้อมูลนั้น
- 3) ถ้าโหนดมีข้อมูลหลายกลุ่มปะปนอยู่ จะต้องวัดค่าผลกำไร (Gain) ของแต่ละคุณลักษณะ (Attribute) เพื่อที่จะใช้เป็นเกณฑ์ในการคัดเลือกคุณลักษณะที่มีความสามารถในการแบ่งแยกข้อมูลออกเป็นกลุ่มต่าง ๆ ได้ดีที่สุด โดยคุณลักษณะที่มีผลกำไรมากที่สุดจะถูกเลือกให้เป็นตัวทดสอบหรือคุณลักษณะที่ใช้ในการตัดสินใจ โดยแสดงในรูปของโหนดบนต้นไม้
- 4) กิ่งของต้นไม้ ถูกสร้างขึ้นจากค่าต่าง ๆ ที่เป็นไปได้ของโหนดทดสอบ และข้อมูลจะถูกแบ่งออกตามกิ่งต่าง ๆ ที่สร้างขึ้น
- 5) ทำการวนซ้ำเพื่อหาคุณลักษณะที่มีผลกำไรมากที่สุด สำหรับข้อมูลที่ถูกแบ่งแยกออกมาในแต่ละกิ่งเพื่อนำคุณลักษณะนี้มาสร้างเป็นโหนดตัดสินใจต่อไป โดยที่คุณลักษณะที่ถูกเลือกมาเป็นโหนดแล้ว จะไม่ถูกเลือกมาอีกสำหรับโหนดในระดับต่อ ๆ ไป
- 6) ทำการวนซ้ำเพื่อแบ่งข้อมูลและแตกกิ่งของต้นไม้ไปเรื่อย ๆ โดยการวนซ้ำจะสิ้นสุดก็ต่อเมื่อเงื่อนไขข้อใดข้อหนึ่งข้างบนนี้เป็นจริง

2.5.3 การคำนวณค่าผลกำไรสารสนเทศ (Information Gain)

ต้นไม้ตัดสินใจเป็นโครงสร้างที่ใช้แสดงกฎที่ได้จากวิธีการจำแนกกลุ่มข้อมูล โดยต้นไม้ตัดสินใจจะมีลักษณะคล้ายโครงสร้างต้นไม้ โดยที่แต่ละโหนดแสดงคุณลักษณะ (Attribute) ในการสร้างต้นไม้ตัดสินใจ ปัญหาสำคัญที่ต้องพิจารณาคือควรจะต้องตัดสินใจเลือกคุณลักษณะใดมาทำหน้าที่เป็นโหนดราก ในแต่ละขั้นตอนของการสร้างต้นไม้และต้นไม้ย่อย (Subtree) ของต้นไม้ตัดสินใจเกณฑ์ที่ใช้ช่วยประกอบการเลือกคุณลักษณะคือการคำนวณค่ามาตรฐานผลกำไร (Gain Criterion) ซึ่งเป็นค่าที่บ่งบอกว่าคุณลักษณะนั้นสามารถจำแนกกลุ่มของข้อมูลได้ดีเพียงใด โดยทดลองเลือกแต่ละคุณลักษณะที่เป็นไปได้จากชุดข้อมูลมาทำหน้าที่เป็นโหนดราก ถ้าคุณลักษณะใดให้ค่าผลกำไรสูงที่สุด แสดงว่าคุณลักษณะนั้นสามารถจำแนกกลุ่มของข้อมูลได้ดีที่สุด การใช้ค่าผลกำไรสารสนเทศจะช่วยลดจำนวนครั้งของการทดสอบในการแยกแยะข้อมูล อีกทั้งยังรับประกันว่าต้นไม้ตัดสินใจที่ได้ไม่มีความซับซ้อนมากเกินไป ซึ่งค่าผลกำไรสารสนเทศนั้นสามารถคำนวณได้จากสมการที่ 2-2

$$I(S_1, S_2, \dots, S_n) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S} \quad (2-2)$$

เมื่อ S แทนเซตของข้อมูลซึ่งประกอบด้วยข้อมูล S ระเบียบ (Record)
 n แทนจำนวนกลุ่มทั้งหมดที่ต่างกันของข้อมูลชุดนั้น
 S_i แทนจำนวนข้อมูลที่เป็นสมาชิกของ S และอยู่ในกลุ่ม C_i
 C_i แทนกลุ่มในลำดับที่ i โดยที่ i มีค่าระหว่าง 1 ถึง n

ค่าเอ็นโทรปี (Entropy) ของคุณลักษณะ A ซึ่งมีค่าของคุณลักษณะเป็น $(a_1, a_2, a_3, \dots, a_v)$ หาได้ในสมการที่ 2-3

$$E(A) = \sum_{j=1}^v \frac{S_{1j} + \dots + S_{nj}}{S} (S_{1j}, S_{2j}, \dots, S_{nj}) \quad (2-3)$$

S_{ij} แทนจำนวนข้อมูลที่เป็นสมาชิกของ S ในกลุ่ม C_i จากการแบ่งข้อมูลด้วยค่าที่เป็นไปได้ของคุณลักษณะ A

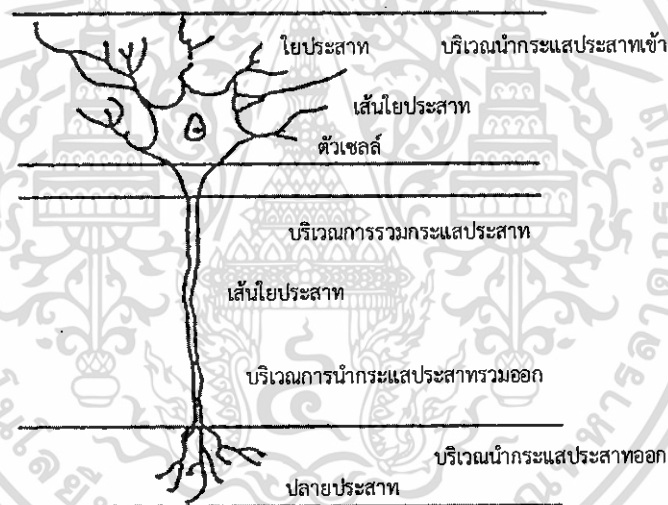
ดังนั้นจะสามารถพิจารณาค่ามาตรฐานผลกำไรได้ดังสมการที่ 2-4

$$\text{Gain (A)} = I(S_1, S_2, \dots, S_n) - E(A) \quad (2-4)$$

2.6 วิธีโครงข่ายประสาทเทียม (Artificial Neural Network) (ศุภกฤต ชูธงชัย, 2546)

โครงข่ายประสาทเทียมเป็นศาสตร์ที่จำลองแบบความสามารถของมนุษย์ด้านการเรียนรู้จดจำและจำแนกสิ่งต่างๆ ซึ่งใช้สมองเป็นส่วนสำคัญ ในการประมวลระบบของโครงข่ายประสาทเทียมนั้นจะเลียนแบบการทำงานของระบบสมอง คือมีการส่งผ่านข้อมูลระหว่างกันโดยมีการเชื่อมต่อของเซลล์ประสาท (Neuron) กันเป็นโครงข่ายร่างแหจำนวนมากและมีการประมวลผลในลักษณะขนาน (Parallel Processing) สาเหตุหลักที่โครงข่ายประสาทเทียมเป็นที่นิยมมากขึ้นเนื่องจากมีความยืดหยุ่นในการทำงานสูงและสามารถปรับตัวเองให้ทำงานในสภาพที่เปลี่ยนแปลง อีกทั้งไม่จำเป็นต้องทราบตัวแบบทางคณิตศาสตร์ (Mathematical Model) ที่แน่นอนของกระบวนการ เพียงแต่ใช้ชุดข้อมูลที่ประกอบด้วยข้อมูลนำเข้า (Input Data) และข้อมูลเป้าหมาย (Target Data) ของกระบวนการในจำนวนที่มากพอมาใช้ในการสอน (Training) โครงข่ายประสาทเทียม

2.6.1 ความรู้พื้นฐานของระบบประสาท (Neural System Knowledge)



รูปที่ 2-3 โครงข่ายของเซลล์ประสาท

ภายในสมองมนุษย์ประกอบด้วยหน่วยประมวลผลขนาดเล็กที่เรียกว่า เซลล์ประสาท (Neurons) ซึ่งจะมีประมาณ 10 เซลล์ ในเซลล์ประสาทแต่ละเซลล์ดังแสดงในรูปที่ 2-3 ประกอบด้วย ใยประสาท (Dendrites) ตัวเซลล์ (Cell Body) และเส้นใยประสาท (Axon) ซึ่งแบ่งออกเป็น 4 บริเวณด้วยกัน คือ

1) บริเวณนำกระแสประสาทเข้า (Input Region) เป็นบริเวณที่จะมีการนำกระแสประสาท (Nerve impulse) จากเซลล์ประสาทอื่นเข้ามาภายในตัวเซลล์โดยผ่านทางใยประสาท ซึ่งมีลักษณะแตกเป็นกิ่งก้านคล้ายต้นไม้และมีจำนวนตั้งแต่ 1 ใยขึ้นไป

2) บริเวณการรวมกระแสประสาท (Integration Region) เป็นบริเวณที่มีการรวมกระแสประสาทก่อนที่จะเข้าสู่บริเวณการนำกระแสประสาทรวมออกจากเซลล์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) บริเวณการนำกระแสประสาทรวมออกจากเซลล์ (Conduction Region) เป็นบริเวณที่จะนำกระแสประสาทรวมออกจากเซลล์ โดยใช้เส้นใยประสาทเป็นทางผ่าน ซึ่งมีเพียง 1 เส้นใยต่อเซลล์เท่านั้น

4) บริเวณนำกระแสประสาทออก (Output Region) เป็นบริเวณส่วนปลายของเส้นใยประสาทที่มีการแตกแขนง ใช้ในการถ่ายทอดกระแสประสาทข้ามเซลล์ไปยังเซลล์ประสาทอื่นโดยผ่านทางใยประสาทของเซลล์ประสาทรุ่น

2.6.2 การเรียนรู้ของโครงข่ายประสาทเทียม (Artificial Neural Network Learning)

การเรียนรู้ของโครงข่ายประสาทเทียมจะมีประสิทธิภาพเพียงใดนั้นขึ้นอยู่กับค่าถ่วงน้ำหนัก (Weight) ของโครงข่ายที่ทำการออกแบบ ซึ่งการฝึกหัด (Training) โครงข่ายคือการหาค่าถ่วงน้ำหนักที่เหมาะสมให้กับโครงข่ายนั้นๆ โดยทั่วไปสามารถจำแนกวิธีการเรียนรู้ของโครงข่ายประสาทเทียมได้เป็น 2 ประเภท คือ การเรียนรู้แบบมีผู้สอนและการเรียนรู้แบบไม่มีผู้สอน

1) การเรียนรู้แบบมีผู้สอน (Supervised Learning)

การเรียนรู้แบบมีผู้สอนจะกำหนดข้อมูลฝึกหัด (Training Data) ให้กับโครงข่าย ซึ่งกลุ่มนี้ประกอบด้วยข้อมูลนำเข้า (Input Data) และข้อมูลเป้าหมาย (Target Data) ที่ต้องการ จากนั้นโครงข่ายจะทำการคำนวณค่าถ่วงน้ำหนักที่เหมาะสมให้กับข้อมูลฝึกหัด โดยคำตอบที่ได้จากโครงข่ายจะถูกคำนวณค่าความผิดพลาด (Error Value) ว่ามีความห่างจากคำตอบที่ต้องการของข้อมูลนำเข้าในชุดเดียวกันมากน้อยเพียงใด ถ้ายังมีความผิดพลาดสูงอยู่ การฝึกหัดจะดำเนินต่อจนกว่าค่าความผิดพลาดจะลดลงต่ำกว่าค่าที่ยอมรับได้ (Accepted Level) จึงจะหยุดฝึกหัด สุดท้ายค่าถ่วงน้ำหนักที่ได้จะเป็นเหมือนฟังก์ชันที่ใช้ในการแปลงข้อมูล

2) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบไม่มีผู้สอนจะอาศัยชุดข้อมูลนำเข้าเพียงอย่างเดียวในการฝึกหัดโครงข่าย โดยไม่มีข้อมูลเป้าหมาย แต่จะใช้ข้อมูลนำออก (Output Data) จากโครงข่ายแทน เมื่อป้อนข้อมูลเข้าสู่โครงข่าย โครงข่ายจะคำนวณค่าความสัมพันธ์ที่มีอยู่ภายในกลุ่มข้อมูลนำเข้า โดยอาศัยค่าถ่วงน้ำหนักเป็นตัวแยกความแตกต่างของข้อมูลนำเข้าและนำไปเก็บไว้ในโหนดข้อมูลนำออกของโครงข่าย ซึ่งมีวัตถุประสงค์เพื่อใช้ในการจำแนกชุดข้อมูล (Clustering)

งานวิจัยขั้นนี้ใช้การเรียนรู้แบบมีผู้สอน เนื่องจากโครงข่ายสามารถระบุกลุ่มของข้อมูลเป้าหมายได้อย่างแน่นอน ทำให้สะดวกในการออกแบบมากกว่าการเรียนรู้แบบไม่มีผู้สอน

2.6.3 การเชื่อมโยงของโครงข่ายประสาทเทียม (Artificial Neural Network Linking)

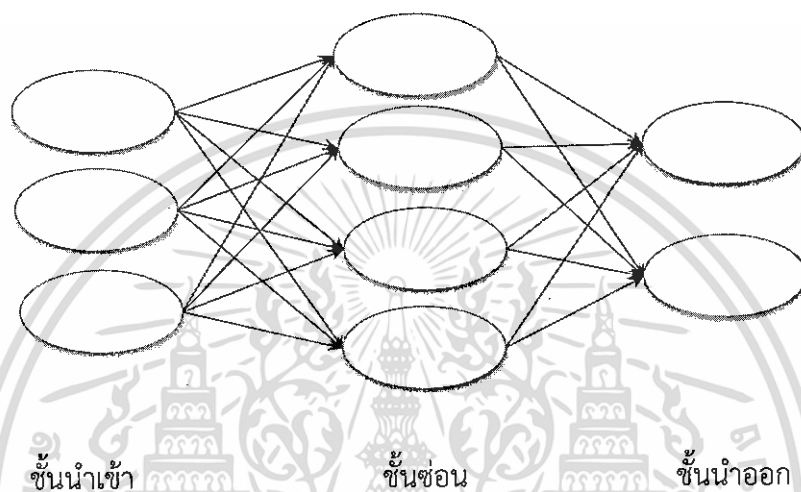
เพื่อให้โครงข่ายประสาทเทียมสามารถเรียนรู้ได้อย่างมีประสิทธิภาพ จำเป็นต้องมีการเชื่อมโยงกันระหว่างเซลล์ประสาท โดยทั่วไปสามารถแบ่งการเชื่อมโยงของโครงข่ายได้ 2 ลักษณะคือ

1) โครงข่ายแบบส่งสัญญาณไปข้างหน้า (Feed-forward Network)

เป็นโครงข่ายที่การประมวลผลจะอาศัยชุดข้อมูลปัจจุบันและส่งค่าที่ประมวลผลได้ไปยังชั้นถัดๆ ไป

กล่าวคือ โครงข่ายชนิดนี้จะประกอบด้วยชั้นต่างๆ โดยชั้นแรกจะเป็นชั้นนำเข้า (Input Layer) และ
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับนักเรียนเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

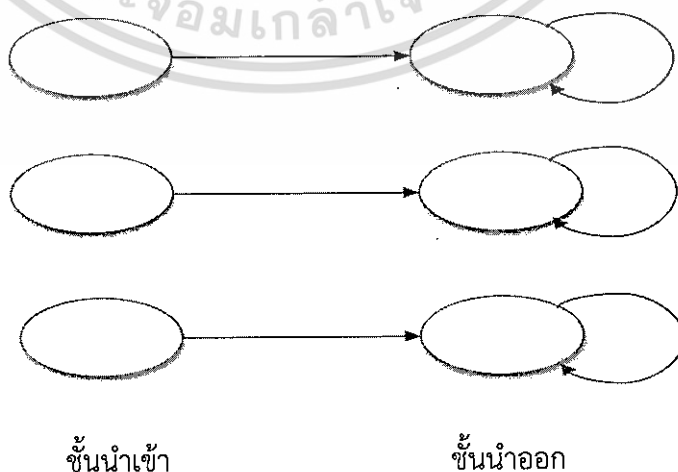
ชั้นสุดท้ายเป็นชั้นนำออก (Output Layer) ส่วนระหว่างชั้นนำเข้ากับชั้นนำออกอาจจะมีหรือไม่มีชั้นซ่อน (Hidden Layers) อยู่ภายใน ซึ่งขึ้นกับกฎการเรียนรู้ (Learning Rule) ที่ใช้ในการสอนโครงข่าย เช่น ถ้าเป็นโครงข่ายเพอร์เซปตรอนแบบหลายชั้น (Multi-layer Perceptron) จะมี ชั้นซ่อนอยู่ระหว่างชั้นนำเข้ากับชั้นนำออก ซึ่งอาจมีมากกว่าหนึ่งชั้นได้ การเชื่อมต่อระหว่างชั้นของโครงข่ายแบบส่งสัญญาณไปข้างหน้าจะมีค่าถ่วงน้ำหนัก (Weight) เป็นตัวเชื่อมและสัญญาณนำเข้าที่เข้ามาจะถูกส่งไปตามทิศทางของลูกศรจนถึงชั้นนำออกโดยไม่มีการป้อนกลับ สามารถแสดงตัวอย่างได้ดังรูปที่ 2-4



รูปที่ 2-4 ลักษณะโครงข่ายประสาทเทียมแบบส่งสัญญาณไปข้างหน้า

2) โครงข่ายแบบมีการย้อนกลับ (Feed-back Network)

โครงข่ายชนิดนี้มีชื่อเรียกอีกชื่อหนึ่งว่า โครงข่ายหันกลับ (Recurrent Network) เป็นโครงข่ายที่จะอาศัยทั้งข้อมูลในปัจจุบันและข้อมูลที่มีการถ่วงเวลามาใช้ในการประมวลผลของโครงข่ายประสาทเทียม สามารถแสดงตัวอย่างโครงข่ายที่มีการย้อนกลับได้ดังรูปที่ 2-5

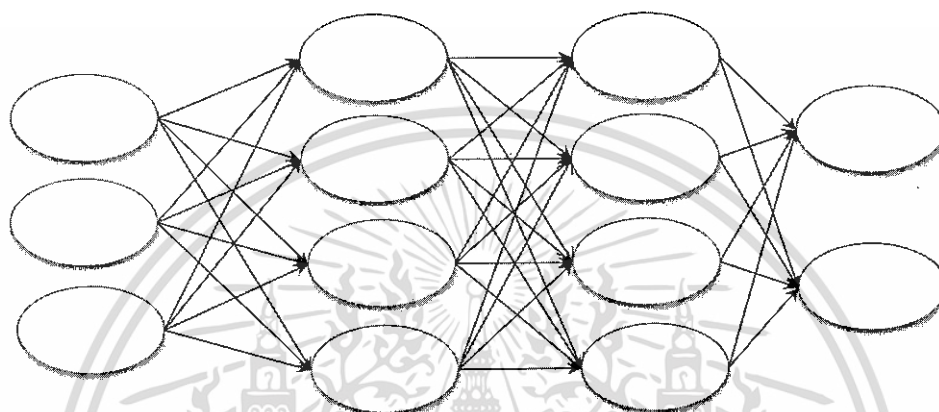


รูปที่ 2-5 ลักษณะโครงข่ายประสาทเทียมแบบมีการย้อนกลับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.4 การแพร่แบบย้อนกลับ (Back-propagation)

การแพร่แบบย้อนกลับเป็นขั้นตอนที่ใช้สอนโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น (Multi-layer Perceptron) ซึ่งแบบจำลองโครงข่ายประสาทเทียมมีการเชื่อมโยงกันเป็นโครงข่ายแบบเป็นชั้นๆ โครงข่ายชนิดนี้มีการเชื่อมโยงกัน 3 ชั้น ประกอบด้วยชั้นนำเข้า (Input Layer) ถัดมาเป็นชั้นซ่อน (Hidden Layer) และชั้นสุดท้ายคือชั้นนำออก (Output Layer) รูปที่ 2-6 แสดงโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้นที่มีชั้นซ่อน 2 ชั้น



ชั้นนำเข้า ชั้นซ่อนที่ 1 ชั้นซ่อนที่ 2 ชั้นนำออก
(Input Layer) (First Hidden Layer) (Second Hidden Layer) (Output Layer)

รูปที่ 2-6 โครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น

ที่มาของชื่อการแพร่แบบย้อนกลับนั้นมาจากจุดที่ว่า วิธีการปรับค่าถ่วงน้ำหนักเพื่อให้ได้ค่าที่เหมาะสมนั้นจะใช้วิธีสอนว่าค่าเป้าหมาย (Target) ของแต่ละข้อมูลนำเข้านั้นคืออะไร และใช้ค่าความผิดพลาด (Error) ของข้อมูลนำออกมาใช้เป็นตัวชี้้นำในการปรับค่าถ่วงน้ำหนัก ดังนั้นการแพร่แบบย้อนกลับจึงเป็นกระบวนการเรียนรู้แบบมีผู้สอน แต่ปัญหาที่เกิดขึ้นคือไม่มีค่าเป้าหมายของสัญญาณที่ออกมาจากแต่ละเซลล์ประสาทในชั้นซ่อน ดังนั้นจึงต้องอาศัยการแพร่ความผิดพลาดจากชั้นนำออกกลับมายังชั้นซ่อนนั่นเอง

2.6.5 ปัจจัยที่ส่งผลต่อการเรียนรู้การแพร่แบบย้อนกลับ

1) การกำหนดค่าเริ่มต้นของค่าถ่วงน้ำหนัก

ก่อนที่จะทำการสอนโครงข่ายประสาทเทียมแบบหลายชั้น จำเป็นต้องกำหนดค่าเริ่มต้นให้กับค่าถ่วงน้ำหนักที่เชื่อมโยงระหว่างชั้นทุกชั้น โดยค่านี้จะเป็นเลขจำนวนจริงที่มีค่าน้อยๆ ที่ได้มาจากการสุ่มค่าเริ่มต้น (Randomness)

2) การกำหนดเกณฑ์การหยุดฝึกหัด

เกณฑ์ในการหยุดฝึกหัดนั้นขึ้นกับผู้ที่ทำกรออกแบบโครงข่ายประสาทเทียมว่าต้องการที่จะให้โครงข่ายประสาทเทียมมีความแม่นยำเพียงใด โดยทั่วไปนิยมใช้ค่าดัชนีที่ชี้ถึงค่าความผิดพลาดของ

เอกสารนี้เบ้ระบบได้ ในงานวิจัยนี้ใช้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error, MSE) ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) อัตราการเรียนรู้ (Learning Rate, η)

อัตราการเรียนรู้เป็นค่าสัมประสิทธิ์ที่แสดงถึงการเรียนรู้ของโครงข่าย โดยทั่วไปค่าที่เหมาะสมจะอยู่ในช่วง 0.05 ถึง 0.5 ถ้าอัตราการเรียนรู้มีค่าสูง แสดงว่ากำหนดให้โครงข่ายมีการเปลี่ยนแปลงค่าถ่วงน้ำหนักที่มาก ในทางตรงกันข้ามถ้ามีอัตราการเรียนรู้ต่ำ แสดงว่ากำหนดให้โครงข่ายมีการเปลี่ยนแปลงค่าถ่วงน้ำหนักที่น้อย ซึ่งจำเป็นต้องใช้เวลาในการเรียนรู้ที่มากขึ้น แต่จะมีข้อดีคือโครงข่ายจะมีเสถียรภาพขณะที่ทำการเรียนรู้

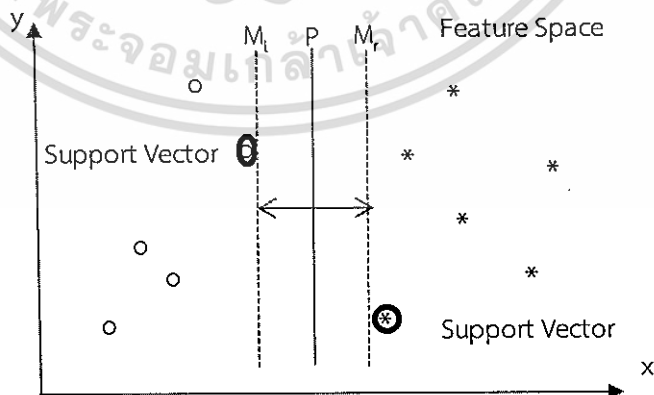
4) ค่าคงที่โมเมนตัม (Momentum Constant, α)

ค่าคงที่โมเมนตัมเป็นค่าสัมประสิทธิ์ที่ช่วยหน่วงไม่ให้เกิดการเปลี่ยนแปลงค่าถ่วงน้ำหนักนั้นมากเกินไป เป็นการเพิ่มเสถียรภาพให้กับโครงข่ายได้อีกทางหนึ่ง ซึ่งค่าโมเมนตัมที่เหมาะสมจะมีค่าเข้าใกล้ 1.0 และควรที่จะกำหนดให้สอดคล้องกับอัตราการเรียนรู้ด้วย เช่น ถ้าอัตราการเรียนรู้สูงก็ควรที่จะมีค่าโมเมนตัมที่ต่ำ ทำให้เกิดการเปลี่ยนแปลงค่าถ่วงน้ำหนักนั้นไม่มากจนเกินไป

2.7 วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) (อานนท์ นามสนิท, 2549)

2.7.1 แนวความคิดของซัพพอร์ตเวกเตอร์แมชชีน

ซัพพอร์ตเวกเตอร์แมชชีนเป็นสมการที่ใช้ในการจำแนกค่าคุณลักษณะของ 2 กลุ่มที่วางตัวอยู่ในพื้นที่คุณลักษณะ (Feature Space) ออกจากกันโดยจะสร้างเส้นแบ่ง (Plane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่ง 2 กลุ่มออกจากกันนั้น เส้นตรงใดที่เป็นเส้นที่ดีที่สุด โดยเส้นตรงนี้จะเพิ่มเส้นขอบ (Margin) ออกไปทั้งสองข้าง โดยเส้นขอบที่เพิ่มนั้นจะขนานกับเส้นเดิมเสมอ เส้นขอบที่เพิ่มขึ้นมานี้จะขยายออกไปจนกว่าจะสัมผัสกับค่าของกลุ่มตัวอย่างที่ใกล้ที่สุด ดังรูปที่ 2-7



รูปที่ 2-7 การขยายตัวของเส้นขอบ

จากรูปที่ 2-7 เส้น M_L และ M_r คือเส้นขอบที่ขยายออกไปด้านซ้ายและขวาตามลำดับ และ P คือเส้นแบ่งข้อมูลทั้ง 2 กลุ่ม เมื่อเส้น M_L และ M_r ขยายออกจนไปสัมผัสค่าข้อมูลที่ใกล้ที่สุด ซึ่งข้อมูลที่อยู่บนเส้นขอบของทั้งสองฝั่งนั้นเรียกว่า ซัพพอร์ตเวกเตอร์ (Support Vector) จะวัดค่าระยะความห่างของเส้นขอบ โดยเส้น P จะเปลี่ยนความชันไปเรื่อย ๆ เพื่อที่จะหาความกว้างสูงสุดของเส้นขอบ กระบวนการโดยรวมของซัพพอร์ตเวกเตอร์แมชชีนนั้น เป็นการหาค่าความชันของเส้น P ที่มีขนาดของเส้นขอบสูงสุดนั่นเอง

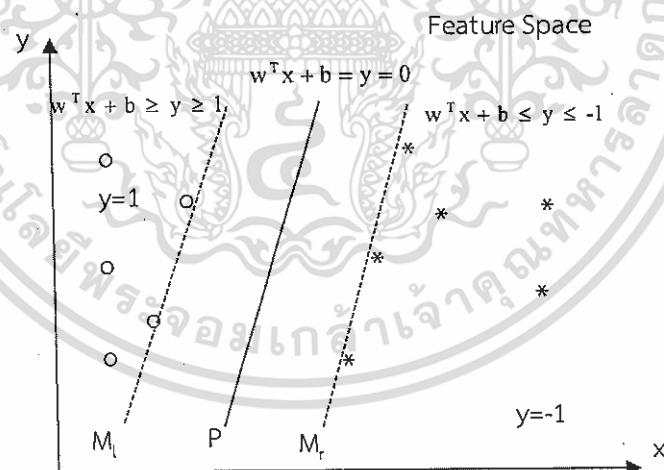
2.7.2 สมการพื้นฐานของซัพพอร์ตเวกเตอร์แมชชีน

ถ้านำแนวคิดของซัพพอร์ตเวกเตอร์แมชชีนที่กล่าวไปแล้วในข้อ 2.7.1 มาเขียนเป็นสมการเพื่อใช้ในการแก้ปัญหา โดยข้อมูลที่นำมาวางลงในพื้นที่คุณลักษณะนั้นเป็นกลุ่มข้อมูลที่อยู่ในรูปของเวกเตอร์

$$x = ((x_1, y_1), \dots, (x_i, y_i)) \quad (2-5)$$

เมื่อ x คือ ชุดค่าคุณลักษณะ

ค่าคุณลักษณะที่วางตัวอยู่ในพื้นที่คุณลักษณะ จะถูกแบ่งด้วยเส้นตรงดังรูปที่ 2-8 และเมื่อนำเส้นตรงมาแทนค่าด้วยสมการเส้นตรง $y = mx + b$ โดยมีการกำหนดกลุ่มของข้อมูลทั้งสองฝั่งเป็นเพียงสองค่าที่ซึ่งแทนด้วยค่า y เพื่อให้ข้อมูลที่อยู่ในกลุ่มเดียวกันที่มาจากหลายค่ากลายเป็นค่าเดียว ดังสมการในรูปที่ 2-8



รูปที่ 2-8 เส้นขอบและเส้นแบ่งเมื่อแทนด้วยสมการเส้นตรง

จากรูปที่ 2-8 เส้นตรง M_L แทนด้วยสมการ $w^T x + b \geq y \geq 1$ ซึ่งข้อมูล y ที่มากกว่า 1 ก็จะถูกกำหนดค่าใหม่โดยให้ y เท่ากับ 1 และ พจน์ w ก็คือค่าความชัน เช่นเดียวกันกับเส้นตรง M_r ที่ค่าของ y จะถูกกำหนดค่าใหม่เมื่อ y น้อยกว่า -1 เป็น -1 ดังนั้นสมการที่เกิดขึ้นใหม่จากสมการเส้นขอบ 2-6 และ 2-7 สามารถกำหนดได้ดังสมการที่ 2-8

$$\text{เมื่อ } w^T x + b \geq y \text{ กำหนด } y = 1 \quad (2-6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ทางการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$w^T x + b \leq y \quad \text{กำหนด } y = -1 \quad (2-7)$$

$$y(w^T x + b) - 1 \geq 0 \quad (2-8)$$

โดย y คือ ค่ากลุ่มข้อมูล (1,-1)

w คือ ค่าความชัน

x คือ ค่าคุณลักษณะ

b คือ ค่าคงที่ (ค่าตัดแกน y)

2.7.3 ค่าความกว้างเส้นขอบ (Margin)

การคำนวณความกว้างของเส้นขอบต้องทำการคำนวณพจน์ w ให้อยู่ในรูปปกติมาตรฐาน (Normalization) โดยคำนวณจากสมการที่ 2-6 และ 2-7 เมื่อแทนค่า y ลงไปแล้ว

$$\begin{aligned} w^T x^+ + b &= 1 \\ w^T x^- + b &= -1 \\ w^T (x^+ - x^-) &= 2 \end{aligned} \quad (2-9)$$

$$\text{จากสูตร} \quad M = \left(\frac{w}{\|w\|} \right)^T (x^+ - x^-) \quad (2-10)$$

นำสมการที่ 2-9 แทนลงในสมการที่ 2-10

$$M = \frac{2}{\|w\|} \quad (2-11)$$

โดยที่ M คือ ความกว้างของเส้นขอบ (Margin)

2.7.4 การแก้ปัญหา Dual Problem

หลังจากที่ได้สมการที่ 2-8 และ 2-11 ของการหาเส้นแบ่งและค่าความกว้างตามลำดับแล้ว จึงทำการแก้สมการนั้นต่อด้วย Lagrangian Dual Problem เพื่อหาค่าในพจน์ของ w ในสมการที่ 2-8

$$L(w, b, \alpha) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^N \alpha_i [y_i ((w_i \cdot w_i) + b) - 1] \quad (2-12)$$

เมื่อ $\alpha_i \geq 0; i = 1, \dots, N$

สมการที่ 2-12 ถูกนำมาหาอนุพันธ์ (Differential)

$$\frac{\partial L(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^N y_i \alpha_i x_i = 0 \quad (2-13)$$

$$\frac{\partial L(w, b, \alpha)}{\partial b} = \sum_{i=1}^N y_i \alpha_i = 0 \quad (2-14)$$

หลังจากได้สมการที่ผ่านแก้ปัญหาด้วย Dual Problem แล้วค่า w จะลดรูป และทำได้จาก

$$w = \sum_{i=1}^N y_i \alpha_i x_i \quad (2-15)$$

เมื่อ α คือ สัมประสิทธิ์คงที่

เมื่อนำค่า w ไปใส่ในสมการที่ 2-8 ซึ่งเป็นสมการในการหาเส้นแบ่ง จะได้
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$y_i \left(\sum_{i=1}^N \alpha_i y_i x_i^T x + b \right) - 1 \geq 0 \quad (2-16)$$

2.7.5 เคอร์เนล (Kernel)

ในโลกความเป็นจริงนั้นข้อมูล 2 กลุ่มไม่ได้วางตัวในพื้นที่คุณลักษณะ และไม่สามารถแบ่งได้โดยเส้นตรง แต่ข้อมูลอาจจะจับกลุ่มกันในตำแหน่งต่างๆ ดังนั้นจึงเป็นปัญหาทำให้ไม่สามารถที่จะใช้สมการซัพพอร์ตเวกเตอร์แมชชีนแบบเชิงเส้นได้ ดังนั้นจะต้องมีเครื่องมือมาช่วยให้ข้อมูลเหล่านั้นเรียงตัวใหม่ในพื้นที่ เรียกว่า พื้นที่หลายมิติ (Higher Dimensional Space)



รูปที่ 2-9 รูปแบบการวางตัวที่ไม่สามารถแบ่งด้วยเส้นตรงได้

ในเคอร์เนลนั้นคือการคูณกันของชุดเวกเตอร์ของ x ใด ๆ

$$K(x_i, x_j) = x_i^T x_j \quad (2-17)$$

เคอร์เนลที่นิยมใช้มีอยู่ 3 ชนิดด้วยกัน คือ

- 1) โพลีโนเมียล (Polynomial)

$$K(x_i, x_j) = (x_i^T x_j + 1)^d \quad (2-18)$$

เมื่อ d คือค่าเลขยกกำลัง

- 2) ฟังก์ชันเบสิสรัศมี (Radial Basis Function : RBF)

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (2-19)$$

เมื่อ σ คือ ค่าพารามิเตอร์

- 3) ซิกมอยด์ (Sigmoid)

$$K(x_i, x_j) = \tanh(kx_i^T x_j + \mu) \quad (2-20)$$

เมื่อ k, μ คือค่าพารามิเตอร์

ดังนั้นจากสมการของเคอร์เนลนั้นสามารถที่จะแทนลงไปในพื้นที่ของ $x_i^T x_j$ ในสมการที่

$$y_i \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) - 1 \geq 0 \quad (2-21)$$

สมการที่ 2-21 เป็นสมการที่ใช้ในขั้นตอนที่จะเรียนรู้ว่าจะวางตำแหน่งเส้นแบ่งไว้ที่ตำแหน่งใดโดยทำงานร่วมกับเคอร์เนล เพื่อแปลงให้ข้อมูลที่ยากต่อการแบ่งแบบเชิงเส้นสามารถแบ่งได้เมื่อทำให้เป็นข้อมูลแบบหลายมิติ (Higher Dimension) ดังนั้นจึงมีอีกสมการหนึ่งที่ใช้ค่า w และ b เดิม มาจัดตำแหน่งของข้อมูลเพื่อให้ทราบว่าข้อมูลนั้นเป็นกลุ่มใด กำหนดได้ดังสมการที่ 2-22

$$f(x) = \text{sgn} \left(\sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \right) \quad (2-22)$$

เมื่อ $f(x)$ คือค่า y หาในรูปของ x

2.8 วิธีฐานกฎ (Rule-Based) (Witten I. H. et al., 2011)

วิธีฐานกฎเป็นวิธีหนึ่งที่นิยมใช้เช่นเดียวกับวิธีต้นไม้ตัดสินใจ ข้อกำหนดหรือเงื่อนไข (Antecedent or Precondition) ของวิธีฐานกฎเป็นการทดสอบคล้ายกับการทดสอบของวิธีต้นไม้ตัดสินใจ แต่ผลของการทดสอบหรือผลลัพธ์ (Consequent or Conclusion) ที่ได้ นั้น จะให้คำตอบ (Class) ที่ใช้กับตัวอย่างที่อยู่ภายใต้กฎนั้น หรือบางครั้งก็จะให้ค่าการแจกแจงความน่าจะเป็นของคำตอบต่างๆ กฎบางสูตรมีข้อกำหนดหรือเงื่อนไขเป็นการแสดงทางตรรกะทั่วไปมากกว่าที่จะเป็นการเชื่อมอย่างง่าย (Simple Conjunction) ถ้ากฎหนึ่งถูกนำไปใช้ คำตอบ (หรือการแจกแจงความน่าจะเป็น) ที่กำหนดในข้อสรุปจะถูกนำไปใช้กับตัวอย่างเช่นกัน อย่างไรก็ตาม จะเกิดข้อขัดแย้งขึ้นเมื่อกฎหลายกฎมีข้อสรุปแตกต่างกัน

2.8.1 กฎการจำแนก (Classification Rules)

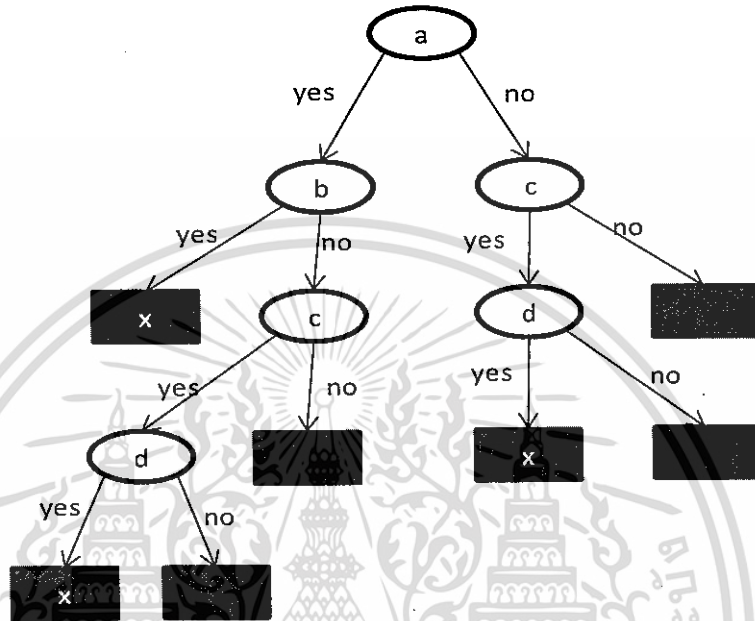
เป็นเรื่องง่ายที่จะอ่านชุดของกฎการจำแนกโดยตรงจากวิธีต้นไม้ตัดสินใจ กฎหนึ่งจะถูกสร้างขึ้นสำหรับแต่ละใบ (Leaf) ข้อกำหนดหรือเงื่อนไข (Antecedent) ของกฎประกอบด้วยเงื่อนไขของทุกโหนดบนเส้นทางจากราก (Root) ไปยังใบ (Leaf) และผลลัพธ์ (Consequence) ของกฎคือคำตอบที่ถูกกำหนดโดยใบ ขั้นตอนการสร้างกฎเป็นไปตามลำดับซึ่งมีการดำเนินงานที่ไม่เกี่ยวข้องกัน แต่อย่างไรก็ตาม โดยทั่วไปกฎที่อ่านได้จากวิธีต้นไม้ตัดสินใจโดยตรงมีความซับซ้อนมากเกินไป และกฎที่ได้จากต้นไม้มีมักมีการปรับปรุงโดยตัดการทดสอบที่ซ้ำซ้อนออกไป

เนื่องจากวิธีต้นไม้ตัดสินใจไม่สามารถแสดงได้โดยง่ายในกฎที่แตกต่างกันในข้อมูลชุดหนึ่งๆ การเปลี่ยนแปลงชุดของกฎเป็นต้นไม้ค่อนข้างจะไม่ตรงไปตรงมา การอธิบายที่ดีของการเกิดขึ้นนี้เมื่อกฎมีโครงสร้างที่เหมือนกัน แต่คุณลักษณะ (Attributes) ต่างกัน เช่น

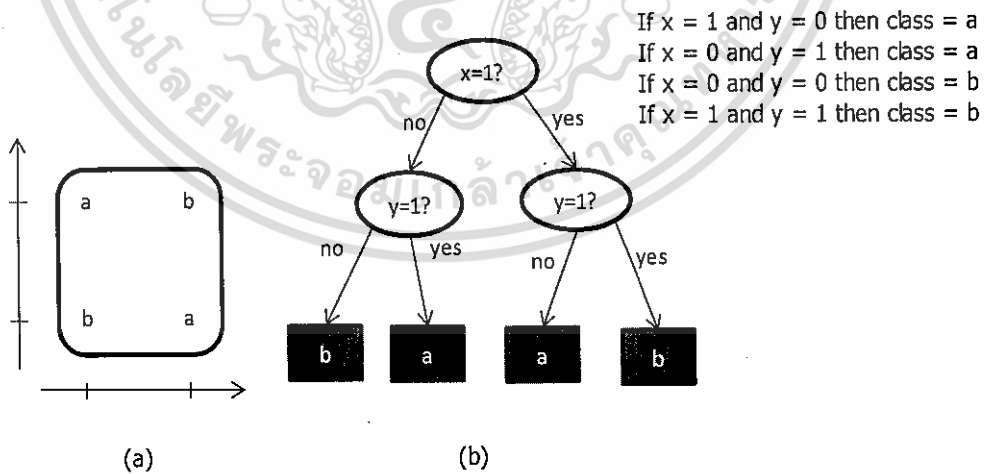
ถ้า a และ b แล้ว x

ถ้า c และ d แล้ว x

นั่นคือ จะเลือกการทดสอบเดียวสำหรับโหนดราก (Root Node) ตัวอย่างเช่น ถ้า a ถูกเลือก กฎข้อที่ 2 ทำซ้ำ 2 ครั้งในต้นไม้ ดังรูปที่ 2-10 เรียกว่า ปัญหาต้นไม้ย่อยที่มีการทำซ้ำ (Replicated Subtree Problem)



รูปที่ 2-10 รูปปัญหาต้นไม้ย่อยที่มีการทำซ้ำ



รูปที่ 2-11 ปัญหาต้นไม้ย่อยที่มีการทำซ้ำในการมองแบบตัวอย่างคู่

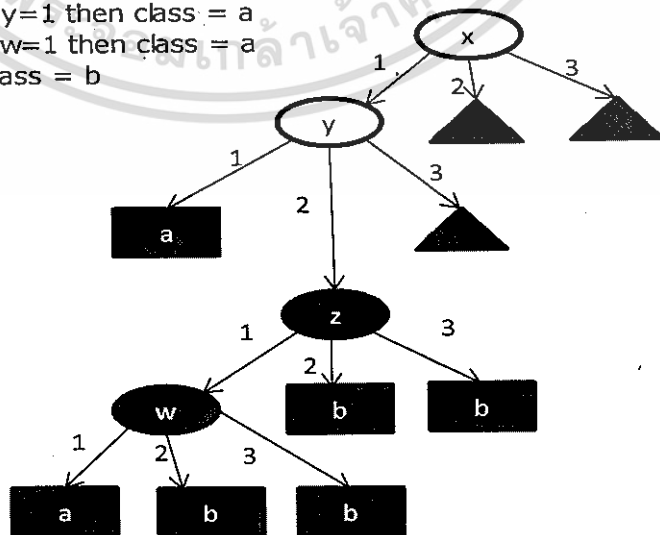
ปัญหาต้นไม้ย่อยที่มีการทำซ้ำมีความสำคัญในการมองแบบตัวอย่างคู่ แผนภาพด้านบนของรูป 2-11 แสดงให้เห็นฟังก์ชันที่มีผลลัพธ์เป็น a ถ้า $x=1$ หรือ $y=1$ แต่ไม่ใช่เป็น 1 ทั้งคู่ หรือไม่ใช่เป็น 0 ทั้งคู่ เพื่อทำสิ่งนี้ลงไปต้นไม้ เราจะต้องแยกคุณลักษณะ (Attributes) แรกที่นำไปสู่โครงสร้างไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เช่นเดียวกับที่แสดงให้เห็นในจุดศูนย์กลาง (Center) ในทางตรงกันข้าม วิธีฐานกฎสามารถสะท้อนให้เห็นถึงปัญหาของความสมมาตรที่แท้จริงเมื่อเทียบกับคุณลักษณะต่างๆ ซึ่งแสดงให้เห็นทางด้านขวา

ในตัวอย่างนี้ กฎไม่ได้มีขนาดกระทัดรัดมากกว่าวิธีต้นไม้ตัดสินใจ ซึ่งในความเป็นจริงทั้ง 2 วิธีต่างเป็นเพียงแค่สิ่งที่ได้มาจากการอ่านกฎของวิธีต้นไม้ตัดสินใจอย่างชัดเจน แต่ในสถานการณ์อื่นๆ กฎมีขนาดกระทัดรัดมากกว่าวิธีต้นไม้ตัดสินใจ โดยเฉพาะอย่างยิ่งถ้าเป็นไปได้ที่จะมีกฎเริ่มต้นที่จะครอบคลุมกรณีต่างๆ ที่ไม่ได้ระบุไว้โดยกฎอื่น ยกตัวอย่างเช่น ภาพอิทธิพลของกฎ ดังรูปที่ 2-12 ซึ่งมี 4 คุณลักษณะคือ x , y , z และ w โดยแต่ละคุณลักษณะคือ 1, 2 หรือ 3 ดังแสดงในรูปต้นไม้ตัดสินใจด้านล่าง รูปสามเหลี่ยมเล็กๆ เทาทั้งสามรูปด้านขวาบนควรจะเป็นต้นไม้ย่อย 3 ระดับ ดังแสดงเป็นสีเทา เป็นตัวอย่างของปัญหาต้นไม้ย่อยที่มีการทำซ้ำ รายละเอียดความซ้ำซ้อนของแนวคิดนี้ค่อนข้างง่าย

เหตุผลหนึ่งที่ทำให้วิธีฐานกฎเป็นที่นิยมคือ แต่ละกฎจะแสดงให้เห็นถึงความรู้ที่เป็นอิสระต่อกัน กฎใหม่สามารถเพิ่มไปยังกฎที่มีอยู่เดิมได้โดยปราศจากการเข้าไปเกี่ยวข้องกับกฎอื่นๆ ที่มีอยู่แล้ว ในขณะที่เพิ่มโครงสร้างต้นไม้อาจจะต้องปรับปรุงต้นไม้ใหม่ทั้งต้น อย่างไรก็ตาม ความเป็นอิสระต่อกันนี้ จะไม่สนใจคำถามที่ว่ากฎชุดนี้จะดำเนินการอย่างไร เราได้อธิบายข้อเท็จจริงไปก่อนหน้านี้แล้วว่าถ้ากฎหมายถึงการตีความตามลำดับในรูปการตัดสินใจ (Decision List) บางอย่างและการออกนอกเนื้อหาอาจไม่ถูกต้อง ในทางตรงกันข้าม ถ้าลำดับของการตีความไม่มีความสำคัญ แล้วไม่มีความชัดเจนที่จะทำอะไร เมื่อกฎที่แตกต่างกันจะนำไปสู่ข้อสรุปที่แตกต่างกันสำหรับตัวอย่างเดียวกัน ในกรณีเช่นนี้จะไม่เกิดขึ้นกับกฎที่อ่านได้โดยตรงจากวิธีต้นไม้ตัดสินใจ เนื่องจากความซ้ำซ้อน (Redundancy) ถือเป็นส่วนหนึ่งของโครงสร้างของกฎที่จะป้องกันความคลุมเครือในการตีความหมาย แต่จะเกิดขึ้นเมื่อกฎถูกสร้างขึ้นโดยวิธีการอื่นๆ

If $x=1$ and $y=1$ then class = a
 If $z=1$ and $w=1$ then class = a
 otherwise class = b



รูปที่ 2-12 อิทธิพลของกฎ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

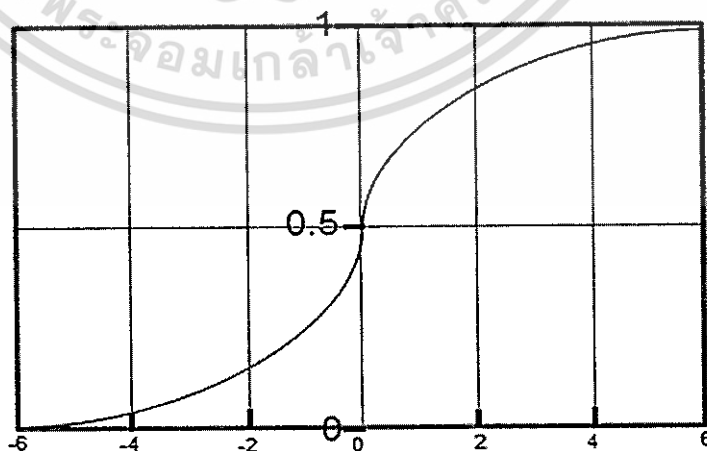
ถ้าชุดของกฎสามารถจำแนกได้หลายประเภทสำหรับตัวอย่างที่เฉพาะ วิธีการหนึ่งในการแก้ปัญหาคือไม่มีข้อสรุป วิธีการอีกวิธีหนึ่งคือให้นับความถี่ของแต่ละกฎที่ใช้ในข้อมูลฝึกหัด (Training Data) ซึ่งเป็นที่นิยมที่สุด กลวิธีเหล่านี้สามารถนำไปสู่ผลลัพธ์ที่แตกต่างกันอย่างสิ้นเชิง ปัญหาที่แตกต่างเกิดขึ้นเมื่อตัวอย่างหนึ่งเผชิญกับวิธีฐานกฎที่ไม่สามารถจำแนกประเภทได้เลย สิ่งนี้ จะไม่สามารถเกิดขึ้นกับวิธีต้นไม้ตัดสินใจหรือกฎที่อ่านได้โดยตรงจากวิธีต้นไม้ตัดสินใจ แต่สามารถเกิดขึ้นได้ง่ายกับของชุดกฎโดยทั่วไป วิธีการหนึ่งในการแก้ปัญหานี้คือการตัดสินใจที่จะไม่จำแนกประเภทตัวอย่างเช่นนี้ วิธีการอีกวิธีหนึ่งคือการเลือกคำตอบที่เกิดบ่อยที่สุดเป็นค่าเริ่มต้น (Default) ผลลัพธ์ที่แตกต่างกันอย่างสิ้นเชิงอาจจะมาจากกลวิธีเหล่านี้ กฎที่เฉพาะและชุดของกฎมีความง่ายเหมือนกัน แต่มีเพียงแค่ชุดของกฎเท่านั้นที่ไม่มีข้อมูลเพิ่มเติม นั่นคือไม่มีความชัดเจนในการตีความหมาย

2.9 วิธีการถดถอยลอจิสติก (Logistic Regression)

การถดถอยลอจิสติก (Logistic Regression) ใช้เพื่อหาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม โดยตัวแปรตามมีเพียงสองค่าคือ 0 และ 1 หากตัวแปรอิสระมีค่าน้อย ค่าของตัวแปรตาม จะมีค่าเท่ากับ 0 และหากค่าตัวแปรอิสระมีค่ามาก ค่าของตัวแปรตามจะมีค่าเท่ากับ 1 เส้นโค้งที่จะสามารถแทนความสัมพันธ์นี้คือเส้นโค้งที่ได้จากฟังก์ชัน

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2-23)$$

ฟังก์ชันนี้มีชื่อว่าฟังก์ชันลอจิสติก (Logistic Function) ตัวแปรตามที่ได้จากฟังก์ชันนี้จะมีค่าอยู่ ระหว่าง 0 และ 1



รูปที่ 2-13 ฟังก์ชันลอจิสติก

จากรูปที่ 2-13 ตัวแบบการวิเคราะห์การถดถอยลอจิสติก ในกรณีที่มีตัวแปรอิสระเพียงตัวเดียว
ตัวแบบการวิเคราะห์การถดถอยลอจิสติกสามารถเขียนได้ดังสมการ

$$P(\text{เกิดเหตุการณ์ที่สนใจ}) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (2-24)$$

$$\text{หรือ } P(\text{เกิดเหตุการณ์ที่สนใจ}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (2-25)$$

เมื่อ β_0 และ β_1 เป็นสัมประสิทธิ์ที่ประมาณได้จากข้อมูล

X เป็นตัวแปรอิสระ

e มีค่าประมาณ 2.718

จากสมการข้างต้น สามารถเขียนใหม่ได้เป็น

$$P(\text{เกิดเหตุการณ์ที่สนใจ}) = \frac{e^w}{1 + e^{-w}} \quad (2-26)$$

$$\text{หรือ } P(\text{เกิดเหตุการณ์ที่สนใจ}) = \frac{1}{1 + e^{-w}} \quad (2-27)$$

$$\text{โดย } W = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2-28)$$

และโอกาสของการไม่เกิดเหตุการณ์จะประมาณได้จากสมการ (2-29)

$$P(\text{ไม่เกิดเหตุการณ์ที่สนใจ}) = 1 - P(\text{เกิดเหตุการณ์ที่สนใจ}) \quad (2-29)$$

ในการวิเคราะห์การถดถอยเชิงเส้น จะประมาณค่าพารามิเตอร์ในตัวแบบโดยใช้วิธีกำลังสองน้อยที่สุด (Least Square) โดยคัดเลือกสัมประสิทธิ์การถดถอยที่ทำให้ค่าผลรวมของกำลังสองของความคลาดเคลื่อนในการทำนาย $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ มีค่าน้อยที่สุด ส่วนการวิเคราะห์การถดถอยลอจิสติก จะประมาณค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด (Maximum-likelihood) อันเป็นการคำนวณเพื่อให้ได้ค่าประมาณของพารามิเตอร์ที่ใกล้เคียงกับข้อมูลเชิงประจักษ์มากที่สุด (ชนัญดาภรณ์ เย็นประเสริฐ, 2557)

การปรับรูปแบบความสัมพันธ์ที่ไม่ได้อยู่ในรูปเชิงเส้นให้อยู่ในรูปเชิงเส้น จากสมการของการวิเคราะห์การถดถอยลอจิสติก จะพบว่าความสัมพันธ์ไม่ได้อยู่ในรูปเชิงเส้น (Nonlinear) จึงทำการปรับให้อยู่ในรูปเชิงเส้น (Linear) ดังนี้

ให้อัตราส่วนออด (Odd Ratio : OR) เป็นอัตราส่วนระหว่างโอกาสที่จะเกิดเหตุการณ์ที่สนใจ กับโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ (สายชล สินสมบูรณ์ทอง, 2559)

$$\begin{aligned} \text{Odd ratio} = \text{OR} &= \frac{P(\text{เกิดเหตุการณ์ที่สนใจ})}{P(\text{ไม่เกิดเหตุการณ์ที่สนใจ})} \\ &= \frac{P(Y = 1)}{P(Y = 0)} \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\begin{aligned}
 &= \frac{e^{\beta_0 + \beta_1 x + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x + \dots + \beta_p x_p}} \\
 &= \frac{1}{1 + e^{\beta_0 + \beta_1 x + \dots + \beta_p x_p}} \\
 &= e^{\beta_0 + \beta_1 x + \dots + \beta_p x_p} \quad (2-30)
 \end{aligned}$$

ดังนั้น อัตราส่วนออดเป็นค่าที่แสดงถึงโอกาสที่จะเกิดเหตุการณ์ที่สนใจว่าเป็นกี่เท่าของโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ

ค่าประมาณของอัตราส่วนออดคือ

$$\hat{OR} = e^{b_0 + b_1 x + \dots + b_p x_p} \quad (2-31)$$

ความหมายของอัตราส่วนออด

ถ้า $OR > 1$ แสดงว่า โอกาสที่จะเกิดเหตุการณ์ที่สนใจมากกว่าโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ

ถ้า $OR = 1$ แสดงว่า โอกาสที่จะเกิดเหตุการณ์ที่สนใจและไม่เกิดเหตุการณ์ที่สนใจเท่ากัน

ถ้า $OR < 1$ แสดงว่า โอกาสที่จะเกิดเหตุการณ์ที่สนใจน้อยกว่าโอกาสที่จะไม่เกิดเหตุการณ์ที่สนใจ

2.10 วิธีนาอิว เบย์ (Naïve Bayes) (วรรณสิริ รุระชน และคณะ, 2557)

นาอิว เบย์เป็นเครื่องจักรเรียนรู้ที่อาศัยหลักการความน่าจะเป็น (Probability) ตามทฤษฎีของเบย์ (Bayes Theorem) ซึ่งมีอัลกอริทึมที่ไม่ซับซ้อน เป็นการจำแนกข้อมูลโดยการเรียนรู้ปัญหาที่เกิดขึ้นเพื่อนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ เป็นการจำแนกข้อมูลโดยใช้ความน่าจะเป็นและคำนวณการแจกแจงความน่าจะเป็นตามสมมติฐานที่ตั้งให้กับข้อมูล จากการคำนวณตัวอย่างใหม่ที่ได้จะถูกนำมาปรับเปลี่ยนการแจกแจง ซึ่งมีผลต่อการเพิ่มหรือลดความน่าจะเป็นของข้อมูล ข้อมูลใหม่ที่เกิดขึ้นและตัวแบบที่ตั้งไว้ให้กับข้อมูลจะถูกปรับเปลี่ยนไปตามข้อมูลใหม่ โดยผนวกกับข้อมูลเดิมที่มีหลักการของนาอิว เบย์ ใช้การคำนวณหาความน่าจะเป็นซึ่งถูกใช้ในการทำนายผล เป็นวิธีในการแก้ปัญหาแบบการจำแนกที่สามารถคาดการณ์ผลลัพธ์ได้ จะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์ นาอิว เบย์เป็นวิธีจำแนกกลุ่มข้อมูลที่มีประสิทธิภาพ มีอัลกอริทึมในการทำงานที่ไม่ซับซ้อน เหมาะกับกรณีของเซตตัวอย่างที่มีจำนวนมากและคุณสมบัติ (Attribute) ของตัวอย่างไม่ขึ้นต่อกันโดยกำหนดให้ความน่าจะเป็นของข้อมูลเท่ากับสมการ (2-32)

$$P(a_1, a_2, \dots, a_n | v_j) = P(v_j) \prod_{i=1}^n P(a_i | v_j) \quad (2-32)$$

โดยที่ $P(v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลที่ให้คำตอบ

$P(a_i | v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลคุณลักษณะที่ i มีค่า a_i และให้คำตอบ v_j

$P(a_1, a_2, \dots, a_n | v_j)$ คือ ค่าความน่าจะเป็นของข้อมูลทดสอบ (ที่มีค่าคุณลักษณะ a_1, a_2, \dots, a_n)

ที่จะให้คำตอบ v_j

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กลุ่ม v_j สำหรับข้อมูลที่มีคุณสมบัติ n ตัว คือ $X = \{a_1, a_2, \dots, a_n\}$ ใช้สัญลักษณ์ว่า $P(a_1, a_2, \dots, a_n | v_j)$

โดยที่ Π หมายถึง ผลคูณของค่า $(a_i | v_j)$ ทั้งหมด $i=1,2,3,\dots,n$ และ $j = 1,2,3,\dots,n$ กำหนดให้ E แทนข้อมูลที่นำมาใช้ในการคำนวณการแจกแจงความน่าจะเป็นหลัง (Posteriori Probability) ของสมมติฐาน h คือ $\Pr(h|E)$ ตามทฤษฎีของนาอ็ฟ เบย์ ดังสมการ (2-33)

$$\Pr(h|E) = \frac{\Pr(E|h)\Pr(h)}{\Pr(E)} \quad (2-33)$$

$\Pr(h)$ คือ ความน่าจะเป็นก่อนหน้าของสมมติฐาน h

$\Pr(E)$ คือ ความน่าจะเป็นก่อนหน้าของชุดข้อมูลตัวอย่าง E

$\Pr(h|E)$ คือ ความน่าจะเป็นของ h เมื่อรู้ E

$\Pr(E|h)$ คือ ความน่าจะเป็นของ E เมื่อรู้ h

2.11 การเปรียบเทียบประสิทธิภาพของวิธีการจำแนกกลุ่ม (Kinzang, W., 2009)

การวัดประเมินผลมีความสำคัญเนื่องจากการนำมาใช้ในการวิเคราะห์ประสิทธิภาพการทำงานของอัลกอริทึม ตัวชี้วัดการประเมินผลสามารถที่พัฒนาจากเมทริกซ์ความสับสน (Confusion Matrix) แสดงให้เห็นในรูปที่ 2-14

		ผลการจำแนก	
		คำตอบเป็นบวก	คำตอบเป็นลบ
ค่าที่แท้จริง	คำตอบเป็นบวก	TP (True Positive)	FN (False Negative)
	คำตอบเป็นลบ	FP (False Positive)	TN (True Negative)

รูปที่ 2-14 เมทริกซ์ความสับสน (Confusion Matrix)

โดยที่ True Positive (TP) คือจำนวนข้อมูลที่จำแนกถูกว่าเป็นบวก

True Negative (TN) คือจำนวนข้อมูลที่จำแนกถูกว่าเป็นลบ

False Positive (FP) คือจำนวนข้อมูลที่จำแนกผิดว่าเป็นบวก ซึ่งค่าที่แท้จริงเป็นลบ

False Negative (FN) คือจำนวนข้อมูลที่จำแนกผิดว่าเป็นลบ ซึ่งค่าที่แท้จริงเป็นบวก

2.11.1 ค่าความถูกต้อง (Accuracy) คือ การแสดงผลการวัดที่ได้มีความถูกต้องในรูปอัตราส่วน

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \% \quad (2-34)$$

2.11.2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยใช้หลักการเดียวกันกับการหาค่าความแปรปรวนในทางสถิติ การวัดค่าความคลาดเคลื่อนด้วยวิธีนี้ได้ค่าความคลาดเคลื่อนที่สูง เนื่องจากการนำความคลาดเคลื่อน ณ เวลาใดๆ มายกกำลังสองก่อนที่จะหาผลรวม แล้วจึงนำมาหาค่าเฉลี่ยอีกครั้งหนึ่ง นั่นคือ ค่า MSE ยิ่งน้อยหมายถึง การพยากรณ์ยิ่งแม่นยำ มีสูตรในการคำนวณดังนี้ (จุฑามาศ สิทธิโชคสถาพร, 2555)

$$\text{MSE} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (2-35)$$

โดยที่ y_i แทน ค่าจริง
 \hat{y}_i แทน ค่าพยากรณ์

2.12 งานวิจัยที่เกี่ยวข้อง

ไกรรุ่ง เสงพระพรหม และคณะ (2547) ได้นำเสนอทดลองกับวิธีทำเหมืองข้อมูล 2 วิธี ได้แก่ วิธีต้นไม้ตัดสินใจ J48 และ วิธีนาอิว เบย์ โดยแบ่งการทดลองออกเป็น 2 ลักษณะ คือ การทดลองกับชุดข้อมูลทั้งหมดทุกคุณลักษณะและการทดลองด้วยการเลือกคุณลักษณะด้วยวิธี SNR (Signal to Noise Ratio) เพื่อเพิ่มประสิทธิภาพในการทำงานของอัลกอริทึมซึ่งจากผลการทดลองทั้งสองแบบยืนยันว่าวิธีต้นไม้ตัดสินใจ J48 ให้ประสิทธิภาพในการทำงานที่ดีกว่าเมื่อเทียบกับวิธีนาอิว เบย์

บุษกร สุคนธวงศาโรจน์ (2551) ได้นำเสนอการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้ โดยข้อมูลที่ใช้ในการทดลองได้จากหน่วยงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง โดยทำการทดลองกับกลุ่มตัวอย่าง 28 หน่วยงาน โดยใช้วิธีการจำแนกกลุ่มความเชี่ยวชาญเชิงเปรียบเทียบด้วยอัลกอริทึม 3 แบบ ได้แก่ Decision Tree (J48), OneR และ Naïve Bayes การทดลองเพื่อหาอัลกอริทึมที่มีความแม่นยำสูงสุดในการทำนายความเชี่ยวชาญ โดยใช้ตัววัดประสิทธิภาพ 4 ตัว ได้แก่ Accuracy, Precision/Recall, F-measure และ Root Mean-Squared Error (RMSE) จากการทดลองพบว่าอัลกอริทึมที่ให้ค่าความถูกต้องในการทำนายคำตอบของความเชี่ยวชาญสูงสุด คือ J48

พลอยพรรณ สอนสุวิทย์ และ ตรัสพงศ์ ไทยอุบลัมภ์ (2552) ได้นำเสนอการเปรียบเทียบประสิทธิภาพการตรวจจับสิ่งผิดกติทางเครือข่าย โดยการสแกนหาจุดอ่อน ซึ่งเป็นการบุกรุกทางเครือข่ายที่มีความสำคัญ งานวิจัยนี้ได้ศึกษาวิธีการจำแนกกลุ่ม (Classification) มาตรวจจับการ

สแกน ได้แก่ SVM, J48, Naïve Bayes และ Artificial Neural Network จากการศึกษาทำให้ทราบถึงประสิทธิภาพของแต่ละวิธี เช่น ร้อยละของความถูกต้อง (%Correct) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error) พบว่า วิธี J48 วิเคราะห์ได้แม่นยำที่สุดและมีค่า Error ต่ำที่สุด

วาทีณี น้อยเพียร และคณะ (2553) ได้ทำการศึกษาเพื่อเปรียบเทียบวิธีการจำแนกกลุ่มข้อมูล โดยเลือกใช้อัลกอริทึมโครงข่ายประสาทเทียมแบบมัลติเลเยอร์เปอร์เซ็ปตรอน ซัพพอร์ตเวกเตอร์แมชชีน นาอิว เบย์และเคเนียร์เรสต์เนเบอร์ เพื่อประเมินประสิทธิภาพค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความถ่วงดุล (F-Measure) ใช้ข้อมูลจาก UCI ประกอบด้วย Ozone Days และ Adult เลือกกลุ่มข้อมูล โดยมีจำนวนกลุ่ม (Class) เท่ากันในข้อมูลแต่ละชุด เป็นการทดลองแบบมีการเรียนรู้ จากผลการวิจัยอัลกอริทึมที่ดีที่สุดของข้อมูล Ozone Days คือซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบ Rbf มีค่าความถูกต้อง 95.83% ค่าความแม่นยำ 96% ค่าความระลึก 96% และค่าความถ่วงดุล 96% ส่วนข้อมูล Adult คือซัพพอร์ตเวกเตอร์แมชชีน ฟังก์ชันเคอร์เนลแบบ Polynomial มีค่าความถูกต้อง 79.66% ค่าความแม่นยำ 80% ค่าความระลึก 80% และค่าความถ่วงดุล 80%

วลัยลักษณ์ สุขสมบุญ และสมชาย ปราการเจริญ (2553) ได้นำเสนอการเปรียบเทียบวิธีการจำแนกกลุ่มของปัญหาสำหรับ Helpdesk ด้วยซัพพอร์ตเวกเตอร์แมชชีน นาอิว เบย์และเคเนียร์เรสต์เนเบอร์ โดยนำข้อมูลมาตัดคำภาษาอังกฤษด้วยวิธีการอ้างอิงพจนานุกรมและคลังคำศัพท์เฉพาะที่ผู้วิจัยสร้างขึ้นใหม่ จากนั้นนำข้อมูลปัญหา Helpdesk แปลงให้อยู่ในรูปของเวกเตอร์ทางคณิตศาสตร์และทำการทดลองด้วยซัพพอร์ตเวกเตอร์แมชชีน นาอิว เบย์และเคเนียร์เรสต์เนเบอร์โดยใช้การตรวจสอบไขว้กัน K-Fold = 10 โดยทดลองปรับค่าพารามิเตอร์ของฟังก์ชันแกน Linear, Polynomial และ Radius Basic Function ในซัพพอร์ตเวกเตอร์แมชชีน นาอิว เบย์และเคเนียร์เรสต์ เนเบอร์ ผลการทดลองสามารถสรุปได้ว่าการทดลองเปรียบเทียบวิธีการจำแนกกลุ่มของปัญหาสำหรับ Helpdesk แสดงให้เห็นว่าวิธีที่นำเสนอด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องที่ดีที่สุดโดยใช้ฟังก์ชันแกน Polynomial

เดช ธรรมศิริ และพยุ่ง มีสังข์ (2554) ได้นำเสนอการจำแนกข้อมูลโดยทำการทดสอบประสิทธิภาพด้วยข้อมูล Austrian Credit และ Bankruptcy Data โดยนำเอาวิธีซัพพอร์ตเวกเตอร์แมชชีนและหาค่าพารามิเตอร์ที่เหมาะสมร่วมด้วยการเลือกคุณลักษณะที่เหมาะสมโดยใช้อัลกอริทึมเชิงพันธุกรรมเปรียบเทียบผลการวิจัยกับวิธีต้นไม้ตัดสินใจ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีนกับวิธีเชิงพันธุกรรม พบว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนที่ใช้อัลกอริทึมเชิงพันธุกรรมจะให้ค่าความแม่นยำสูงที่สุด

อรนุช พันโท และมนชัย เทียนทอง (2557) ได้นำเสนอการเปรียบเทียบประสิทธิภาพการจำแนกรูปแบบการเรียนรู้ VARK ด้วยวิธีนาอิว เบย์ (Naive Bayes) วิธีต้นไม้ตัดสินใจ (Decision Tree) และวิธีฐานกฎ (Rules Based) ผลการศึกษาพบว่าประสิทธิภาพการจำแนกข้อมูลแบบต้นไม้ตัดสินใจมีประสิทธิภาพสูงสุด นั่นคืออัลกอริทึม J48 ให้ค่าความถูกต้องเท่ากับ 82.78% และ

NBTree เท่ากับ 81.78% รองลงมาคือวิธีฐานกฎ และวิธีแบบ เบย์ ตามลำดับ

ชณัฐตาภรณ์ เย็นประเสริฐ (2557) ได้นำเสนอการทดสอบประสิทธิภาพการพยากรณ์ สถานะการชำระหนี้ของลูกค้าธนาคารด้วยวิธีต้นไม้ตัดสินใจ (Decision Tree) วิธีการถดถอย ลอจิสติก (Logistic Regresstion) และวิธีนาอ็ฟ เบย์ (Naive Bayes) ผลลัพธ์จากวิธีการพยากรณ์ ความแม่นยำของสถานการณ์ชำระหนี้ของลูกค้าพบว่า วิธีต้นไม้ตัดสินใจมีความแม่นยำเท่ากับ 89.80% มีค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ 10.20% รากที่สองของค่าเฉลี่ยคลาดเคลื่อนกำลังสอง เท่ากับ 0.30 รองลงมาคือวิธีการถดถอยลอจิสติกมีความแม่นยำเท่ากับ 89.80% ค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์เท่ากับ 10.20% รากที่สองของค่าเฉลี่ยคลาดเคลื่อนกำลังสองเท่ากับ 0.31 และวิธี นาอ็ฟ เบย์ มีความแม่นยำเท่ากับ 88.78% ค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์เท่ากับ 11.22% รากที่ สองของค่าเฉลี่ยคลาดเคลื่อนกำลังสองเท่ากับ 0.33 จากผลการทดลองกรณีศึกษาพบว่าวิธีต้นไม้ ตัดสินใจและวิธีการถดถอยลอจิสติกมีค่าความแม่นยำ ค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ และรากที่สอง ของค่าเฉลี่ยคลาดเคลื่อนกำลังสองใกล้เคียงกัน ในขณะที่วิธีนาอ็ฟ เบย์ มีค่าความแม่นยำต่ำ ค่าเฉลี่ย ความคลาดเคลื่อนสัมพัทธ์และรากที่สองของค่าเฉลี่ยคลาดเคลื่อนกำลังสองสูงกว่า

เบญจกัท จงหมื่นไวย (2558) ได้นำเสนอการเปรียบเทียบปัจจัยโรคประจำตัวผู้สูงอายุจาก ข้อมูลผู้สูงอายุของศูนย์บริการสาธารณสุขโพธิ์กลาง จังหวัดนครราชสีมา โดยข้อมูลที่นำมาวิเคราะห์มี จำนวนผู้สูงอายุทั้งหมด 594 ราย ซึ่งใช้อัลกอริทึมการจำแนกกลุ่ม J48 และ NaiveBayes ผลการ ทดสอบพบว่าวิธีต้นไม้ตัดสินใจโดยใช้อัลกอริทึม J48 ให้ค่าความถูกต้องสูงกว่าวิธีนาอ็ฟ เบย์

Vijayarani, S. and Dhayanand, S. (2015) ได้นำเสนอการทำนายผลของข้อมูลโรคไต เรื้อรัง ซึ่งเก็บรวบรวมข้อมูลจากกรมสาธารณสุขด้วยเทคนิคการทำเหมืองข้อมูล โดยใช้วิธีซัพพอร์ต เวกเตอร์แมชชีนและโครงข่ายประสาทเทียม จากการศึกษาพบว่าวิธีโครงข่ายประสาทเทียมให้ค่า ความถูกต้องคือ 87.70% และวิธีซัพพอร์ตเวกเตอร์แมชชีนให้ค่าความถูกต้องคือ 76.32% จึงสรุปได้ ว่าวิธีโครงข่ายประสาทเทียมมีประสิทธิภาพในการทำนายผลได้ดีกว่าวิธีซัพพอร์ตเวกเตอร์แมชชีน

Lakshmi, K. R. et al. (2014) ได้นำเสนอการเปรียบเทียบประสิทธิภาพด้วยเทคนิคการทำ เหมืองข้อมูลเพื่อทำนายผู้มีชีวิตรอดจากการล้างไต ซึ่งเก็บรวบรวมข้อมูลจากเว็บไซต์ฟอกไต ในการศึกษารังนี้ใช้การเปรียบเทียบวิธีจำแนกกลุ่ม 3 วิธี ได้แก่ วิธีโครงข่ายประสาทเทียม วิธีต้นไม้ ตัดสินใจและวิธีการถดถอยลอจิสติก พบว่าวิธีโครงข่ายประสาทเทียมให้ค่าความถูกต้องและค่าความ ระวังมากที่สุดคือ 93.85% และ 0.9387 รองลงมาวิธีต้นไม้ตัดสินใจให้ค่าความถูกต้องและค่าความ ระวังคือ 78.45% และ 0.8008 และวิธีการถดถอยลอจิสติกให้ค่าความถูกต้องและค่าความระวังคือ 74.75% และ 0.7642 ตามลำดับ

Kharya, S. (2012) ได้นำเสนอการวิเคราะห์และทำนายผลเกี่ยวกับโรคมะเร็งเต้านมสำหรับ ผู้หญิงในประเทศอินเดีย ด้วยเทคนิคการทำเหมืองข้อมูล จากการศึกษาพบว่าวิธีต้นไม้ตัดสินใจให้ค่า ความถูกต้องมากที่สุดคือ 86.7% รองลงมาคือวิธีโครงข่ายประสาทเทียมให้ค่าความถูกต้อง 86.5% และวิธีนาอ็ฟ เบย์ให้ค่าความถูกต้องคือ 84.5% ตามลำดับ

บทที่ 3

วิธีการดำเนินงานวิจัย

3.1 การเก็บรวบรวมข้อมูล

โรคไตเรื้อรัง (Chronic Kidney Disease) เป็นปัญหาทางด้านสาธารณสุขที่สำคัญของคนทั่วไป มีผลกระทบต่อประชาชนทุกอายุ เชื้อชาติและทุกสถานะทางเศรษฐกิจ ความชุกและอุบัติการณ์ของโรคที่เพิ่มขึ้นเนื่องมาจากโรคเบาหวาน ความดันโลหิตสูงและโรคอ้วน นอกจากนี้ผู้ป่วยโรคไตเรื้อรังมีความเสี่ยงเพิ่มขึ้นที่จะเกิดภาวะหลอดเลือดตีบแข็ง ทำให้เกิดโรคหัวใจและโรคหลอดเลือดตามมา และเป็นสาเหตุสำคัญที่ทำให้ผู้ป่วยเสียชีวิต นอกจากนี้โรคไตเรื้อรังเรื้อรังถือเป็นปัญหาสาธารณสุขที่สำคัญของทั่วโลก รวมถึงประเทศไทย ผู้วิจัยจึงได้มีความสนใจที่จะนำข้อมูลเกี่ยวกับโรคไตเรื้อรังมาทำการวิเคราะห์ในการทำวิทยานิพนธ์ครั้งนี้ ซึ่งได้มีการค้นคว้าข้อมูลจากฐานข้อมูล UCI Machine Learning Repository จำนวน 1 ชุด คือ ข้อมูลผู้ป่วยโรคไตเรื้อรังของโรงพยาบาลอพอลโล ซึ่งเป็นโรงพยาบาลเอกชนชื่อดังในกรุงนิวเดลี ประเทศอินเดีย (Chronic Kidney Disease of Apollo Hospitals, India) รวบรวมข้อมูลผู้ป่วยโรคไตเรื้อรังโดย L.Jerlin Rubini นักวิชาการจากมหาวิทยาลัย อลากัปปา (L.Jerlin Rubini Research Scholar Alagappa University) บริจาคข้อมูล ณ วันที่ 3 กรกฎาคม 2558 ซึ่งได้มีการเก็บรวบรวมข้อมูลเป็นระยะเวลา 2 เดือน จำนวนข้อมูลที่เก็บรวบรวมทั้งหมด 400 ระเบียบ ประกอบด้วยตัวแปรอิสระ 24 ตัวแปร และตัวแปรตาม 1 ตัวแปร โดยมีรายละเอียดดังนี้

(http://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease)

ตารางที่ 3.1 ตัวแปรอิสระและตัวแปรตามของข้อมูลผู้ป่วยโรคไตเรื้อรัง

ตัวแปร	ความหมาย	มาตรวัด	หน่วย
X ₁	อายุ (age)	numerical	ปี (year)
X ₂	ความดันโลหิต (blood pressure)	numerical	มิลลิเมตรต่อปรอท (mm/Hg)
X ₃	ความถ่วงจำเพาะของปัสสาวะ (urine specific gravity)	nominal	ไม่มีหน่วย
X ₄	ระดับโปรตีนอัลบูมิน (albumin)	nominal	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₅	ระดับน้ำตาล (sugar)	nominal	ความเข้มข้นโมลาร์ (molar concentration)
X ₆	เซลล์เม็ดเลือดแดง (red blood cells)	nominal	ปกติ, ไม่ปกติ (normal, abnormal)
X ₇	เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่น (pus cell)	nominal	ปกติ, ไม่ปกติ (normal, abnormal)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 ตัวแปรอิสระและตัวแปรตามของข้อมูลผู้ป่วยโรคไตเรื้อรัง (ต่อ)

ตัวแปร	ความหมาย	มาตรวัด	หน่วย
X ₈	เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่นเป็นกลุ่มก้อน (pus cell clumps)	nominal	ปรากฏ, ไม่ปรากฏ (present, notpresent)
X ₉	แบคทีเรีย (bacteria)	nominal	ปรากฏ, ไม่ปรากฏ (present, notpresent)
X ₁₀	ปริมาณน้ำตาลในเลือดแบบสุ่ม (blood glucose random)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₁	ปริมาณไนโตรเจนในกระแสเลือด (blood urea nitrogen)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₂	ปริมาณครีเอตินินในเลือด (serum creatinine)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₃	ปริมาณโซเดียมในเลือด (sodium)	numerical	มิลลิอิกวาเลนซ์ต่อลิตร (mEq/L)
X ₁₄	ปริมาณโปแตสเซียมในเลือด (potassium)	numerical	มิลลิอิกวาเลนซ์ต่อลิตร (mEq/L)
X ₁₅	ปริมาณฮีโมโกลบิน (hemoglobin)	numerical	กรัมต่อเดซิลิตร (gms/dl)
X ₁₆	ปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมด (packed cell volume)	numerical	ยูนิต (Unit)
X ₁₇	จำนวนเซลล์เม็ดเลือดขาว (white blood cell count)	numerical	เซลล์ต่อลูกบาศก์มิลลิเมตร (cells/cumm)
X ₁₈	จำนวนเซลล์เม็ดเลือดแดง (red blood cell count)	numerical	ล้านเซลล์ต่อลูกบาศก์มิลลิเมตร (millions/cumm)
X ₁₉	ภาวะความดันโลหิตสูง (hypertension)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₀	เบาหวาน (diabetes mellitus)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₁	โรคหลอดเลือดหัวใจ (coronary artery disease)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₂	ความอยากอาหาร (appetite)	nominal	ต้องการมาก, ต้องการน้อย (good, poor)
X ₂₃	อาการเท้าบวม (pedal edema)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₄	โรคโลหิตจาง (anemia)	nominal	เป็น, ไม่เป็น (yes, no)
Y	การเป็นโรคไตเรื้อรัง (Chronic kidney)	nominal	0 = ผู้ป่วยไม่เป็นโรคไตเรื้อรัง 1 = ผู้ป่วยเป็นโรคไตเรื้อรัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 เครื่องมือที่ใช้ในการวิจัย

โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) เวอร์ชัน 3.7.12 ซึ่งเป็นโปรแกรมที่สามารถดาวน์โหลดได้จากเว็บไซต์ ซึ่งอยู่ภายใต้การควบคุมของ GPL License ซึ่งโปรแกรม WEKA ได้ถูกพัฒนามาจากภาษาจาวาทั้งหมด ซึ่งเป็นที่นิยมในการใช้งานด้านการทำเหมืองข้อมูล รายละเอียดการใช้โปรแกรมในภาคผนวก ก (ดังรูปที่ ก-1 ถึง ก-9)

3.3 ขั้นตอนในการดำเนินงานวิจัย

- 1) ศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้องกับงานวิจัย
- 2) ศึกษาและหาข้อมูลเพื่อนำมาเปรียบเทียบประสิทธิภาพในการทำนายผลการจำแนกกลุ่ม ด้วยวิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) วิธีต้นไม้ตัดสินใจ (Decision Tree) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) วิธีฐานกฎ (Rule-Based) วิธีการถดถอยลอจิสติก (Logistic Regression) และวิธีนาอิวเบย์ (Naïve Bayes)
- 3) ศึกษาอัลกอริทึมของแต่ละวิธี ดังนี้

3.1) วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) ใช้อัลกอริทึมชนิด IBk เนื่องจากเป็นฟังก์ชันหลักที่สนใจ ซึ่งเป็นพื้นฐานของอัลกอริทึม อย่างไรก็ตาม อัลกอริทึม IBk ยังสามารถกำหนดน้ำหนัก ระยะทางและทางเลือก (Option) เพื่อกำหนดค่า k โดยใช้ cross-validation (Wu, X. and Kumar, V., 2009) ซึ่งในการศึกษาครั้งนี้ ผู้วิจัยได้กำหนดค่า $k = 1$ เนื่องจากให้ค่าความถูกต้องมากที่สุดและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุด รายละเอียดในบทที่ 4 (ดังรูปที่ 4-4)

3.2) วิธีต้นไม้ตัดสินใจ (Decision Trees) ใช้อัลกอริทึมชนิด J48 ซึ่งพัฒนาจาก ID3 สามารถใช้ได้กับข้อมูลแบบไม่ต่อเนื่องและแบบต่อเนื่อง ต่างจาก ID3 ที่ใช้ได้เพียงข้อมูลแบบไม่ต่อเนื่องเท่านั้น (รุจิรา ธรรมสมบัติ, 2554)

3.3) วิธีโครงข่ายประสาทเทียม (Artificial Neural Networks) ใช้อัลกอริทึมชนิดเพอร์เซปตรอนหลายชั้น (Multilayer Perceptron) โดยกำหนดค่าอัตราการเรียนรู้ (Learning Rate) เป็น 0.1 ค่าโมเมนตัม (Momentum) เป็น 0.9 จำนวนรอบการสอน (Training Time) 20,000 รอบ การวิจัยครั้งนี้ใช้อัลกอริทึมของวิธีโครงข่ายประสาทเทียมแบบเพอร์เซปตรอนหลายชั้น ที่มีชั้นซ่อน (Hidden Layer) 1 ชั้น แม้ว่าโครงสร้างโครงข่ายประสาทเทียมที่ซับซ้อนสามารถมีชั้นซ่อนมากกว่า 1 ชั้น แต่ในทางปฏิบัติการกำหนดชั้นซ่อน 1 ชั้นก็เพียงพอต่อการวิเคราะห์ข้อมูล (Berson, A. and Smith, S. J., 1997)

3.4) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) ใช้อัลกอริทึมชนิดโพลิโนเมียล เคอร์เนล (Polynomial Kernel) เนื่องจากงานวิจัยที่อ้างอิงจาก วาทีนี นุ้ยเพียร และคณะ (2553) ได้ผลลัพธ์ว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนที่ใช้อัลกอริทึมชนิด Polynomial Kernel ดีที่สุด

3.5) วิธีฐานกฎ (Rule-Based) ใช้อัลกอริทึม Decision Table เป็นเครื่องมือที่ใช้แสดงเงื่อนไขการตัดสินใจและเลือกการทำงานหรือกระทำกิจกรรมภายใต้เหตุการณ์ของเงื่อนไขที่ระบุ วิธีการตัดสินใจแบบ Decision Table จะเป็นตาราง 2 มิติอย่างง่าย (อรนุช พันโทและมนชัย เทียนทอง, 2557)

3.6) วิธีการถดถอยลอจิสติก (Logistic Regression) เนื่องจากงานวิจัยที่อ้างอิงจากชนัฐดาภรณ์ เย็นประเสริฐ (2557) กรณีศึกษาวิธีการถดถอยลอจิสติกเปรียบเทียบค่าความแม่นยำ ค่าเฉลี่ยความคลาดเคลื่อนสัมพัทธ์ และรากที่สองของค่าเฉลี่ยคลาดเคลื่อนกำลังสอง

3.7) วิธีนาอีฟ เบย์ (Naïve Bayes) เป็นวิธีจำแนกกลุ่มข้อมูลที่มีประสิทธิภาพ มีอัลกอริทึมในการทำงานที่ไม่ซับซ้อน เหมาะกับกรณีของเซตตัวอย่างที่มีจำนวนมากและคุณสมบัติ (Attribute) (วรพจน์ สุเมธาวัฒนา และคณะ, 2557)

4) แบ่งข้อมูลออกเป็น 2 ส่วน สุ่มด้วยวิธีการสุ่มตัวอย่างอย่างง่าย (Simple Random Sampling, SRS) โดยใช้โปรแกรม SPSS ในการสุ่มข้อมูลครั้งนี้ ข้อมูลส่วนที่ 1 สำหรับการสร้างตัวแบบ 70 เปอร์เซ็นต์ หรือจำนวน 280 ระเบียบ และข้อมูลส่วนที่ 2 สำหรับการทดสอบตัวแบบ 30 เปอร์เซ็นต์ หรือจำนวน 120 ระเบียบ (รุจิรา ธรรมสมบัติ, 2554) รายละเอียดในบทที่ 4 (ดังรูปที่ 4-1 ถึง 4-3)

5) นำข้อมูลหลังจากแบ่งออกเป็น 2 ส่วนแล้ว มาทำการวิเคราะห์ด้วยโปรแกรม WEKA

6) นำผลการวิเคราะห์ของแต่ละวิธีทั้ง 7 วิธี มาเปรียบเทียบประสิทธิภาพในวิธีการจำแนกกลุ่ม ซึ่งในงานวิจัยอาศัยค่าความถูกต้อง (Accuracy) เนื่องจากเป็นค่าที่บ่งบอกถึงความสามารถของเครื่องมือวัด (Instrument) ในการอ่านค่าหรือแสดงค่าที่วัดได้เข้าใกล้ค่าจริง นอกจากนี้ งานวิจัยที่ผู้วิจัยได้ไปศึกษาค้นคว้ามาส่วนใหญ่แล้วได้ใช้ค่าความถูกต้อง (Accuracy) ทุกงานวิจัย และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) เนื่องจากเป็นมาตรวัดการประเมินค่าได้ดี เนื่องจากความคลาดเคลื่อนกำลังสองเฉลี่ยประกอบด้วยทั้งความแปรปรวนและความเอนเอียง (Larose D. T., 2005) โดยผู้วิจัยมีความสนใจที่จะนำมาใช้วิเคราะห์ในการทำวิทยานิพนธ์ครั้งนี้

7) สรุปผลงานวิจัย อภิปรายผล และข้อเสนอแนะ

3.4 วิธีการวิเคราะห์ข้อมูล

ในการทำงานวิจัยครั้งนี้ ผู้วิจัยได้จัดข้อมูลออกเป็น 2 ส่วน ใช้โปรแกรม SPSS ในการสุ่มแบ่งข้อมูล โดยส่วนที่ 1 ใช้ข้อมูล 70 เปอร์เซ็นต์ของข้อมูลทั้งหมดในการสร้างตัวแบบ ส่วนที่ 2 ใช้ข้อมูล 30 เปอร์เซ็นต์ของข้อมูลทั้งหมดในการทดสอบตัวแบบ แปลงไฟล์ข้อมูลให้เป็นนามสกุล *.csv เพื่อใช้วิเคราะห์ประสิทธิภาพการจำแนกกลุ่มข้อมูลในโปรแกรม WEKA ซึ่งเป็นโปรแกรมที่สามารถนำมาทดสอบอัลกอริทึมของวิธีการจำแนกกลุ่มได้ เนื่องจากมีอัลกอริทึมที่ได้รับไว้ให้เลือกใช้ในโปรแกรมครบตามที่กำหนด ผู้วิจัยได้กำหนดวิธีการจำแนกกลุ่มเพื่อนำผลที่ได้ของแต่ละวิธีมาเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มของโรคไตเรื้อรัง โดยใช้เกณฑ์ในการวัดคือ ค่าความถูกต้อง (Accuracy)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) โดยที่ค่าความถูกต้องมีค่ามากแสดงว่าวิธีการจำแนกกลุ่มนั้นมีประสิทธิภาพสูง และค่าความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าต่ำแสดงว่าวิธีการจำแนกกลุ่มนั้นมีประสิทธิภาพสูง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการวิเคราะห์ข้อมูล

งานวิจัยครั้งนี้ผู้วิจัยใช้การจำแนกกลุ่มโดยใช้วิธีการทำเหมืองข้อมูล โดยนำข้อมูลผู้ป่วยโรคไตเรื้อรัง ประเทศอินเดีย จำนวน 400 ระเบียบ มาทำการวิเคราะห์ข้อมูล ซึ่งจะสุ่มแบ่งข้อมูลออกเป็น 2 ส่วน คือ ข้อมูลส่วนที่ 1 ข้อมูลเรียนรู้ (Training Data) นำไปสร้างตัวแบบ 280 ระเบียบ และข้อมูลส่วนที่ 2 ข้อมูลการทดสอบ (Testing Data) นำไปทดสอบตัวแบบ 120 ระเบียบ และผู้วิจัยได้นำวิธีเหล่านี้มาประยุกต์กับงานวิจัย โดยนำมาเปรียบเทียบประสิทธิภาพผลการจำแนกกลุ่มในการทำวิทยานิพนธ์ครั้งนี้ ผู้วิจัยได้สนใจศึกษาการจำแนกกลุ่ม (Classification) เพื่อเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มจากค่าความถูกต้อง (Accuracy) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) ซึ่งวิธีที่ใช้ในการทดสอบในครั้งนี้มี 7 วิธี ผลการวิเคราะห์ข้อมูลแบ่งเป็นหัวข้อดังนี้

- 4.1 ผลการแบ่งข้อมูลด้วยโปรแกรม SPSS
- 4.2 ผลการวิเคราะห์ด้วยวิธีความใกล้เคียงกันมากที่สุดโดยใช้อัลกอริทึมชนิด IBk
- 4.3 ผลการวิเคราะห์ด้วยวิธีต้นไม้ตัดสินใจโดยใช้อัลกอริทึมชนิด J48
- 4.4 ผลการวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียมโดยใช้อัลกอริทึมชนิด Multilayer Perceptron
- 4.5 ผลการวิเคราะห์ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนโดยใช้อัลกอริทึมชนิด Polynomial Kernel
- 4.6 ผลการวิเคราะห์ด้วยวิธีฐานกฎโดยใช้อัลกอริทึมชนิด Decision Table
- 4.7 ผลการวิเคราะห์ด้วยวิธีการถดถอยลอจิสติก
- 4.8 ผลการวิเคราะห์ด้วยวิธีนาอีฟ เบย์

รายละเอียดของข้อมูลผู้ป่วยโรคไตเรื้อรัง

จากข้อมูลผู้ป่วยโรคไตเรื้อรังของโรงพยาบาลอพอลโล ประเทศอินเดีย (Chronic Kidney Disease of Apollo Hospitals, India) ทั้งหมด 400 ระเบียบน ประกอบด้วย 24 ตัวแปรอิสระ และ 1 ตัวแปรตาม ดังนี้

ตัวแปรอิสระ

1) ตัวแปรอายุ (Age) ประกอบด้วยทุกช่วงวัย ได้แก่ น้อยกว่า 11 ปี, 11-20 ปี, 21-30 ปี, 31-40 ปี, 41-50 ปี, 51-60 ปี, 61-70 ปี และมากกว่า 70 ปี ซึ่งส่วนใหญ่จะอยู่ในช่วงวัย 51-60 ปี รองลงมา 61-70 ปี ตามลำดับ อายุเฉลี่ยในการวิเคราะห์ครั้งนี้ คือ 52 ปี

2) ความดันโลหิต (Blood Pressure) คือ ความดันในหลอดเลือดเมื่อหัวใจบีบตัวสูบเลือดเข้าสู่หลอดเลือด ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นี้ส่วนใหญ่แล้วจะเข้าเกณฑ์ปกติ

3) ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) คือ ค่าความหนาแน่นของปัสสาวะเมื่อเปรียบเทียบกับความหนาแน่นของน้ำในอัตราส่วนปริมาณที่เท่ากัน โดยค่าความถ่วงจำเพาะของปัสสาวะปกติจะอยู่ในช่วง 1.001 - 1.030 ซึ่งค่านี้จะเปลี่ยนแปลงตามปริมาณน้ำที่ดื่ม ถ้าดื่มน้ำมากค่าก็จะต่ำลง ถ้าขาดน้ำหรือดื่มน้ำน้อยค่าก็จะสูงขึ้น ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้มีค่าความถ่วงจำเพาะของปัสสาวะอยู่ในเกณฑ์ปกติ

4) ระดับโปรตีนอัลบูมิน (Albumin) คือ เป็นโปรตีนที่สร้างจากตับเป็นส่วนประกอบที่สำคัญของน้ำเหลืองในเลือด ระดับอัลบูมินเป็นการบ่งบอกความสมดุลของการสร้างที่ตับและการขับอัลบูมินออกจากไตพบว่า อัลบูมินจะเป็นสัดส่วนได้ประมาณ 60 เปอร์เซ็นต์ ของโปรตีนทั้งหมดที่ผลิตจากตับ อัลบูมินทำหน้าที่สร้างความเข้มข้นขึ้นในหลอดเลือดจะเกิดความดันออสโมติก (Osmotic Pressure) ป้องกันสารอาหารมิให้รั่วซึมออกมาภายนอกหลอดเลือด นอกจากนี้ อัลบูมินยังมีบทบาทที่สำคัญ ได้แก่ การทำหน้าที่เป็นพาหะขนส่งฮอร์โมน เอนไซม์ แร่ธาตุ หรือแม้แต่ยารักษาโรค ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นี้ส่วนใหญ่แล้วระดับโปรตีนอัลบูมินจะอยู่ในเกณฑ์ปกติ

5) ระดับน้ำตาล (Sugar) คือ เป็นการบอกระดับน้ำตาลที่อยู่ในเลือด โดยปกติแล้วหลังจากการรับประทานอาหาร ร่างกายจะดูดซึมที่ลำไส้เล็กและนำไปใช้ทั่วร่างกาย โดยน้ำตาลกลูโคส (Glucose) เป็นแหล่งพลังงานแรกที่ร่างกายนำไปใช้ โดยเฉพาะอย่างยิ่งคือ ระบบสมองและระบบประสาท ซึ่งฮอร์โมนที่เป็นตัวนำน้ำตาลไปให้ร่างกายใช้เป็นพลังงานได้นั้นคือ ฮอร์โมนอินซูลิน แต่หากร่างกายขาดฮอร์โมนอินซูลินนี้หรือดื้อต่ออินซูลิน (Insulin Resistance) ซึ่งพบในผู้ป่วยเบาหวาน จะทำให้น้ำตาลอยู่ในกระแสเลือดไม่สามารถนำไปให้อวัยวะต่างๆ ในร่างกายใช้ได้ ซึ่งการมีระดับน้ำตาลในกระแสเลือดสูงเป็นเวลานาน ทำให้ส่งผลต่อโรคแทรกซ้อน เช่น ตาเสื่อม ไตเสื่อม โรคหลอดเลือดหัวใจ และระบบประสาทเสื่อมตามมา หากร่างกายมีระดับน้ำตาลต่ำมาก สามารถเป็นอันตรายต่อสมองและระบบประสาทถูกทำลายและอาจมีผลต่อชีวิตได้ ดังนั้นการตรวจวัดระดับน้ำตาลด้วยวิธีนี้จึงสามารถบอกถึงระดับน้ำตาลที่ถูกต้องในช่วงเวลานั้นๆ เพื่อป้องกันและแก้ไขปัญหาต่างๆ ได้ทันทั่วทั้ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6) เซลล์เม็ดเลือดแดง (Red Blood Cells) มีหน้าที่ในการส่งถ่ายออกซิเจนไปเลี้ยงเซลล์ต่างๆ ทั่วร่างกาย ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะส่วนใหญ่แล้วจะเข้าเกณฑ์ปกติ

7) เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่น (Pus Cell) ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะส่วนใหญ่แล้วจะเข้าเกณฑ์ปกติ

8) เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่นเป็นกลุ่มก้อน (Pus Cell Clumps) ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะส่วนใหญ่แล้วจะไม่ปรากฏเป็นกลุ่มก้อน

9) แบคทีเรีย (Bacteria) คือ ประเภทของสิ่งมีชีวิตประเภทใหญ่ประเภทหนึ่ง มีขนาดเล็กมองด้วยตาเปล่าไม่เห็น ส่วนใหญ่มีเซลล์เดี่ยวและมีโครงสร้างเซลล์ที่ไม่ซับซ้อนมาก ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะส่วนใหญ่แล้วจะพบแบคทีเรียในร่างกาย

10) ปริมาณน้ำตาลในเลือดแบบสุ่ม (Blood Glucose Random) คือ การเจาะเลือดตรวจโดยไม่ต้องงดอาหาร เหมาะสำหรับผู้ที่มีอาการของเบาหวาน เช่น หิวบ่อย น้ำหนักลด ปัสสาวะบ่อย หรือผู้ที่มีอาการเหมือนน้ำตาลในเลือดต่ำ ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะพบปริมาณน้ำตาลในเลือดแบบสุ่มผิดปกติ (ค่าปกติ 80-120 mg/dl)

11) ปริมาณไนโตรเจนในกระแสเลือด (Blood Urea Nitrogen) คือ การตรวจค่า Blood Urea Nitrogen หรือที่เรียกว่า BUN เป็นการวัดปริมาณไนโตรเจนในกระแสเลือด ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะพบปริมาณไนโตรเจนในกระแสเลือดผิดปกติ (ค่าปกติ 10-20 mg/dl)

12) ปริมาณครีเอตินินในเลือด (Serum Creatinine) คือ ค่าครีเอตินิน (Creatinine) เป็นการวัดค่าการทำงานของไต ซึ่งสามารถวัดได้จากซีรัมในเลือด เรียกว่า ซีรัมครีเอตินิน (Serum Creatinine) โดยครีเอตินินเกิดจากการออกแรงยืดหดหรือการใช้กล้ามเนื้อในชีวิตประจำวัน โดยสารครีเอตินินจะเกิดขึ้นทุกวันและมีค่าคงที่ คือมีในปริมาณเท่าๆ กันทุกวัน ในแต่ละคน ซึ่งสารนี้จะถูกขับออกจากร่างกายผ่านไต ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะพบปริมาณครีเอตินินในเลือดอยู่ในเกณฑ์ไม่ปกติ

13) ปริมาณโซเดียมในเลือด (Sodium) คือ ภาวะที่มีเกลือโซเดียมในเลือดน้อยกว่า 135 mEq/L จะทำให้เกิดอาการคลื่นไส้ อาเจียน กล้ามเนื้ออ่อนแรงเป็นตะคริว และยังทำให้เกิดอาการทางสมอง ได้แก่ ซึม ชัก หรือหมดสติ ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะพบปริมาณโซเดียมในเลือดอยู่ในเกณฑ์ปกติ

14) ปริมาณโพแทสเซียมในเลือด (Potassium) คือ ส่วนประกอบที่สำคัญของเซลล์ เนื้อเยื่อ และอวัยวะต่างๆ นอกจากนั้นยังทำหน้าที่ร่วมกับเกลือ โซเดียม คลอไรด์ แคลเซียม แมกนีเซียม ในการนำไฟฟ้า โพแทสเซียมยังเป็นแร่ที่สำคัญในการบีบตัวของกล้ามเนื้อหัวใจและระบบทางเดินอาหาร การที่มีค่าโพแทสเซียมสูงเกินไปหรือต่ำเกินไปอาจจะส่งผลเสียต่อร่างกาย เช่น หัวใจเต้นผิดปกติหรือหยุดเต้น กล้ามเนื้ออ่อนแรง เป็นต้น ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้น่าจะพบปริมาณโพแทสเซียมในเลือดในเกณฑ์ปกติ (ค่าปกติ 10-20 mg/dl)

15) ปริมาณฮีโมโกลบิน (Hemoglobin) ประกอบไปด้วยโมเลกุลของโปรตีน Globin 4 โมเลกุล คือ 2 Alpha-Globulin Chains และ 2 Beta-Globulin Chains และธาตุเหล็กที่เป็นองค์ประกอบของเซลล์เม็ดเลือดแดงมีหน้าที่พาออกซิเจนจากปอดไปสู่ยังเนื้อเยื่อทั่วร่างกายและรับคาร์บอนไดออกไซด์ซึ่งเกิดจากการเผาผลาญอาหาร (เลือดดำ) มาส่งให้ปอดเพื่อทำการพอกให้เป็นเลือดแดง ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะพบปริมาณฮีโมโกลบินในเกณฑ์ปกติ

16) ปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมด (Packed Cell Volume) คือ การวัดว่าร่างกายมีเม็ดเลือดแดงเพียงพอหรือไม่ หากไม่เพียงพอก็เกิดภาวะโลหิตจาง ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วอยู่ในเกณฑ์ปกติ

17) จำนวนเซลล์เม็ดเลือดขาว (White Blood Cell Count) คือ ในเลือดของคนเราประกอบไปด้วยน้ำเหลือง เม็ดเลือดแดง เม็ดเลือดขาว และเกล็ดเลือด แม้จะมีปริมาณเพียงประมาณ 1 เปอร์เซ็นต์ของน้ำเลือด แต่เม็ดเลือดขาวก็มีบทบาทในการต่อต้านโรคร้ายต่างๆ และโรคติดเชื้อ ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะมีจำนวนเซลล์เม็ดเลือดขาวอยู่ในเกณฑ์ปกติ

18) จำนวนเซลล์เม็ดเลือดแดง (Red Blood Cell Count) คือ เม็ดเลือดแดงจะนำออกซิเจนจากปอดไปสู่ร่างกายและนำคาร์บอนไดออกไซด์จากร่างกายไปขับออกที่ปอด จำนวนเม็ดเลือดแดงมีปริมาณน้อยเรียกว่าโลหิตจาง ร่างกายไม่สามารถที่จะนำพาออกซิเจนไปเลี้ยงร่างกายได้เพียงพอ หากเม็ดเลือดแดงมีปริมาณมากเรียกว่า Polycythemia จะมีโอกาสเกิดเม็ดเลือดแดงอุดหลอดเลือดฝอย ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะมีจำนวนเซลล์เม็ดเลือดแดงอยู่ในเกณฑ์ผิดปกติ

19) ภาวะความดันโลหิตสูง (Hypertension) คือ ผู้ป่วยโรคความดันโลหิตสูงมักจะไม่มีอาการอะไร เนื่องจากร่างกายมีการปรับตัว ผู้ป่วยมักจะมาด้วยโรคหัวใจวาย กล้ามเนื้อหัวใจขาดเลือด ไตวาย ตามัว สำหรับผู้ที่ความดันปกติให้วัดความดันโลหิตทุก 2 ปี ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะไม่มีภาวะการเป็นความดันโลหิตสูง

20) เบาหวาน (Diabetes Mellitus) คือ การเป็นโรคเรื้อรังที่รักษาไม่หาย ต้องได้รับการดูแลรักษาตลอดชีวิต ทั้งนี้เกิดจากการที่ในเลือดมีน้ำตาลสูงกว่าปกติ ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะไม่ใช่เบาหวาน

21) โรคหลอดเลือดหัวใจ (Coronary Artery Disease) คือ ภาวะที่หลอดเลือดแดง (Coronary Artery) ที่เลี้ยงหัวใจตีบหรือตัน ทำให้กล้ามเนื้อหัวใจขาดเลือดหรือทำให้กล้ามเนื้อหัวใจตาย ผู้ป่วยจะมีอาการเจ็บหน้าอกและอาจมีภาวะหัวใจวายหรือเสียชีวิตได้ ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วจะไม่ใช่โรคหลอดเลือดหัวใจ

22) ความอยากอาหาร (Appetite) ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วกินอาหารได้ตามปกติ

23) อาการเท้าบวม (Pedal Edema) ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้นับว่าส่วนใหญ่แล้วไม่ได้มีอาการเท้าบวม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

24) โรคโลหิตจาง (Anemia) คือ ภาวะที่ร่างกายหรือเลือดมีเม็ดเลือดแดงต่ำ เม็ดเลือดแดงคือเม็ดเลือดชนิดหนึ่งสร้างจากไขกระดูก เม็ดเลือดแดงประกอบด้วยสารสำคัญมีสีแดงที่เรียกว่าฮีโมโกลบิน (Hemoglobin) ซึ่งประกอบด้วยธาตุ เหล็กและโปรตีน มีหน้าที่สำคัญในการนำออกซิเจนจากปอดไปหล่อเลี้ยงเซลล์ต่างๆ ทั่วร่างกาย เพื่อให้เซลล์นำออกซิเจนไปเผาผลาญพลังงานเพื่อการทำงานต่างๆ ของเซลล์และนำคาร์บอนไดออกไซด์ที่เกิดจากการใช้พลังงานของเซลล์กลับสู่ปอด เพื่อกำจัดออกนอกร่างกาย ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้อยู่ส่วนใหญ่แล้วไม่ได้เป็นโรคโลหิตจาง

ตัวแปรตาม

1) โรคไตเรื้อรัง (Chronic Kidney Disease) คือ ภาวะที่มีการทำลายไตนานกว่า 3 เดือน ส่งผลให้ไตทำงานผิดปกติ ซึ่งทราบได้จากการตรวจเลือด ตรวจปัสสาวะ เอกซเรย์ หรือตรวจพยาธิสภาพของชิ้นเนื้อไต นอกจากนี้ผู้ป่วยโรคไตเรื้อรังมีความเสี่ยงเพิ่มขึ้นที่จะเกิดภาวะหลอดเลือดตีบแข็งทำให้เกิดโรคหัวใจและโรคหลอดเลือดตามมาและเป็นสาเหตุสำคัญที่ทำให้ผู้ป่วยเสียชีวิต ซึ่งข้อมูลผู้ป่วยโรคไตเรื้อรังในครั้งนี้อยู่ส่วนใหญ่แล้วเป็นโรคไตเรื้อรัง มีทั้งหมด 250 ระเบียบ

4.1 ผลการแบ่งข้อมูลด้วยโปรแกรม SPSS

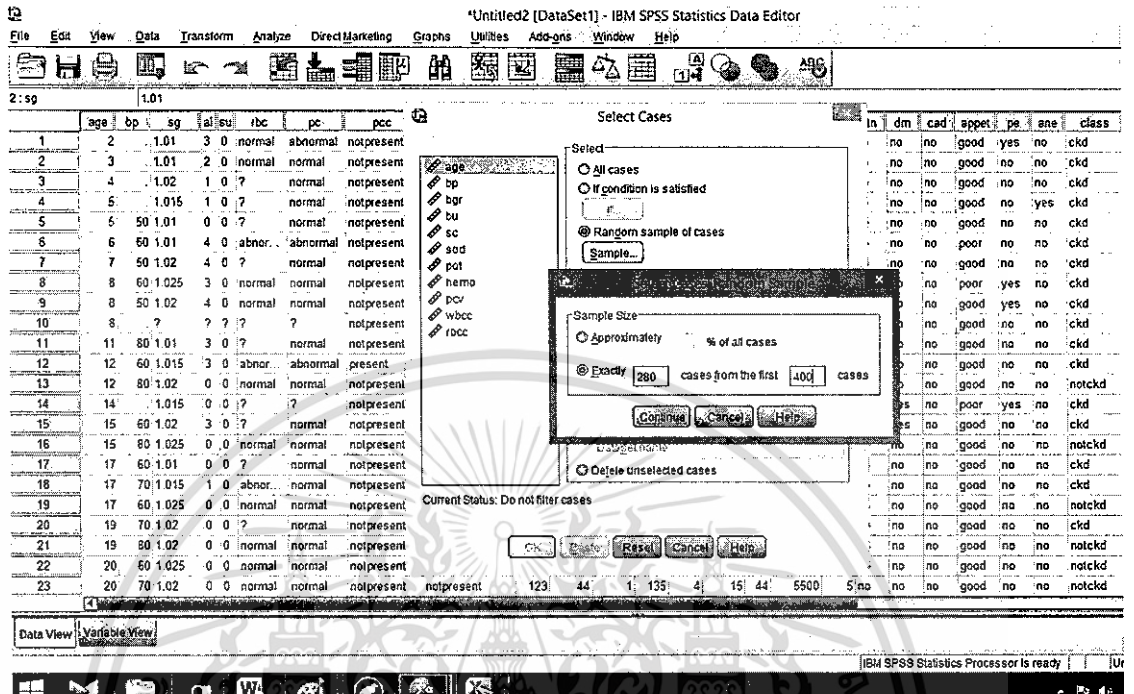
1) เข้าสู่โปรแกรม SPSS โดยสุ่มแบบ Simple Random Sampling (SRS) จากนั้นเลือก File >> open data chronic kidney disease หลังจากนั้นเลือก Data >> Select Cases

	ba	bgr	bu	sc	sod	pot	hemo	pcv	wbcc	rbc	htn	dm	cad	appel	pe	ana	class
1	notpresent										no	no	no	good	yes	no	ckd
2	notpresent		22	1			11	34	12300		no	no	no	good	no	no	ckd
3	notpresent	99	23	1	138	4	12	34			no	no	no	good	no	no	ckd
4	notpresent	16	1	138	3	8					no	no	no	good	no	yes	ckd
5	notpresent		25	1			12	38	12400		no	no	no	good	no	no	ckd
6	present	94	67	1	135	5	10	30	16700	5	no	no	no	poor	no	no	ckd
7	notpresent	18	1				11	38	6000		no	no	no	good	no	no	ckd
8	notpresent	78	27	1			12	41	6700		no	no	no	poor	yes	no	ckd
9	notpresent	46	1	135	4						no	no	no	good	yes	no	ckd
10	notpresent	80	66	3	142	4	12	38			no	no	no	good	no	no	ckd
11	notpresent	17	1				15	45	8600		no	no	no	good	no	no	ckd
12	notpresent		51	2			12		10300		no	no	no	good	no	no	ckd
13	notpresent	100	26	1	137	4	16	49	6600	5	no	no	no	good	no	no	notckd
14	notpresent	192	15	1	137	4	14	40	9500	5	no	yes	no	poor	yes	no	ckd
15	notpresent	85	15	1	138	4	11	33	7700	4	yes	yes	no	good	no	no	ckd
16	notpresent	93	17	1	136	4	17	50	6200	5	no	no	no	good	no	no	notckd
17	notpresent	92	32	2	141	4	14	52	7000		no	no	no	good	no	no	ckd
18	notpresent	22	2	7	145	3	13	41	11200		no	no	no	good	no	no	ckd
19	notpresent	114	50	1	135	5	14	51	7200	6	no	no	no	good	no	no	notckd
20	notpresent						12		6900		no	no	no	good	no	no	ckd
21	notpresent	107	23	1	141	4	14	44			no	no	no	good	no	no	notckd
22	notpresent				137	5	14	41	4500	6	no	no	no	good	no	no	notckd
23	notpresent	123	44	1	135	4	15	44	5500	5	no	no	no	good	no	no	notckd

รูปที่ 4-1 การเลือกคำสั่ง Select Cases

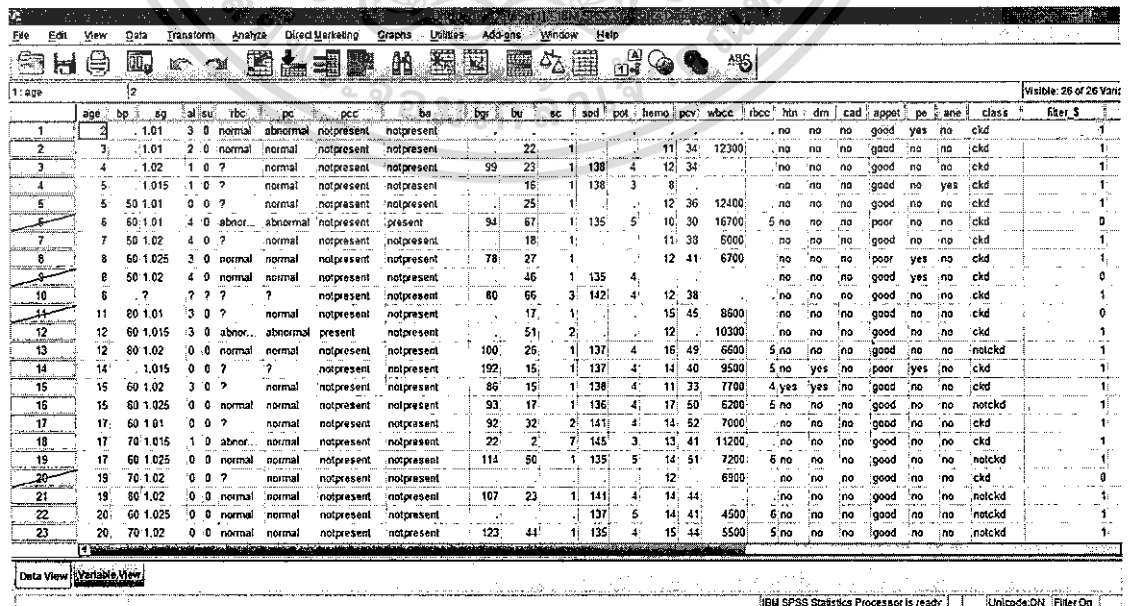
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) เลือก Random sample of cases >> Exactly 280 cases from the first 400 cases >> continue >> OK



รูปที่ 4-2 การเลือกคำสั่ง Random Sample

3) ผลจากการสุ่มตัวอย่าง ซึ่งในช่อง filter_\$ จะมีหมายเลข 1 จำนวน 280 ระเบียบ ตามที่กำหนดไว้ข้างต้น เพื่อนำไปสร้างตัวแบบ ที่เหลือจะเป็นหมายเลข 0 จำนวน 120 ระเบียบ เพื่อนำไปทดสอบตัวแบบ

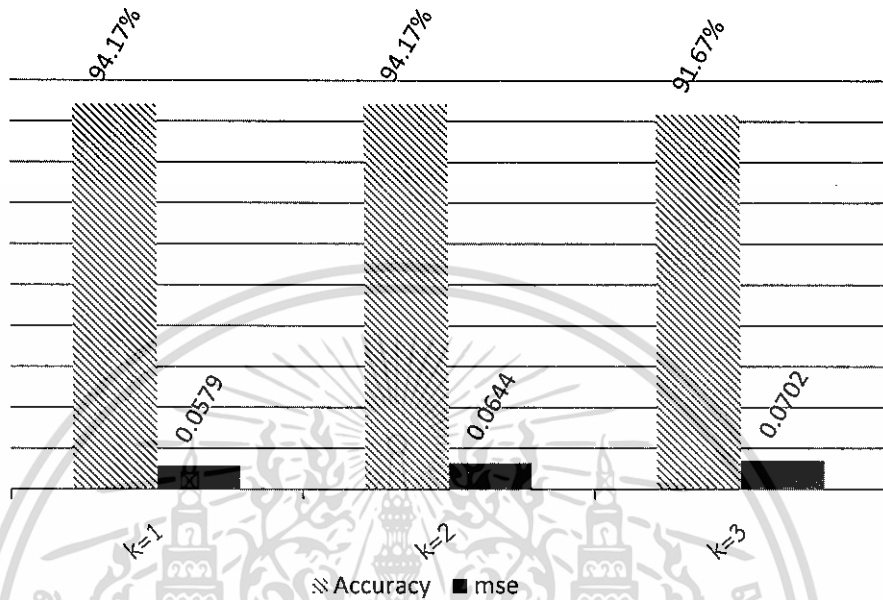


รูปที่ 4-3 ผลจากการสุ่มตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเฉพาะกิจของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี ไม่ควรเผยแพร่โดยไม่ได้รับอนุญาตให้ไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ผลการวิเคราะห์ด้วยวิธีความใกล้เคียงกันมากที่สุด

กำหนดค่า $k = 1, 2, 3$ ตามลำดับ ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-1 ถึง รูปที่ ข-6)



รูปที่ 4-4 ผลการเปรียบเทียบค่า k ของวิธีความใกล้เคียงกันมากที่สุด

จากรูปที่ 4-4 พบว่าค่า $k = 1$ และ $k = 2$ ให้ค่าความถูกต้องมากที่สุดเท่ากันคือ 94.17% แต่ $k = 1$ ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยกว่า $k = 2$ ดังนั้นผู้วิจัยจึงใช้ค่า $k = 1$ มาทำการวิเคราะห์ข้อมูลในครั้งนี้

Classifier

Choose [IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last]"

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 50

(Nom) class

Result list (right-click for options)

17:48:46 - lazy.IBk

17:50:52 - msc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.06 seconds

=== Summary ===

Correctly Classified Instances	113	94.1667 %
Incorrectly Classified Instances	7	5.8333 %
Kappa statistic	0.8726	
Mean absolute error	0.0615	
Root mean squared error	0.2407	
Relative absolute error	13.2131 %	
Root relative squared error	50.4877 %	
Coverage of cases (0.95 level)	94.1667 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.937	0.049	0.974	0.937	0.955	0.874	0.944	0.954	ckd
	0.951	0.063	0.886	0.951	0.918	0.874	0.944	0.860	notckd
Weighted Avg.	0.942	0.054	0.944	0.942	0.942	0.874	0.944	0.922	

=== Confusion Matrix ===

a b <-- classified as

74 5 | a = ckd

2 39 | b = notckd

รูปที่ 4-5 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-1) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-5 พบว่าข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 113 ระเบียบ คิดเป็น 94.17% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 74 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 39 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 7 ระเบียบ คิดเป็น 5.83% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 2 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 5 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.2407)^2 = 0.0579$ ดังรายละเอียดในภาคผนวก ข (ตัวอย่างที่ 1) ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 2, 3, 17, 38, 48, 57 และ 98 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-7) โดยข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังคือระเบียบที่ 38 และ 48 ส่วนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังคือระเบียบที่ 2, 3, 17, 57 และ 98

4.3 ผลการวิเคราะห์ด้วยวิธีต้นไม้ตัดสินใจ

Classifier

Choose J48 - C0.25 - M2

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) class

Result list (right-click for options)

18:12:56 - trees.J48

18:14:20 - misc.input\SkippedClassifier

Classifier output

=== EVALUATION on test set ===

Time taken to test model on supplied test set: 0.02 seconds

=== Summary ===

Correctly Classified Instances	120	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0254		
Root mean squared error	0.0771		
Relative absolute error	5.4554 %		
Root relative squared error	16.1636 %		
Coverage of cases (0.95 level)	100	%	
Mean rel. region size (0.95 level)	52.9167 %		
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	ckd
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	notckd
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

=== Confusion Matrix ===

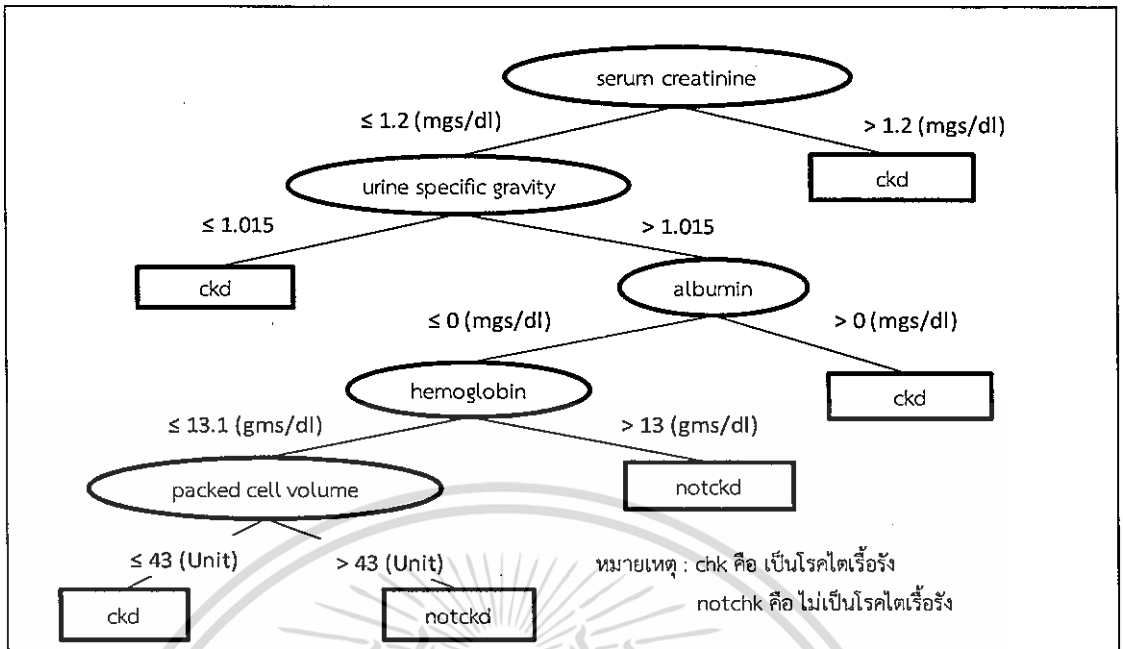
a b <-- classified as

79 0 | a = ckd

0 41 | b = notckd

รูปที่ 4-6 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีต้นไม้ตัดสินใจ

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีต้นไม้ตัดสินใจ ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-8) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-6 พบว่า ข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 120 ระเบียบ คิดเป็น 100% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 79 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 41 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 0 ระเบียบ คิดเป็น 0% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 0 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 0 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.0771)^2 = 0.0059$ ซึ่งไม่มีข้อมูลที่จำแนกผิด ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-10)



รูปที่ 4-7 แบบจำลองของกฎต้นไม้ตัดสินใจเพื่อใช้ในการทำนาย

กฎข้อที่ 1 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) ≤ 1.2 มิลลิกรัมต่อเดซิลิตร และ ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) ≤ 1.015 สถานะ คือ เป็นโรคไตเรื้อรัง

กฎข้อที่ 2 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) ≤ 1.2 มิลลิกรัมต่อเดซิลิตร และ ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) > 1.015 และระดับโปรตีนอัลบูมิน (Albumin) ≤ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบิน (Hemoglobin) ≤ 13.1 กรัมต่อเดซิลิตร และปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมด (Packed Cell Volume) ≤ 43 ยูนิต สถานะ คือ เป็นโรคไตเรื้อรัง

กฎข้อที่ 3 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) ≤ 1.2 มิลลิกรัมต่อเดซิลิตร และ ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) > 1.015 และระดับโปรตีนอัลบูมิน (Albumin) ≤ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบิน (Hemoglobin) ≤ 13.1 กรัมต่อเดซิลิตร และปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมด (Packed Cell Volume) > 43 ยูนิต สถานะ คือ ไม่เป็นโรคไตเรื้อรัง

กฎข้อที่ 4 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) ≤ 1.2 มิลลิกรัมต่อเดซิลิตร และ ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) > 1.015 และระดับโปรตีนอัลบูมิน (Albumin) ≤ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบิน (Hemoglobin) > 13.1 กรัมต่อเดซิลิตร สถานะ คือ ไม่เป็นโรคไตเรื้อรัง

กฎข้อที่ 5 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) ≤ 1.2 มิลลิกรัมต่อเดซิลิตร และ ความถ่วงจำเพาะของปัสสาวะ (Urine Specific Gravity) > 1.015 และระดับโปรตีนอัลบูมิน (Albumin) > 0 มิลลิกรัมต่อเดซิลิตร สถานะ คือ เป็นโรคไตเรื้อรัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กฎข้อที่ 6 : ถ้าปริมาณครีเอตินินในเลือด (Serum Creatinine) > 1.2 มิลลิกรัมต่อเดซิลิตร สถานะ คือ เป็นโรคไตเรื้อรัง

4.4 ผลการวิเคราะห์ด้วยวิธีโครงข่ายประสาทเทียม

Classifier

Choose **MultilayerPerceptron-L0.1-M0.9-N20000-Y0-S0-E20-H4**

Test options

Use training set

Supplied test set

Cross-validation Folds:

Percentage split %:

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:00:50 - results.onSuppliedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	118	98.3333 %
Incorrectly Classified Instances	2	1.6667 %
Kappa statistic	0.9634	
Mean absolute error	0.0138	
Root mean squared error	0.1046	
Relative absolute error	2.9592 %	
Root relative squared error	21.9386 %	
Coverage of cases (0.95 level)	99.1667 %	
Mean rel. region size (0.95 level)	50.8333 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

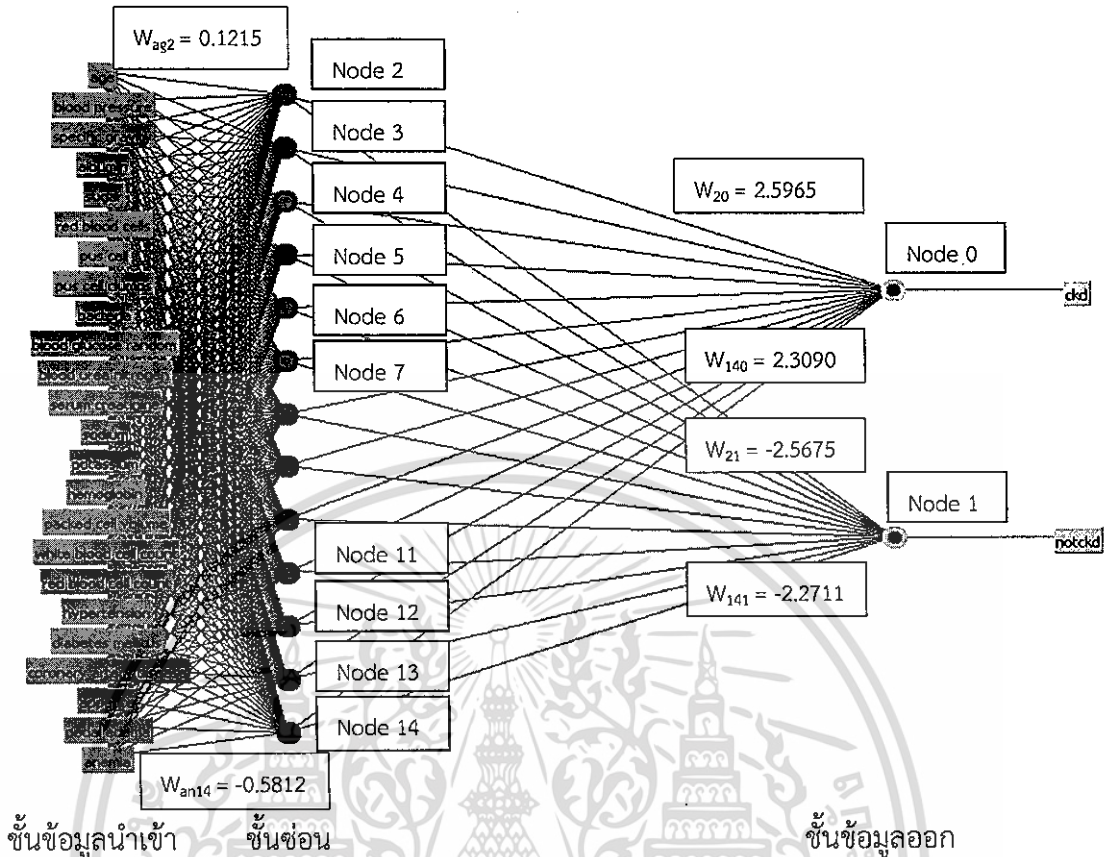
	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	ERC Area	Class
	0.975	0.000	1.000	0.975	0.987	0.964	0.999	1.000	ckd
	1.000	0.025	0.953	1.000	0.976	0.964	0.999	0.998	notckd
Weighted Avg.	0.983	0.009	0.984	0.983	0.983	0.964	0.999	0.999	

=== Confusion Matrix ===

a \ b	-- classified as	
77 a = ckd	2	0
0 b = notckd	41	2

รูปที่ 4-8 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีโครงข่ายประสาทเทียม

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีโครงข่ายประสาทเทียม ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-11) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-8 พบว่าข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 118 ระเบียบ คิดเป็น 98.33% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 77 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 41 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 2 ระเบียบ คิดเป็น 1.67% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 0 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 2 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.1046)^2 = 0.0109$ ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 10 และ 50 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-13) โดยข้อมูลทั้งสองระเบียบจำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรัง



รูปที่ 4-9 ผลค่าถ่วงน้ำหนักบางค่าของโครงข่ายประสาทเทียม

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set Set...

Cross-validation Folds

Percentage split %

More options...

(Nom) Class variable ▼

Start Stop

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 0

Inputs	Weights
Threshold	-3.28914874971958
Node 2	2.5965052166719853
Node 3	4.251135407104274
Node 4	-7.6629311305486265
Node 5	2.7271213494608717
Node 6	2.7920823934499297
Node 7	1.6751056766851142
Node 8	2.005494241165975
Node 9	0.9361906920674167
Node 10	4.211784918610165
Node 11	1.3098749271480157
Node 12	2.33916591698401
Node 13	1.8187883542488679
Node 14	2.30902760769966

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น มิใช่เพื่อเผยแพร่โดยไม่ได้รับอนุญาต
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

- 18:58:45 - functions.MultilayerPerceptron
- 19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 1

Inputs	Weights
Threshold	3.283836603328014
Node 2	-2.5675047500956434
Node 3	-4.274268715858908
Node 4	7.6715284679995746
Node 5	-2.7790007177479557
Node 6	-2.8275073579939125
Node 7	-1.6107253834579398
Node 8	-1.9963533938867026
Node 9	-0.9056654043072142
Node 10	-4.144500197052689
Node 11	-1.3232381290963873
Node 12	-2.357318300285787
Node 13	-1.896847899803941
Node 14	-2.271129569811121

รูปที่ 4-11 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 1

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

- 18:58:45 - functions.MultilayerPerceptron
- 19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 2

Inputs	Weights
Threshold	2.087082383487422
Attrib age	0.12146535545498612
Attrib blood pressure	0.5271286592678875
Attrib specific gravity	-1.28204700078949
Attrib albumin	0.6030013314390629
Attrib sugar	-1.1420991107454657
Attrib red blood cells	1.133277395743126
Attrib pus cell	-0.3830339643578746
Attrib pus cell clumps	0.21871270200028733
Attrib bacteria	-0.31784680973735874
Attrib blood glucose random	0.7115479377643559
Attrib blood urea nitrogen	-0.44663323931762827
Attrib serum creatinine	-0.3066853702925956
Attrib sodium	-0.3212937905069914
Attrib potassium	-0.08439056343231466
Attrib hemoglobin	-1.2995632772614105
Attrib packed cell volume	-0.5265688095871468
Attrib white blood cell count	0.49047738355941906
Attrib red blood cell count	-0.6778432213459796
Attrib hypertension	0.9062047466089669
Attrib diabetes mellitus	0.4457591444475076
Attrib coronary artery disease	-0.042331771067689025
Attrib appetite	0.7189321077386726
Attrib pedal edema	-0.5791499113666924
Attrib anemia	-0.5907194871423561

รูปที่ 4-12 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 3

Inputs	Weights
Threshold	2.2357941751752817
Attrib age	0.024174708494738108
Attrib blood pressure	1.25350209865298
Attrib specific gravity	-3.2810607416504913
Attrib albumin	1.0355851067127795
Attrib sugar	-2.1516259963700444
Attrib red blood cells	2.90257845902176
Attrib pus cell	1.272925173568244
Attrib pus cell clumps	-0.33590778941297067
Attrib bacteria	-0.6546979919329616
Attrib blood glucose random	0.8777986136992523
Attrib blood urea nitrogen	-0.39476878367055646
Attrib serum creatinine	0.2519902294949971
Attrib sodium	-0.7751288089818963
Attrib potassium	-0.279631457730251
Attrib hemoglobin	-1.9598567196236811
Attrib packed cell volume	-0.6985257501978984
Attrib white blood cell count	0.44147460960636203
Attrib red blood cell count	-0.8341053868702751
Attrib hypertension	1.3740716119452068
Attrib diabetes mellitus	0.963192354187308
Attrib coronary artery disease	0.24738354070140947
Attrib appetite	0.582563287925604
Attrib pedal edema	0.2500354940297729
Attrib anemia	-1.0684371179513348

รูปที่ 4-13 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 3

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 4

Inputs	Weights
Threshold	-3.4004732250194496
Attrib age	0.35249721039172566
Attrib blood pressure	-1.6173326301087922
Attrib specific gravity	4.703756947026818
Attrib albumin	-1.3776790396668583
Attrib sugar	3.4754699784895555
Attrib red blood cells	-3.1366783802111926
Attrib pus cell	0.5010600587610108
Attrib pus cell clumps	-1.1119946929943723
Attrib bacteria	0.6321793217775933
Attrib blood glucose random	-1.0752190199900473
Attrib blood urea nitrogen	0.3147005960475743
Attrib serum creatinine	-0.5732453046047269
Attrib sodium	0.8794212398902551
Attrib potassium	-0.6256051887112539
Attrib hemoglobin	2.9244954228745677
Attrib packed cell volume	1.0746180854749443
Attrib white blood cell count	-1.0549411610154489
Attrib red blood cell count	1.0619054628298665
Attrib hypertension	0.39721060175828204
Attrib diabetes mellitus	0.9454515201431763
Attrib coronary artery disease	1.9717144208245152
Attrib appetite	-1.1944989024680897
Attrib pedal edema	1.615387296210873
Attrib anemia	0.992344250209157

รูปที่ 4-14 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 5

Inputs	Weights
Threshold	1.9937798852745907
Attrib age	0.1790413157130218
Attrib blood pressure	0.6185444771956786
Attrib specific gravity	-1.7013182672389249
Attrib albumin	0.5140804744443243
Attrib sugar	-1.319164238587694
Attrib red blood cells	1.5656194189784682
Attrib pus cell	0.23979162634325313
Attrib pus cell clumps	-0.4396809188481631
Attrib bacteria	-0.6158334260365729
Attrib blood glucose random	0.6487806816212306
Attrib blood urea nitrogen	-0.2850720165441862
Attrib serum creatinine	-0.023101550230759268
Attrib sodium	-0.2976663116933
Attrib potassium	-0.07228754418030882
Attrib hemoglobin	-1.386288976106711
Attrib packed cell volume	-0.5290540716040767
Attrib white blood cell count	0.26481142973407884
Attrib red blood cell count	-0.5976080107843669
Attrib hypertension	0.9414294597104534
Attrib diabetes mellitus	0.5835266294115972
Attrib coronary artery disease	-0.06377154373780364
Attrib appetite	0.9432273892061401
Attrib pedal edema	-0.7979602049178338
Attrib anemia	-0.5441523646654081

รูปที่ 4-15 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 5

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 6

Inputs	Weights
Threshold	2.248340556726532
Attrib age	-0.07883559855036386
Attrib blood pressure	0.4423101340052829
Attrib specific gravity	-1.2551272711157144
Attrib albumin	0.8070346483173535
Attrib sugar	-1.0468290293936766
Attrib red blood cells	0.8362419845501791
Attrib pus cell	-0.8501102315675967
Attrib pus cell clumps	0.36599204521555034
Attrib bacteria	-0.19610750345255315
Attrib blood glucose random	0.7176222426448127
Attrib blood urea nitrogen	-0.30970331254521055
Attrib serum creatinine	-0.5205404513306532
Attrib sodium	-0.2590498635715849
Attrib potassium	0.10335363614121842
Attrib hemoglobin	-1.5198967284881275
Attrib packed cell volume	-0.7096418175749383
Attrib white blood cell count	0.6729132739338182
Attrib red blood cell count	-0.7341461163679177
Attrib hypertension	0.7322842561140266
Attrib diabetes mellitus	0.35426712549356365
Attrib coronary artery disease	-0.12767080943939818
Attrib appetite	0.6857140839101172
Attrib pedal edema	-0.7218676421221447
Attrib anemia	-0.7605089769902826

รูปที่ 4-16 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds **10**

Percentage split % **66**

More options...

(Nom) Class variable **v**

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 7

Inputs	Weights
Threshold	1.0929701279368393
Attrib age	0.05032592955448207
Attrib blood pressure	0.1413167435426155
Attrib specific gravity	-0.4386952997680658
Attrib albumin	0.2778662405408795
Attrib sugar	-0.39500176188056335
Attrib red blood cells	0.40468201599610754
Attrib pus cell	-0.7959926464005687
Attrib pus cell clumps	0.43488211983519753
Attrib bacteria	0.0504181734746205
Attrib blood glucose random	0.43552392209152574
Attrib blood urea nitrogen	-0.1941394926671878
Attrib serum creatinine	-0.16462418339246773
Attrib sodium	-0.06110698244031043
Attrib potassium	0.19143734312432278
Attrib hemoglobin	-0.7541092672865483
Attrib packed cell volume	-0.45932770652998656
Attrib white blood cell count	0.4098840015159492
Attrib red blood cell count	-0.42696596550249777
Attrib hypertension	0.39608996990748707
Attrib diabetes mellitus	0.2534318903263691
Attrib coronary artery disease	-0.04138715338444066
Attrib appetite	0.42212083391429284
Attrib pedal edema	-0.5672920138488818
Attrib anemia	-0.21567965118575452

รูปที่ 4-17 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 7

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds **10**

Percentage split % **66**

More options...

(Nom) Class variable **v**

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 8

Inputs	Weights
Threshold	1.440899839317434
Attrib age	0.13465110208842768
Attrib blood pressure	0.2532461651246057
Attrib specific gravity	-0.7645164543508676
Attrib albumin	0.3674197077859333
Attrib sugar	-0.5064194285926309
Attrib red blood cells	0.5901526172441545
Attrib pus cell	-0.7598220267943753
Attrib pus cell clumps	0.38184908135625634
Attrib bacteria	0.02237186410275924
Attrib blood glucose random	0.5484892861121179
Attrib blood urea nitrogen	-0.2011696711947669
Attrib serum creatinine	-0.31835718843046107
Attrib sodium	-0.164682577251509636
Attrib potassium	0.131217349110253
Attrib hemoglobin	-1.0800650712849778
Attrib packed cell volume	-0.5624658505810016
Attrib white blood cell count	0.5554781572700043
Attrib red blood cell count	-0.562177695514353
Attrib hypertension	0.6133357728254188
Attrib diabetes mellitus	0.35069643058869604
Attrib coronary artery disease	0.06209426671457267
Attrib appetite	0.3557716807013535
Attrib pedal edema	-0.37874771236772586
Attrib anemia	-0.5164007638758498

รูปที่ 4-18 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 9

Inputs	Weights
Threshold	0.3413610964973998
Attrib age	0.061927782455956405
Attrib blood pressure	0.14159621874154432
Attrib specific gravity	-0.2588504418669379
Attrib albumin	0.09181693651003833
Attrib sugar	0.004057750725393055
Attrib red blood cells	0.21798671993203259
Attrib pus cell	-0.4496264902794778
Attrib pus cell clumps	0.34124909242698714
Attrib bacteria	0.22603558854909198
Attrib blood glucose random	0.3137498893673126
Attrib serum urea nitrogen	0.04000719943391736
Attrib serum creatinine	0.03981852046656999
Attrib sodium	0.018273658150119654
Attrib potassium	0.1977375478695382
Attrib hemoglobin	-0.3579264405080019
Attrib packed cell volume	-0.22836066682305287
Attrib white blood cell count	0.2513497476143087
Attrib red blood cell count	-0.1387163924503722
Attrib hypertension	0.14656850561190307
Attrib diabetes mellitus	0.2070384856981989
Attrib coronary artery disease	0.14971564294183354
Attrib appetite	0.30008814515445587
Attrib pedal edema	-0.3125085575419978
Attrib anemia	-0.07632048036086371

รูปที่ 4-19 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 9

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 10

Inputs	Weights
Threshold	2.56623951591146
Attrib age	0.03062048901450421
Attrib blood pressure	1.0799024239274284
Attrib specific gravity	-3.0519806507014717
Attrib albumin	1.0680918328564375
Attrib sugar	-2.291441935554884
Attrib red blood cells	2.651822155014778
Attrib pus cell	0.9937952656240904
Attrib pus cell clumps	-0.6739161441858277
Attrib bacteria	-0.9443621756495761
Attrib blood glucose random	0.8144037264702337
Attrib blood urea nitrogen	-0.38198534726177474
Attrib serum creatinine	0.057175006582447584
Attrib sodium	-0.6449583067585115
Attrib potassium	-0.14530446518699833
Attrib hemoglobin	-1.9465536188247707
Attrib packed cell volume	-0.6640712841972651
Attrib white blood cell count	0.3220299558718019
Attrib red blood cell count	-0.8575677527195065
Attrib hypertension	1.3217340018001462
Attrib diabetes mellitus	0.8344057230417502
Attrib coronary artery disease	0.07107754562992667
Attrib appetite	1.1676411695792075
Attrib pedal edema	-0.6714387349853135
Attrib anemia	-0.9945603672805495

รูปที่ 4-20 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 11

Inputs	Weights
Threshold	0.7267840847277557
Attrib age	0.05199899721368518
Attrib blood pressure	0.1290829868087467
Attrib specific gravity	-0.3505313510244884
Attrib albumin	0.18345255042670014
Attrib sugar	-0.14807899183803822
Attrib red blood cells	0.21290296988529306
Attrib pus cell	-0.7623519983911121
Attrib pus cell clumps	0.4033409603639983
Attrib bacteria	0.20380667852550563
Attrib blood glucose random	0.3237593885532226
Attrib blood urea nitrogen	-0.08309913691054033
Attrib serum creatinine	-0.05924481099120779
Attrib sodium	-0.05231140369591474
Attrib potassium	0.21169163410234182
Attrib hemoglobin	-0.5878972141632111
Attrib packed cell volume	-0.38333794279995326
Attrib white blood cell count	0.34574622203229904
Attrib red blood cell count	-0.2622825751439709
Attrib hypertension	0.30906702404633835
Attrib diabetes mellitus	0.15851340052591445
Attrib coronary artery disease	0.01729713928643241
Attrib appetite	0.33928713574526026
Attrib pedal edema	-0.4408989393183272
Attrib anemia	-0.13526966212294328

รูปที่ 4-21 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 11

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 12

Inputs	Weights
Threshold	1.8754558496068447
Attrib age	0.19612451142859666
Attrib blood pressure	0.4364274396980125
Attrib specific gravity	-1.2342256959857663
Attrib albumin	0.3451770196601063
Attrib sugar	-0.9189348590065809
Attrib red blood cells	0.993103053592132
Attrib pus cell	-0.6082245667921679
Attrib pus cell clumps	0.29695651948162505
Attrib bacteria	-0.1731187125428928
Attrib blood glucose random	0.6649530620660279
Attrib blood urea nitrogen	-0.3247665896032145
Attrib serum creatinine	-0.23420342689278947
Attrib sodium	-0.20321934854624824
Attrib potassium	0.10960966356636492
Attrib hemoglobin	-1.2251923094194561
Attrib packed cell volume	-0.616993137266547
Attrib white blood cell count	0.5166569556547855
Attrib red blood cell count	-0.6520524294475029
Attrib hypertension	0.7293040918035273
Attrib diabetes mellitus	0.4023809690678955
Attrib coronary artery disease	-0.06801772421244823
Attrib appetite	0.4295938356083535
Attrib pedal edema	-0.5103526794141803
Attrib anemia	-0.6115916776172126

รูปที่ 4-22 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 12

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds **10**

Percentage split % **66**

More options...

(Nom) Class variable **v**

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 13

Inputs	Weights
Threshold	1.34683568431244
Attrib age	0.12710432517122675
Attrib blood pressure	0.2994018126930319
Attrib specific gravity	-0.729003496391316
Attrib albumin	0.2929508952209519
Attrib sugar	-0.5151860194126641
Attrib red blood cells	0.7697612515675308
Attrib pus cell	-0.3801547715301979
Attrib pus cell clumps	0.09534743823488334
Attrib bacteria	-0.13454760003545135
Attrib blood glucose random	0.5604780719386186
Attrib blood urea nitrogen	-0.21166898317461696
Attrib serum creatinine	-0.2634162524189905
Attrib sodium	-0.1762803067598469
Attrib potassium	0.05328343064664389
Attrib hemoglobin	-0.9644937238507831
Attrib packed cell volume	-0.45189187900667066
Attrib white blood cell count	0.3684695926958752
Attrib red blood cell count	-0.41151740511564694
Attrib hypertension	0.7078724552346147
Attrib diabetes mellitus	0.5019259520789198
Attrib coronary artery disease	0.05948599020534105
Attrib appetite	0.4916464276003738
Attrib pedal edema	-0.3692982337183128
Attrib anemia	-0.3758687147067908

รูปที่ 4-23 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 13

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a -R**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds **10**

Percentage split % **66**

More options...

(Nom) Class variable **v**

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:17:51 - misc.InputMappedClassifier

Classifier output

Sigmoid Node 14

Inputs	Weights
Threshold	1.8183218577868123
Attrib age	0.24454304351891984
Attrib blood pressure	0.40444439532171217
Attrib specific gravity	-1.1016339336002428
Attrib albumin	0.31386096160723065
Attrib sugar	-0.7381811154414655
Attrib red blood cells	0.8722251446370741
Attrib pus cell	-0.6548801268163934
Attrib pus cell clumps	0.27808711753858295
Attrib bacteria	-0.129097391828911162
Attrib blood glucose random	0.6906666017239603
Attrib blood urea nitrogen	-0.3203406082245148
Attrib serum creatinine	-0.2717373444196398
Attrib sodium	-0.1705142401114674
Attrib potassium	-0.002636005418905875
Attrib hemoglobin	-1.2037162012408804
Attrib packed cell volume	-0.5786454509512635
Attrib white blood cell count	0.5803744606401705
Attrib red blood cell count	-0.5740277530318536
Attrib hypertension	0.7719610773136806
Attrib diabetes mellitus	0.3893759642194622
Attrib coronary artery disease	-0.0240134782800974
Attrib appetite	0.5219540901541311
Attrib pedal edema	-0.45045241279927717
Attrib anemia	-0.5811901471701685

รูปที่ 4-24 ผลค่าถ่วงน้ำหนักของโครงข่ายประสาทเทียม โหนดที่ 14

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 ผลการวิเคราะห์ด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน

The screenshot shows the Weka Classifier interface. The 'Classifier' dropdown is set to 'SMD - C 1.0 - L 0.001 - P 1.0E-12 - N 0 - V - 1 - W 1 - K weka.classifiers.functions.supportVector.PolyKernel - E 1.0 - C 250007'. Under 'Test options', 'Supplied test set' is selected with 'Set...' button, 'Cross-validation' is set to 10 folds, and 'Percentage split' is 66%. The 'Classifier output' pane shows the following results:

```

==== Evaluation on test set ====
Time taken to test model on supplied test set: 0.03 seconds

==== Summary ====
Correctly Classified Instances      118      98.3333 %
Incorrectly Classified Instances    2        1.6667 %
Kappa statistic                    0.9634
Mean absolute error                 0.0167
Root mean squared error             0.1291
Relative absolute error             3.5828 %
Root relative squared error        27.0801 %
Coverage of cases (0.95 level)     98.3333 %
Mean rel. region size (0.95 level) 50
Total Number of Instances          120

==== Detailed Accuracy By Class ====
              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC   ROC Area  PRC Area  Class
0.975   0.000   1.000   0.975   0.987   0.964   0.987   0.991   ckd
1.000   0.025   0.953   1.000   0.976   0.964   0.987   0.953   notckd
Weighted Avg.   0.983   0.009   0.984   0.983   0.983   0.964   0.987   0.978

==== Confusion Matrix ====
a b  <-- classified as
77 2 | a = ckd
0 41| b = notckd
    
```

รูปที่ 4-25 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-14) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-25 พบว่าข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 118 ระเบียบ คิดเป็น 98.33% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 77 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 41 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 2 ระเบียบ คิดเป็น 1.67% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 0 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 2 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.1291)^2 = 0.0167$ ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 4 และ 22 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-16) โดยข้อมูลทั้งสองจำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรัง

Classifier

Choose SMO - C 1.0 - I 0.001 - P 1.0E-12 - N 0 - V -1 - W 1 - K "weka.classifiers.functions.supportVector.PolyKernel-E 1.0 - C 250007"

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) Class variable

Result list (right-click for options)

12:16:30 - Functions.SMO

12:17:12 - Misc:Input Mapped Classifier

Classifier output

Machine linear: showing attribute weights, not support vectors.

```

0.0746 * (normalized) age
+ -0.6325 * (normalized) blood pressure
+ 2.8029 * (normalized) specific gravity
+ -1.789 * (normalized) albumin
+ -0.2 * (normalized) sugar
+ -0.7031 * (normalized) red blood cells
+ 0.3968 * (normalized) pus cell
+ -0.2929 * (normalized) pus cell clumps
+ -0.4798 * (normalized) blood glucose random
+ -0.0959 * (normalized) blood urea nitrogen
+ -0.5512 * (normalized) serum creatinine
+ 0.3234 * (normalized) sodium
+ 0.0202 * (normalized) potassium
+ 1.6135 * (normalized) hemoglobin
+ 1.3857 * (normalized) packed cell volume
+ -0.1804 * (normalized) white blood cell count
+ 0.6794 * (normalized) red blood cell count
+ -1 * (normalized) hypertension
+ -0.8047 * (normalized) diabetes mellitus
+ 0 * (normalized) coronary artery disease
+ -1.0034 * (normalized) appetite
+ 0.3003 * (normalized) pedal edema
+ -0.1039 * (normalized) anemia
- 4.3185

```

รูปที่ 4-26 ผลสมการซัพพอร์ตเวกเตอร์แมชชีนของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ

จากรูปที่ 4-26 สามารถเขียนสมการทำนายได้ดังนี้

Machine Linear : $-4.3185 + 0.0746 * (\text{normalized}) \text{ age} - 0.6325 * (\text{normalized}) \text{ blood pressure} + 2.8029 * (\text{normalized}) \text{ specific gravity} - 1.789 * (\text{normalized}) \text{ albumin} - 0.2 * (\text{normalized}) \text{ sugar} - 0.7031 * (\text{normalized}) \text{ red blood cells} + 0.3968 * (\text{normalized}) \text{ pus cell} - 0.2929 * (\text{normalized}) \text{ pus cell clumps} - 0.4798 * (\text{normalized}) \text{ blood glucose random} - 0.0959 * (\text{normalized}) \text{ blood urea} - 0.5512 * (\text{normalized}) \text{ serum creatinine} + 0.3234 * (\text{normalized}) \text{ sodium} + 0.0202 * (\text{normalized}) \text{ potassium} + 0.1635 * (\text{normalized}) \text{ hemoglobin} + 1.3857 * (\text{normalized}) \text{ packed cell volume} - 0.1804 * (\text{normalized}) \text{ white blood cell count} + 0.6794 * (\text{normalized}) \text{ red blood cell count} - 1 * (\text{normalized}) \text{ hypertension} - 0.8047 * (\text{normalized}) \text{ diabetes mellitus} - 1.0034 * (\text{normalized}) \text{ appetite} + 0.3003 * (\text{normalized}) \text{ pedal edema} - 0.1039 * (\text{normalized}) \text{ anemia}$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6 ผลการวิเคราะห์ด้วยวิธีฐานกฎ

Classifier

Choose **DecisionTable -X1-5-weka.attributeSelection.BestFirst-D1-N5**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

19:02:58 - rules.DecisionTable

19:14:07 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	114	95 %
Incorrectly Classified Instances	6	5 %
Kappa statistic	0.8889	
Mean absolute error	0.1192	
Root mean squared error	0.213	
Relative absolute error	25.6196 %	
Root relative squared error	44.6786 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	67.0833 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.962	0.073	0.962	0.962	0.962	0.889	0.987	0.990	ckd
	0.927	0.038	0.927	0.927	0.927	0.889	0.987	0.973	notckd
Weighted Avg.	0.950	0.061	0.950	0.950	0.950	0.889	0.987	0.984	

=== Confusion Matrix ===

```

a b <-- classified as
76 3 | a = ckd
3 38 | b = notckd

```

รูปที่ 4-27 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีฐานกฎ

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีฐานกฎ ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-17) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-27 พบว่าข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 114 ระเบียบ คิดเป็น 95% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 76 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 38 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 6 ระเบียบ คิดเป็น 5% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 3 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 3 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.213)^2 = 0.0454$ ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 2, 10, 26, 38, 48 และ 51 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-19) โดยข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังคือระเบียบที่ 26, 38 และ 48 ส่วนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังคือระเบียบที่ 2, 10 และ 51

4.7 ผลการวิเคราะห์ด้วยวิธีการถดถอยลอจิสติก

Classifier

Choose Logistic-R 1.0E-8-M-1

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) Class variable

Result list (right-click for options)

19:24:09 - functions.Logistic

19:25:15 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.02 seconds

=== Summary ===

Correctly Classified Instances	112	93.3333 %
Incorrectly Classified Instances	8	6.6667 %
Kappa statistic	0.8535	
Mean absolute error	0.0689	
Root mean squared error	0.257	
Relative absolute error	14.8155 %	
Root relative squared error	53.9184 %	
Coverage of cases (0.95 level)	94.1667 %	
Mean rel. region size (0.95 level)	50.8333 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.937	0.073	0.961	0.937	0.949	0.854	0.945	0.970	ckd
	0.927	0.063	0.884	0.927	0.905	0.854	0.963	0.878	notckd
Weighted Avg.	0.933	0.070	0.935	0.933	0.934	0.854	0.951	0.939	

=== Confusion Matrix ===

```

a b  <-- classified as
74  5 | a = ckd
 3 38 | b = notckd

```

รูปที่ 4-28 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีการถดถอยลอจิสติก

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีการถดถอยลอจิสติก ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-20) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-28 พบว่าข้อมูล 120 ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 112 ระเบียบ คิดเป็น 93.33% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 74 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 38 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 8 ระเบียบ คิดเป็น 6.67% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 3 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 5 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.257)^2 = 0.066$ ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 11, 19, 22, 38, 39, 48, 90 และ 99 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-22) โดยข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังคือระเบียบที่ 19, 38 และ 48 ส่วนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังคือระเบียบที่ 11, 22, 39, 90 และ 99

Choose Logistic -R 1.0E-8 -M-1

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable v
 Start Stop

Result list (right-click for options)
 18: 17:45 - functions.Logistic
 18: 17:50 - misc.lrp/MapClassifier

Classifier output		Classifier output	
Logistic Regression with ridge parameter of 1.0E-8		Odds Ratios...	
Coefficients...			
Variable	Class	Variable	Class
	ckd		ckd
age	0.1432	age	1.154
blood pressure	0.578	blood pressure	1.7824
specific gravity	-4045.661	specific gravity	0
albumin	48.5437	albumin	1.2085498420489953E21
sugar	5.9432	sugar	381.1692
red blood cells	28.7858	red blood cells	3.1734419222180166E12
pus cell	-44.2879	pus cell	0
pus cell clumps	12.5021	pus cell clumps	268905.8664
bacteria	-11.0782	bacteria	0
blood glucose random	0.267	blood glucose random	1.3061
blood urea nitrogen	0.0071	blood urea nitrogen	1.0071
serum creatinine	11.7729	serum creatinine	129690.9316
sodium	0.5626	sodium	1.7552
potassium	-10.3519	potassium	0
hemoglobin	-22.0174	hemoglobin	0
packed cell volume	2.228	packed cell volume	9.2814
white blood cell count	-0.0002	white blood cell count	0.9998
red blood cell count	-2.1271	red blood cell count	0.1192
hypertension	-11.2721	hypertension	0
diabetes mellitus	86.3947	diabetes mellitus	3.3170726504355596E37
coronary artery disease	-51.0916	coronary artery disease	0
appetite	82.5477	appetite	7.079345012592387E35
pedal edema	108.7582	pedal edema	1.7103276837055434E47
anemia	-36.3689	anemia	0
Intercept	4122.3792		

รูปที่ 4-29 ผลสมการถดถอยลอจิสติกของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ

จากรูปที่ 4-29 สามารถเขียนสมการทำนายได้ดังนี้

$$W = 4122.38 + 0.1432(\text{age}) + 0.578(\text{blood pressure}) - 4045.66(\text{specific gravity}) + 48.54(\text{albumin}) + 5.94(\text{sugar}) + 28.79(\text{red blood cells}) - 44.29(\text{pus cell}) + 12.5(\text{pus cell clumps}) - 11.08(\text{bacteria}) + 0.27(\text{blood glucose random}) + 0.01(\text{blood urea}) + 11.77(\text{serum creatinine}) + 0.56(\text{sodium}) - 10.35(\text{potassium}) - 22.02(\text{hemoglobin}) + 2.23(\text{packed cell volume}) - 0.0002(\text{white blood cell count}) - 2.13(\text{red blood cell count}) - 11.27(\text{hypertension}) + 86.39(\text{diabetes mellitus}) - 51.09(\text{coronary artery disease}) + 82.55(\text{appetite}) + 108.76(\text{pedal edema}) - 36.37(\text{anemia})$$

และสามารถแสดงค่าประมาณของอัตราส่วนออด (odds ratio) ได้ดังตัวอย่าง

$$\text{ตัวอย่างที่ 1 } \hat{OR} = e^{b_1} = e^{0.1432} = 1.154$$

หมายความว่า ถ้าอายุเพิ่มขึ้น 1 ปี จะทำให้ odds ratio เพิ่มขึ้น หรือโอกาสที่จะเป็นโรคไตเรื้อรังเพิ่มขึ้น

$$\text{ตัวอย่างที่ 2 } \hat{OR} = e^{b_{17}} = e^{-0.0002} = 0.9998 \approx 1$$

หมายความว่า ถ้าจำนวนเซลล์เม็ดเลือดขาวเพิ่มขึ้น 1 เซลล์ เทบจะไม่ทำให้ odds ratio เพิ่มขึ้นหรือลดลง หรือโอกาสที่จะเป็นโรคไตเรื้อรังเท่าเดิม

$$\text{ตัวอย่างที่ 3 } \hat{OR} = e^{b_{18}} = e^{-2.1271} = 0.1192$$

หมายความว่า ถ้าจำนวนเซลล์เม็ดเลือดแดงเพิ่มขึ้น 1 เซลล์ จะทำให้ odds ratio ลดลง หรือโอกาสที่จะเป็นโรคไตเรื้อรังลดลง

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการเรียนการสอนเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.8 ผลการวิเคราะห์ด้วยวิธีนาอิว เบย์

Classifier

Choose NaiveBayes

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 55

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

19:36:14 - bayes.NaiveBayes

19:37:58 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	114	95	%
Incorrectly Classified Instances	6	5	%
Kappa statistic	0.8926		
Mean absolute error	0.049		
Root mean squared error	0.2023		
Relative absolute error	10.5249 %		
Root relative squared error	42.4244 %		
Coverage of cases (0.95 level)	97.5 %		
Mean rel. region size (0.95 level)	52.0833 %		
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.924	0.000	1.000	0.924	0.961	0.898	1.000	1.000	ckd
	1.000	0.076	0.872	1.000	0.932	0.898	1.000	1.000	notckd
Weighted Avg.	0.950	0.026	0.956	0.950	0.951	0.898	1.000	1.000	

=== Confusion Matrix ===

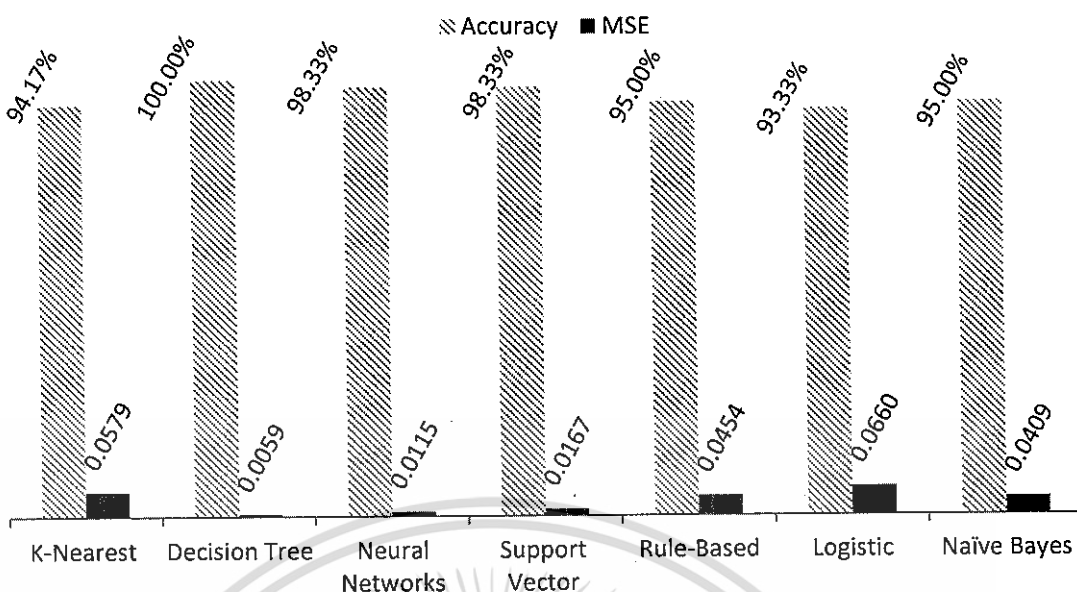
a b <-- classified as

73 6 | a = ckd

0 41 | b = notckd

รูปที่ 4-30 ผลจากการเปรียบเทียบประสิทธิภาพของข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบด้วยวิธีนาอิว เบย์

จากข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีนาอิว เบย์ ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-23) จากนั้นนำข้อมูลที่ได้จากการสร้างตัวแบบมาทำการทดสอบตัวแบบ จากรูปที่ 4-30 พบว่าข้อมูล 120 ระเบียบ ตัวแบบสามารถทำนายข้อมูลได้ถูกต้อง 114 ระเบียบ คิดเป็น 95% โดยมีจำนวนข้อมูลที่จำแนกถูกว่าเป็นโรคไตเรื้อรังมี 73 ระเบียบ และจำนวนข้อมูลที่จำแนกถูกว่าไม่เป็นโรคไตเรื้อรังมี 41 ระเบียบ ตัวแบบทำนายข้อมูลไม่ถูกต้องมี 6 ระเบียบ คิดเป็น 5% โดยมีจำนวนข้อมูลที่จำแนกผิดว่าเป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วไม่ได้เป็นโรคไตเรื้อรังมี 0 ระเบียบ และจำนวนข้อมูลที่จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรังมี 6 ระเบียบ โดยความคลาดเคลื่อนกำลังสองเฉลี่ยมีค่าเท่ากับ $(0.2023)^2 = 0.0409$ ซึ่งข้อมูลที่จำแนกผิดมีดังนี้คือระเบียบที่ 4, 17, 22, 40, 54 และ 64 ดังรายละเอียดในภาคผนวก ข (รูปที่ ข-25) โดยข้อมูลทั้ง 6 ระเบียบ จำแนกผิดว่าไม่เป็นโรคไตเรื้อรัง ซึ่งแท้จริงแล้วเป็นโรคไตเรื้อรัง



รูปที่ 4-31 การเปรียบเทียบค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการจำแนกกลุ่มข้อมูลผู้ป่วยโรคไตเรื้อรัง

ตารางที่ 4-1 การเปรียบเทียบค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของวิธีการจำแนกกลุ่มผู้ป่วยโรคไตเรื้อรัง

วิธีการจำแนกกลุ่ม	ค่าความถูกต้อง	ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย
วิธีความใกล้เคียงกันมากที่สุด	94.17%	0.0579
วิธีต้นไม้ตัดสินใจ	100.00%	0.0059
วิธีโครงข่ายประสาทเทียม	98.33%	0.0109
วิธีซัพพอร์ตเวกเตอร์แมชชีน	98.33%	0.0167
วิธีฐานกฎ	95.00%	0.0454
วิธีถดถอยลอจิสติก	93.33%	0.0660
วิธีนาอิว เบย์	95.00%	0.0409

จากตารางที่ 4-1 เมื่อพิจารณาจากค่าความถูกต้องและค่าความคลาดเคลื่อนกำลังสองเฉลี่ยประกอบกันพบว่า วิธีต้นไม้ตัดสินใจมีค่าความถูกต้องมากที่สุดคือ 100% และยังมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยน้อยที่สุดคือ 0.0059 จึงสรุปว่าวิธีการจำแนกกลุ่มที่ดีที่สุดที่ใช้สำหรับจำแนกผู้ป่วยโรคไตเรื้อรังของโรงพยาบาลอพลโล ประเทศอินเดีย คือ วิธีต้นไม้ตัดสินใจ

สรุปผล อภิปรายผล และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ในการจัดทำวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพวิธีการจำแนกกลุ่มการเป็นโรคไตเรื้อรัง 7 วิธี ได้แก่ วิธีความใกล้เคียงกันมากที่สุด (K-Nearest Neighbor) วิธีต้นไม้ตัดสินใจ (Decision Tree) วิธีโครงข่ายประสาทเทียม (Artificial Neural Network) วิธีซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) วิธีฐานกฎ (Rule-Based) วิธีการถดถอยลอจิสติก (Logistic Regression) และวิธีนาอิว์ เบย์ (Naïve Bayes) เนื่องจากวิธีดังกล่าวข้างต้นเป็นวิธีที่นิยมนำมาใช้ในการจำแนกกลุ่มข้อมูลที่ทราบผลลัพธ์แน่นอน โดยใช้หลักการของการทำเหมืองข้อมูลมาใช้ในการจำแนกกลุ่มข้อมูล ซึ่งสรุปผลได้ดังนี้

วิธีความใกล้เคียงกันมากที่สุด กรณีกำหนดค่า $k=1$ มีค่าความถูกต้องคือ 94.17% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0579 วิธีต้นไม้ตัดสินใจมีค่าความถูกต้องคือ 100% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0059 วิธีโครงข่ายประสาทเทียมมีค่าความถูกต้องคือ 98.33% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0109 วิธีซัพพอร์ตเวกเตอร์แมชชีนมีค่าความถูกต้องคือ 98.33% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0167 วิธีฐานกฎมีค่าความถูกต้องคือ 95% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0454 วิธีการถดถอยลอจิสติกมีค่าความถูกต้องคือ 93.33% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0660 และวิธีนาอิว์ เบย์มีค่าความถูกต้องคือ 95% และมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยคือ 0.0409 ซึ่งวิธีการจำแนกกลุ่มที่มีประสิทธิภาพในการจำแนกดีที่สุดสำหรับข้อมูลโรคไตเรื้อรังคือวิธีต้นไม้ตัดสินใจ โดยใช้อัลกอริทึมชนิด J48 เนื่องจากมีค่าความถูกต้องสูงที่สุดและความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

5.2 อภิปรายผลการวิจัย

จากการสรุปผลการวิจัยครั้งนี้ วิธีต้นไม้ตัดสินใจมีประสิทธิภาพในการจำแนกกลุ่ม การเป็นโรคไตเรื้อรัง โรงพยาบาลอพลโล ประเทศอินเดียมากที่สุด เพราะว่ามีค่าความถูกต้องสูงที่สุดและความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ซึ่งสรุปผลจากแบบจำลองของกฎต้นไม้ตัดสินใจเพื่อใช้ในการทำนายได้ว่า ถ้าผู้ป่วยเป็นโรคไตเรื้อรังจะปรากฏตามกฎข้อที่ 1 คือ ปริมาณครีเอตินินในเลือดน้อยกว่าเท่ากับ 1.2 มิลลิกรัมต่อเดซิลิตร และความถ่วงจำเพาะของปัสสาวะน้อยกว่าเท่ากับ 1.015 หรือกฎข้อที่ 2 คือ ปริมาณครีเอตินินในเลือดน้อยกว่าเท่ากับ 1.2 มิลลิกรัมต่อเดซิลิตร และความถ่วงจำเพาะของปัสสาวะมากกว่า 1.015 และระดับโปรตีนอัลบูมินน้อยกว่าเท่ากับ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบินน้อยกว่าเท่ากับ 13.1 กรัมต่อเดซิลิตร และปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมดน้อยกว่าเท่ากับ 43 ยูนิต หรือกฎข้อที่ 5 คือ ปริมาณครีเอตินินในเลือดน้อยกว่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เท่ากับ 1.2 มิลลิกรัมต่อเดซิลิตร และความถ่วงจำเพาะของปัสสาวะมากกว่า 1.015 และระดับโปรตีนอัลบูมินมากกว่า 0 มิลลิกรัมต่อเดซิลิตร และกฎข้อที่ 6 คือ ปริมาณครีเอตินินในเลือดมากกว่า 1.2 มิลลิกรัมต่อเดซิลิตร และถ้าผู้ป่วยไม่เป็นโรคไตเรื้อรังจะปรากฏตามกฎข้อที่ 3 คือ ปริมาณครีเอตินินในเลือดน้อยกว่าเท่ากับ 1.2 มิลลิกรัมต่อเดซิลิตร และความถ่วงจำเพาะของปัสสาวะมากกว่า 1.015 และระดับโปรตีนอัลบูมินน้อยกว่าเท่ากับ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบินน้อยกว่าเท่ากับ 13.1 กรัมต่อเดซิลิตร และปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมดมากกว่า 43 ยูนิต หรือกฎข้อที่ 4 คือ ปริมาณครีเอตินินในเลือดน้อยกว่าเท่ากับ 1.2 มิลลิกรัมต่อเดซิลิตร และความถ่วงจำเพาะของปัสสาวะมากกว่า 1.015 และระดับโปรตีนอัลบูมินน้อยกว่าเท่ากับ 0 มิลลิกรัมต่อเดซิลิตร และปริมาณฮีโมโกลบินมากกว่า 13.1 กรัมต่อเดซิลิตร ซึ่งผลสรุปดังกล่าวสามารถนำไปใช้เป็นแนวทางในการตัดสินใจเลือกวิธีการจำแนกกลุ่มที่เหมาะสมที่สุดสำหรับการใช้งานข้อมูลผู้ป่วยโรคไตเรื้อรัง และสามารถนำผลลัพธ์ที่ได้ไปสนับสนุนการตัดสินใจทางการแพทย์เกี่ยวกับการวินิจฉัยโรคไตเรื้อรัง เพื่อให้ได้ผลการวินิจฉัยที่รวดเร็ว และแม่นยำมากขึ้น จะทำให้กลุ่มผู้ป่วยโรคไตเรื้อรังลดน้อยลง

ผลการสรุปดังกล่าวใกล้เคียงกับในช่วงที่ผ่านมาที่มีผู้วิจัยได้ศึกษาวิธีการจำแนกกลุ่มด้วยข้อมูลต่าง ๆ ซึ่งผู้วิจัยส่วนใหญ่ได้สรุปว่าวิธีต้นไม้ตัดสินใจเป็นวิธีที่มีประสิทธิภาพมากที่สุดในการจำแนกกลุ่ม เช่น การเปรียบเทียบประสิทธิภาพการจำแนกกลุ่มรูปแบบการเรียนรู้ VARK ด้วยวิธีนาอิว เบย์ วิธีต้นไม้ตัดสินใจ และวิธีฐานกฎ ผลการศึกษาพบว่าวิธีต้นไม้ตัดสินใจให้ประสิทธิภาพสูงสุดคืออัลกอริทึม J48 (อรนุช และมนชัย, 2557) หรือการจำแนกกลุ่มบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้ โดยข้อมูลที่ใช้ในการทดลองได้จากหน่วยงานการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย สายงานรองผู้ว่าการบัญชีและการเงิน สายงานรองผู้ว่าการผลิตไฟฟ้า และสายงานรองผู้ว่าการเชื้อเพลิง ด้วยอัลกอริทึม 3 แบบ ได้แก่ J48, OneR และ Naïve Bayes จากการทดลองพบว่าอัลกอริทึมที่ให้ค่าความถูกต้องในการทำนายค่าตอบของความสัมพันธ์สูงสุด คือ J48 (บุษกร, 2551) หรือการเปรียบเทียบปัจจัยโรคประจำตัวผู้สูงอายุโดยใช้อัลกอริทึมการจำแนกกลุ่ม J48 และ Naïve Bayes ผลการทดสอบพบว่าวิธีต้นไม้ตัดสินใจโดยใช้อัลกอริทึม J48 ให้ค่าความถูกต้องสูงกว่าวิธีนาอิว เบย์ (เบญจภาคี, 2558) และการทดสอบประสิทธิภาพการทำนายสถานะการชำระหนี้ของลูกค้าหนี้ธนาคารด้วยวิธีต้นไม้ตัดสินใจ วิธีถดถอยลอจิสติก และวิธีนาอิว เบย์ ผลการทดสอบพบว่าวิธีต้นไม้ตัดสินใจให้ค่าความแม่นยำสูงที่สุด รองลงมาคือวิธีถดถอยลอจิสติกและนาอิว เบย์ ตามลำดับ (ชนัญฐดาภรณ์, 2557)

นอกจากนี้ งานวิจัยที่ผ่านมาได้มีการศึกษาเกี่ยวกับข้อมูลโรคไตเรื้อรัง โดยใช้วิธีซัพพอร์ตเวกเตอร์แมชชีนและโครงข่ายประสาทเทียม จากการศึกษาพบว่าวิธีโครงข่ายประสาทเทียมมีประสิทธิภาพในการทำนายผลได้ดีกว่าวิธีซัพพอร์ตเวกเตอร์แมชชีน (Vijayarani and Dhayanand, 2015) ซึ่งงานวิจัยดังกล่าวได้ให้ผลลัพธ์ที่สอดคล้องกัน เนื่องจากในงานวิจัยในครั้งนี้วิธีโครงข่ายประสาทเทียมมีประสิทธิภาพในการทำนายผลได้ดีกว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนเช่นกัน แต่อาจจะมีการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำแนกกลุ่มอื่น ๆ ที่มีประสิทธิภาพในการวิเคราะห์ข้อมูลโรคไตเรื้อรังมากกว่าวิธีดังกล่าว ดังนั้นผู้วิจัย จึงได้ศึกษาวิธีการจำแนกกลุ่มด้วยวิธีอื่น ๆ ทั้งหมด 7 วิธี เพื่อหาวิธีที่ให้ประสิทธิภาพมากที่สุดใน การวิเคราะห์ข้อมูลสำหรับโรคไตเรื้อรัง ทำให้ได้ผลการวิจัยที่แตกต่างกัน

5.3 ข้อเสนอแนะ

1) เพื่อให้ได้ข้อสรุปของผลการวิเคราะห์ข้อมูลที่มีความสมบูรณ์มากขึ้น ดังนั้นผู้วิจัยอาจ วิเคราะห์ข้อมูลด้วยอัลกอริทึมประเภทอื่นๆ ได้แก่ วิธีความใกล้เคียงกันมากที่สุดมีอัลกอริทึมที่ใช้ใน การจำแนกกลุ่มอีก 2 แบบ คือ KStar และ LWL วิธีต้นไม้ตัดสินใจมีอัลกอริทึมที่ใช้ในการจำแนกกลุ่ม อื่นๆ เช่น ID3, REPTree, Random Forest และ Random Tree วิธีโครงข่ายประสาทเทียมแบบ เพอร์เซปตรอนหลายชั้นสามารถกำหนดพารามิเตอร์ได้หลายค่า วิธีซัพพอร์ตเวกเตอร์แมชชีนมี Kernel Function อีก 2 แบบ คือ Radial Basis Function Kernel และ Sigmoid Kernel และ วิธีฐานภูมิมีอัลกอริทึมที่ใช้ในการจำแนกกลุ่มอื่นๆ เช่น JRip, OneR, PART และ ZeroR เป็นต้น

2) อาจนำข้อมูลผู้ป่วยโรคไตเรื้อรังจากประเทศอื่นๆ มาวิเคราะห์เพื่อเปรียบเทียบว่าผลที่ได้ นั้น สอดคล้องกันมากน้อยเพียงใด

3) ผู้ป่วยโรคไตเรื้อรังควรระมัดระวังปริมาณครีเอตินินในเลือดไม่ควรมากกว่า 1.2 มิลลิกรัม ต่อเดซิลิตร ค่าความถ่วงจำเพาะของปัสสาวะไม่ควรน้อยกว่าหรือเท่ากับ 1.015 ระดับโปรตีนอัลบูมิน ไม่ควรมากกว่า 0 มิลลิกรัมต่อเดซิลิตร ปริมาณฮีโมโกลบินไม่ควรน้อยกว่าหรือเท่ากับ 13.1 กรัมต่อ เดซิลิตร และปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมดไม่ควรน้อยกว่าหรือเท่ากับ 43 ยูนิต

เอกสารอ้างอิง

- ไกรรุ่ง เองพระพรหม และคณะ. 2547. การค้นหาความรู้คุณลักษณะสำคัญของนักศึกษาที่มีต่อผลสัมฤทธิ์ทางการเรียนในกลุ่มโปรแกรมคอมพิวเตอร์ มหาวิทยาลัยราชภัฏนครปฐม ด้วยเทคนิคเหมืองข้อมูล. โปรแกรมมหาวิทยาลัยการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม.
- จุฑามาส ลิทธิโชคสถาพร. 2555. พยากรณ์ราคาขายแผ่นดิบ โดยใช้ซัพพอร์ตเวกเตอร์แมชชีน. สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยสงขลานครินทร์.
- ชนัญฐดาภรณ์ เย็นประเสริฐ. 2557. การเปรียบเทียบความแม่นยำการพยากรณ์สถานการณ์ชำระหนี้ของลูกหนี้ โดยใช้เทคนิคการถดถอยลอจิสติก นาอ์ฟีย์เบย์ และต้นไม้การตัดสินใจ. สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- ฐานข้อมูล. (ม.ป.ป.). เข้าถึงได้ : archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease (วันที่ค้นหาข้อมูล : 1 มีนาคม 2559).
- เดช ธรรมศิริ และพยุ่ง มีสัจ. 2552. การจำแนกข้อมูลด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีนโดยการปรับพารามิเตอร์และเลือกคุณลักษณะที่เหมาะสมด้วยขั้นตอนวิธีเชิงพันธุกรรม. ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- บุษกร สุคนธวงศาโรจน์. 2551. อัลกอริทึมในการจำแนกบุคลากรในองค์กรสำหรับการสร้างแผนที่ความรู้. ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์.
- เบญจภัก จงหมื่นไวย. 2558. การเปรียบเทียบปัจจัยโรคประจำตัวผู้สูงอายุโดยใช้อัลกอริทึมการจัดกลุ่ม J48 และ Naïve Bayes กรณีศึกษาสาธารณสุขโพธิ์กลางนครราชสีมา. นักศึกษาโปรแกรมมหาวิทยาลัยการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา.
- พยุ่น พาณิษฐ์กุล. 2548. การพัฒนาระบบดัดไม้หนึ่งโดยใช้ Decision Tree. โครงการพัฒนาระบบงานปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ แขนงวิทยาการสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- พลอยพรรณ สอนสุวิทย์ และตรีสพงศ์ ไทยอุปถัมภ์. 2552. การเปรียบเทียบประสิทธิภาพการตรวจจับสิ่งผิดปกติทางเครือข่ายชนิด Probing Detection Efficiency Comparison of Probing-Type Network Anomaly. ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเชียงใหม่.
- พรคิด อ้นขาว และพยุ่ง มีสัจ. 2558. เทคนิคการใช้ดัดไม้หนึ่งสำหรับการพยากรณ์ผลสัมฤทธิ์ทางการศึกษา. สาขาวิชาระบบสารสนเทศ คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลพระนครและสาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

- ภัทรพงศ์ พงศ์ภัทรกานต์. 2552. การเปรียบเทียบการจำแนกข้อมูลแบบจำลอง CART, SVM, C5.0 และแบบผสมผสานกัน. ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเลย.
- รุจิรา ธรรมสมบัติ. 2554. ระบบสนับสนุนการตัดสินใจในการเลือกใช้แพคเกจอินเทอร์เน็ตมือถือโดยใช้ต้นไม้ตัดสินใจ. สาขาคอมพิวเตอร์ธุรกิจ คณะบริหารธุรกิจ วิทยาลัยราชพฤกษ์.
- วลัยลักษณ์ สุขสมบูรณ์ และสมชาย ปรากฏเจริญ. 2553. การเปรียบเทียบประสิทธิภาพการจำแนกประเภทปัญหาสำหรับระบบถามตอบโดยใช้ซอฟต์แวร์เวกเตอร์แมชชีน นาอิว เบย์ ความใกล้เคียงกันมากที่สุด. ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- วรรณสิริ ฐระชน และคณะ. 2557. ระบบการจำแนกพันธุ์ยางพาราโดยใช้ตัวจำแนกนาอิวเบย์. สาขาวิชาวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุดรธานี.
- วาทีณี น้อยเพียร และคณะ. 2553. การเปรียบเทียบประสิทธิภาพและวิเคราะห์การจำแนกข้อมูลด้วยโครงข่ายประสาทเทียม. ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- ศุภกฤษ์ ชูธงชัย. 2546. ระบบเครือข่ายสมองเทียมสำหรับการพยากรณ์อากาศ. วิทยาลัยนวัตกรรมการอุดมศึกษา มหาวิทยาลัยธรรมศาสตร์.
- สายชล สีนสมบูรณ์ทอง. 2559. การทำเหมืองข้อมูลเล่ม 2 : วิธีการและตัวแบบ. กรุงเทพฯ : จามจุรีโปรดักส์.
- สายชล สีนสมบูรณ์ทอง. 2559. การวิเคราะห์หลายตัวแปร. กรุงเทพฯ : จามจุรีโปรดักส์.
- สุขฤทัย นิมาชาติ. 2556. ผลของโปรแกรมการขึ้นเาะเพื่อสุขภาพต่อพฤติกรรมกรรมการบริโภคอาหารและผลลัพธ์ทางคลินิกของผู้ป่วยโรคไตเรื้อรัง. สาขาวิชาการพยาบาลผู้ใหญ่ คณะพยาบาลศาสตร์ มหาวิทยาลัยบูรพา.
- สมชาย ประสิทธิ์จตุระกุล. 2545. การออกแบบและวิเคราะห์อัลกอริทึม. (ม.ป.ป.). เข้าถึงได้ : th.wikipedia.org/wiki/ขั้นตอนวิธี (วันที่ค้นหาข้อมูล : 11 มีนาคม 2559).
- อานนท์ นามสนิท. 2549. การจำแนกกลุ่มเพลงไทยโดยใช้ซอฟต์แวร์เวกเตอร์แมชชีน. สาขาวิชาเทคโนโลยีคอมพิวเตอร์ ภาควิชาคอมพิวเตอร์ศึกษามหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.
- อรนุช พันโท และมนชัย เทียนทอง. 2557. การเปรียบเทียบประสิทธิภาพการจำแนกรูปแบบการเรียนรู้ VARK ด้วยเทคนิคเหมืองข้อมูล. คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

- Berson, A. and Smith, S. J. 2001. **Data Warehousing, Data Mining, & OLAP.** Boston : McGraw-Hill.
- Connolly, T. M. and Begg, C. E. 2005. **Database Systems: A Practical Approach to Design, Implementation, and Management.** New York : Addison-Wesley.
- Kharya, S. 2012. **Using Data Mining Techniques for Diagnosis and Prognosis of Cancer Disease.** Bhilai Institute of Technology, Durg-491 001, Chhattisgarh, India.
- Kinzang Wangdi. **Classification of Loan Borrowers of National Pension and Provident Fund Bhutan : A Case Study.** Information Technology King Mongkut's University of Technology North Bangkok. 2009.
- Lakshmi, K. R. et al. 2014. **Performance Comparison of Three Data Mining Techniques for Predicting Kidney Dialysis Survivability.** Maddur Nagar, Kurnool, Andhra Pradesh, India.
- Larose, D. T. 2005. **Discovering Knowledge in Data : An Introduction to Data Mining.** New Jersey : John Wiley and Sons.
- Troyanskaya, O. et al. **Missing values estimation methods for DNA microarrays.** *Bioinformatics*, 17, 2001.
- Vijayarani, S. and Dhayanand, S. 2015. **Kidney Disease Prediction Using SVM and Algorithms.** Department of Computer Science, School of Computer Science and Engineering, Bharathiar University, Coimbatore, Tamilnadu, India.
- Witten, I. H. et al. 2011. 3rd ed. **Data Mining : Practical Machine Learning Tools and Techniques.** Amsterdam : Morgan Kaufmann.
- Wu, X. and Kumar, V. 2009. **The Top Ten Algorithms in Data Mining.** University of Minnesota, Department of Computer Science and Engineering, Minneapolis, Minnesota, U.S.A.

ภาคผนวก ก

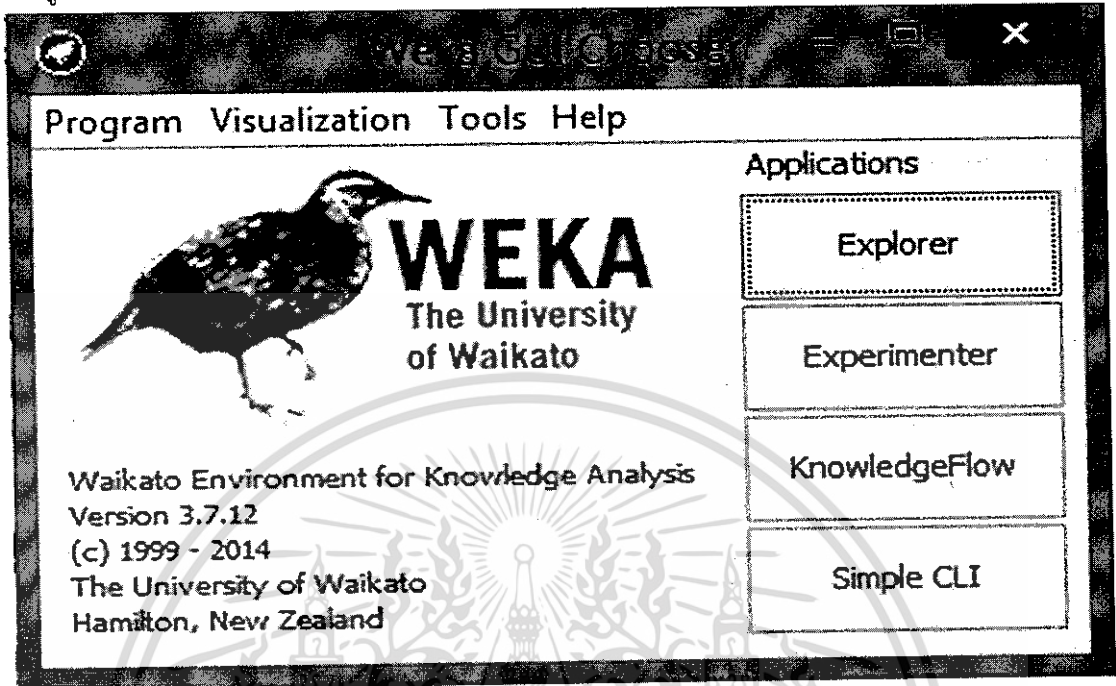
รายละเอียดขั้นตอนการใช้โปรแกรม
และตัวอย่างข้อมูลที่นำมาวิเคราะห์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

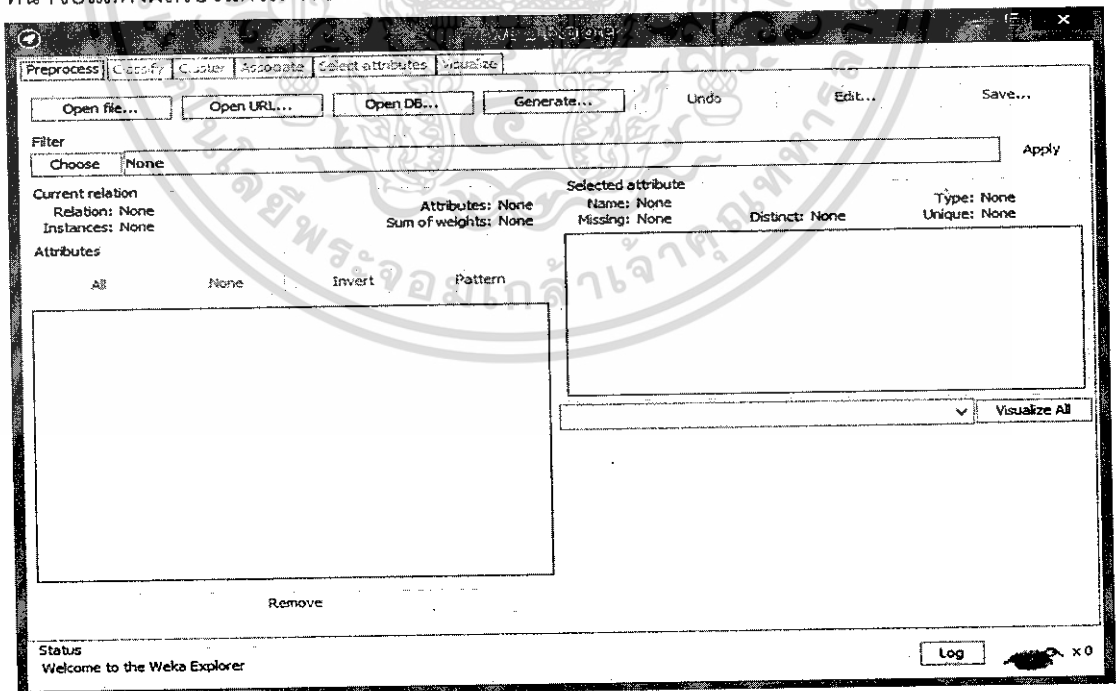
รายละเอียดขั้นตอนการใช้โปรแกรม

เข้าสู่โปรแกรม Weka คลิก Explorer



รูปที่ ก-1 โปรแกรม Weka

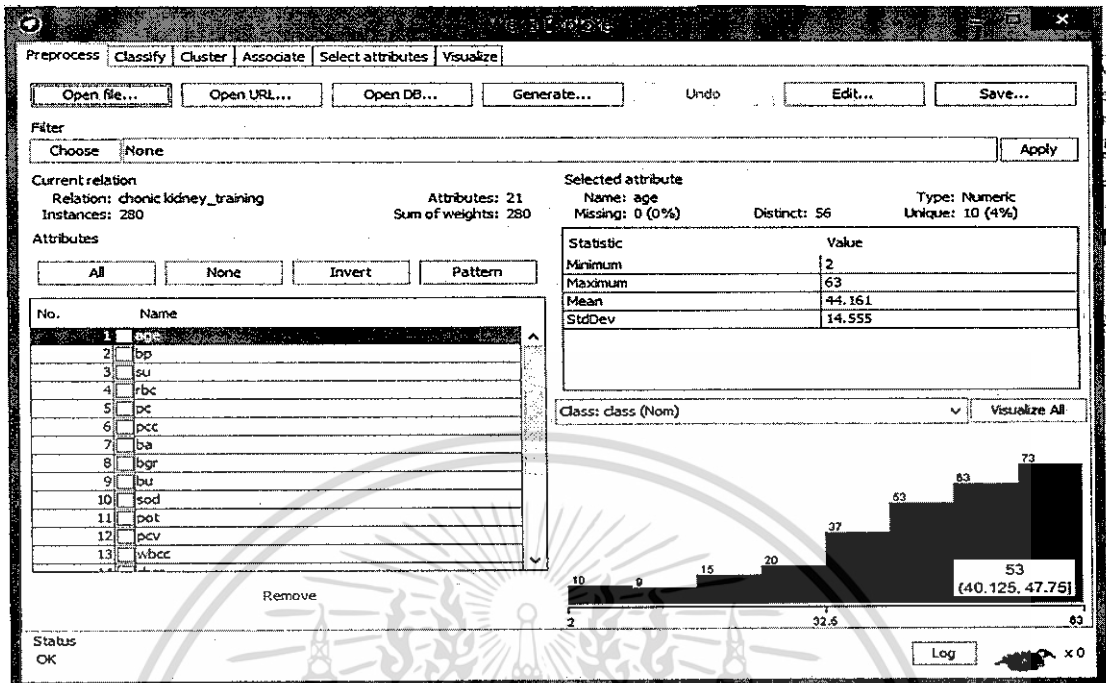
หน้าจอแสดงผลโปรแกรม Weka



รูปที่ ก-2 หน้าจอแสดงผลโปรแกรม Weka

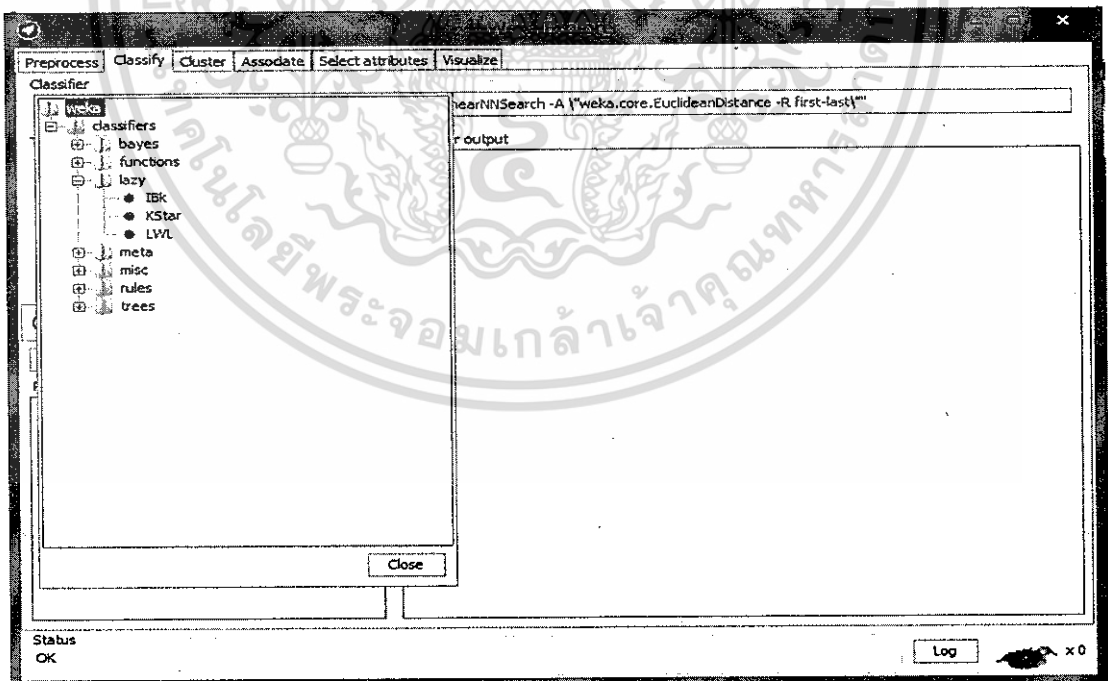
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปิดไฟล์ chronic_kidney_training



รูปที่ ก-3 ไฟล์การสร้างตัวแบบ

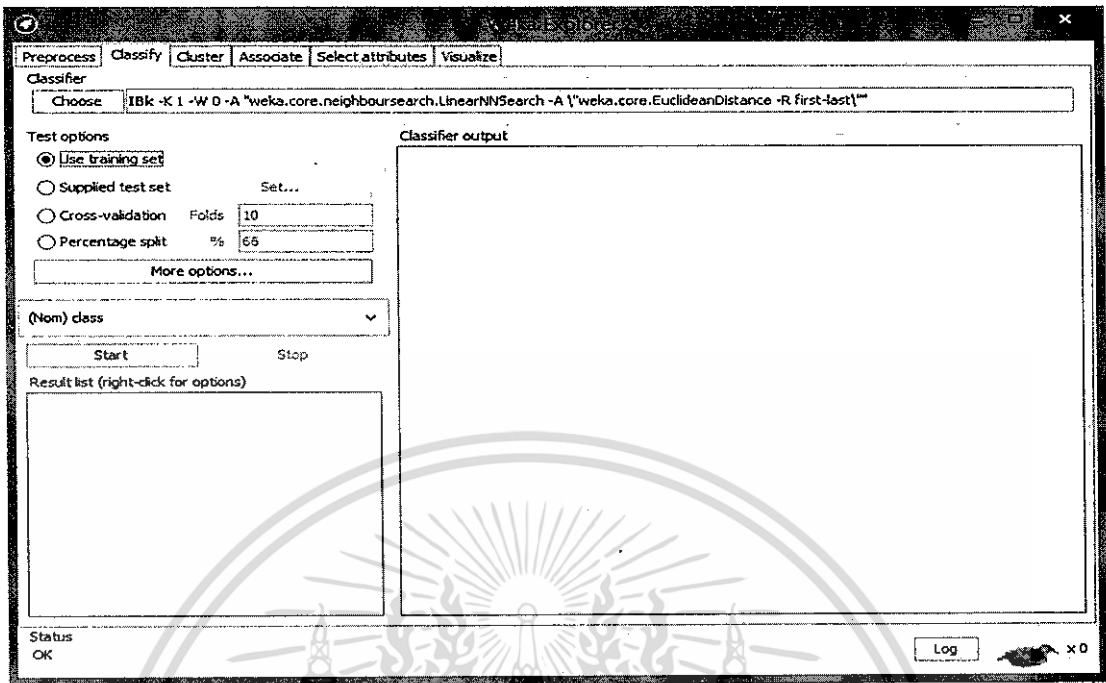
เลือก Classify จะมีอัลกอริทึมแต่ละวิธีให้เลือก



รูปที่ ก-4 หน้าจอแสดงวิธีการจำแนกกลุ่ม

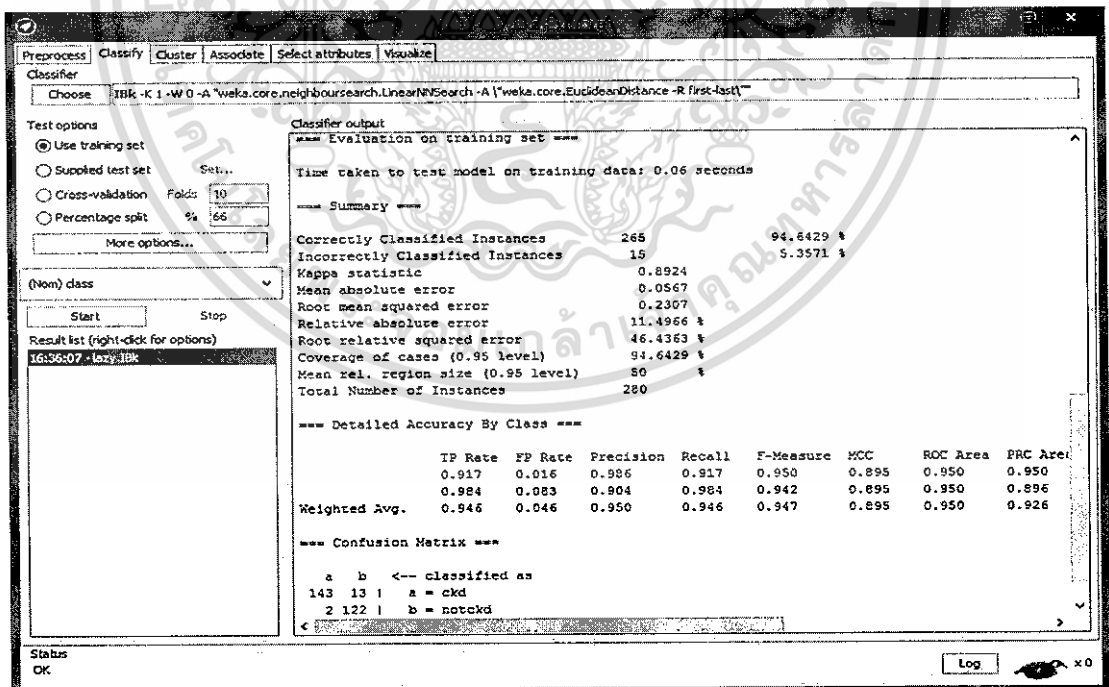
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากเลือกอัลกอริทึมที่ต้องการ คลิก Use training set เพื่อสร้างตัวแบบ



รูปที่ ก-5 ขั้นตอนการสร้างตัวแบบ

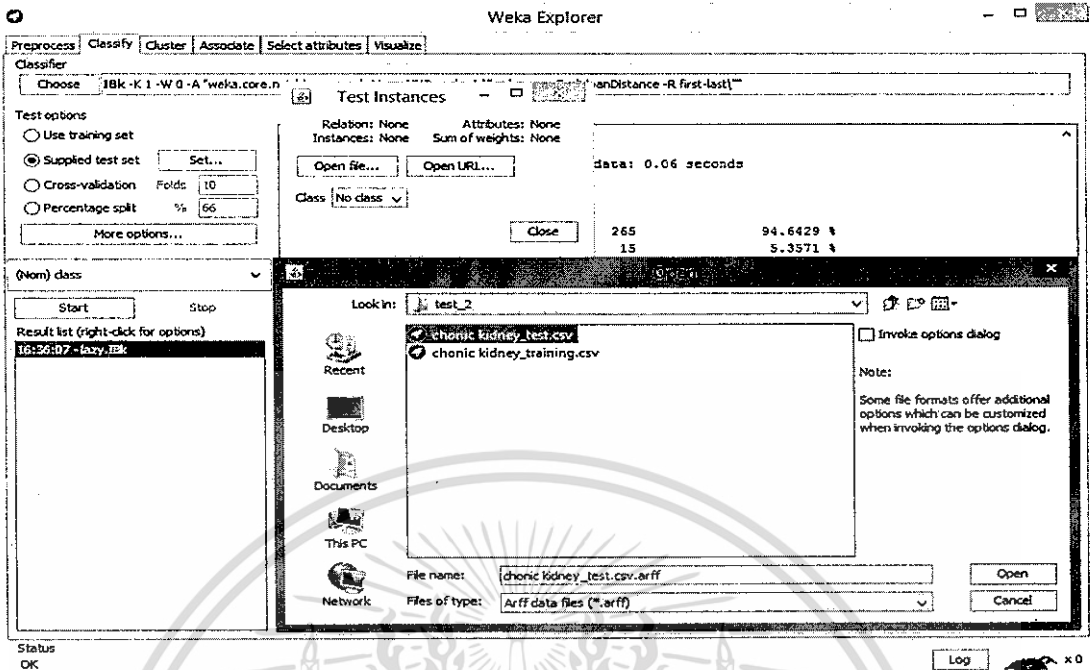
หน้าจอแสดงผลการสร้างตัวแบบ



รูปที่ ก-6 หน้าจอแสดงผลการสร้างตัวแบบ

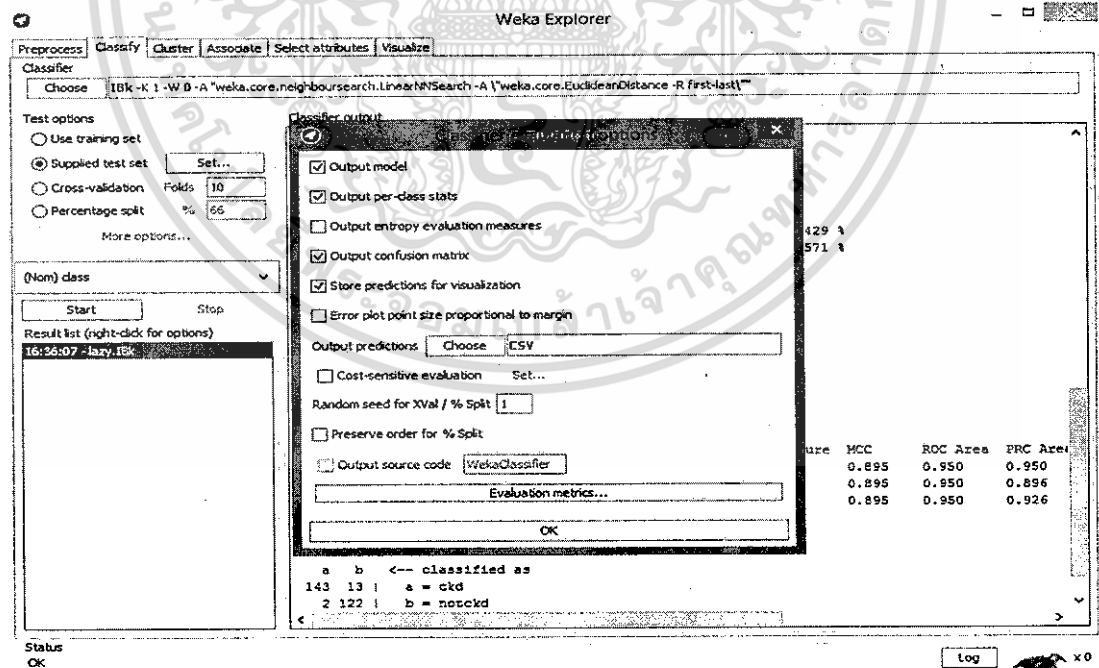
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หลังจากนั้นคลิก Supplied test set เลือกไฟล์ choric_kidney_test เพื่อทดสอบตัวแบบ



รูปที่ ก-7 ขั้นตอนการทดสอบตัวแบบ

คลิก More options เลือก CSV เสร็จแล้ว คลิกปุ่ม Start



รูปที่ ก-8 ขั้นตอนการเลือก Classifier evaluation options

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หน้าจอบ่งแสดงผลการทดสอบตัวแบบ

The screenshot shows the Weka Classifier output window. The classifier used is lazy.IBk. The output includes evaluation on a test set, a summary of performance metrics, detailed accuracy by class, and a confusion matrix.

Classifier output

==== Evaluation on test set ====

Time taken to test model on supplied test set: 0.09 seconds

==== Summary ====

Correctly Classified Instances	114	95	%
Incorrectly Classified Instances	6	5	%
Kappa statistic	0.2567		
Mean absolute error	0.0532		
Root mean squared error	0.2228		
Relative absolute error	11.3694	%	
Root relative squared error	47.3949	%	
Coverage of cases (0.95 level)	95	%	
Mean rel. region size (0.95 level)	50	%	
Total Number of Instances	120		

==== Detailed Accuracy By Class ====

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
	0.957	0.077	0.978	0.957	0.968	0.858	0.940	0.970
	0.923	0.043	0.857	0.923	0.889	0.858	0.940	0.808
Weighted Avg.	0.950	0.069	0.952	0.950	0.951	0.858	0.940	0.935

==== Confusion Matrix ====

a b <-- classified as

90	4	a = ckd
2	24	b = nonekd

Status: OK

รูปที่ ก-9 หน้าจอบ่งแสดงผลการทดสอบตัวแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก-1 คุณลักษณะและรายละเอียดของข้อมูลผู้ป่วยโรคไตเรื้อรัง

ตัวแปร	ความหมาย	มาตราวัด	หน่วย
X ₁	อายุ (age)	numerical	ปี (year)
X ₂	ความดันโลหิต (blood pressure)	numerical	มิลลิเมตรต่อปรอท (mm/Hg)
X ₃	ความถ่วงจำเพาะของปัสสาวะ (urine specific gravity)	nominal	ไม่มีหน่วย
X ₄	ระดับโปรตีนอัลบูมิน (albumin)	nominal	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₅	ระดับน้ำตาล (sugar)	nominal	ความเข้มข้นโมลาร์ (concentration)
X ₆	เซลล์เม็ดเลือดแดง (red blood cells)	nominal	ปกติ, ไม่ปกติ (normal, abnormal)
X ₇	เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่น (pus cell)	nominal	ปกติ, ไม่ปกติ (normal, abnormal)
X ₈	เซลล์เม็ดเลือดขาวปนปัสสาวะขุ่นเป็นกลุ่มก้อน (pus cell clumps)	nominal	ปรากฏ, ไม่ปรากฏ (present, notpresent)
X ₉	แบคทีเรีย (bacteria)	nominal	ปรากฏ, ไม่ปรากฏ (present, notpresent)
X ₁₀	ปริมาณน้ำตาลในเลือดแบบสุ่ม (blood glucose random)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₁	ปริมาณไนโตรเจนในกระแสเลือด (blood urea nitrogen)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₂	ปริมาณครีเอตินินในเลือด (serum creatinine)	numerical	มิลลิกรัมต่อเดซิลิตร (mgs/dl)
X ₁₃	ปริมาณโซเดียมในเลือด (sodium)	numerical	มิลลิอิควิวาเลนซ์ต่อลิตร (mEq/L)
X ₁₄	ปริมาณโปแตสเซียมในเลือด (potassium)	numerical	มิลลิอิควิวาเลนซ์ต่อลิตร (mEq/L)
X ₁₅	ปริมาณฮีโมโกลบิน (hemoglobin)	numerical	กรัมต่อเลือด 100 มิลลิลิตร (gms/dl)
X ₁₆	ปริมาณเม็ดเลือดแดงต่อปริมาณเลือดทั้งหมด (packed cell volume)	numerical	ยูนิต (Unit)
X ₁₇	จำนวนเซลล์เม็ดเลือดขาว (white blood cell count)	numerical	เซลล์ต่อลูกบาศก์มิลลิเมตร (cells/cumm)
X ₁₈	จำนวนเซลล์เม็ดเลือดแดง (red blood cell count)	numerical	ล้านเซลล์ต่อลูกบาศก์มิลลิเมตร (millions/cumm)
X ₁₉	ภาวะความดันโลหิตสูง (hypertension)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₀	เบาหวาน (diabetes mellitus)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₁	โรคหลอดเลือดหัวใจ (coronary artery disease)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₂	ความอยากอาหาร (appetite)	nominal	ต้องการมาก, ต้องการน้อย (good, poor)
X ₂₃	อาการเท้าบวม (pedal edema)	nominal	เป็น, ไม่เป็น (yes, no)
X ₂₄	โรคโลหิตจาง (anemia)	nominal	เป็น, ไม่เป็น (yes, no)
Y	การเป็นโรคไตเรื้อรัง (Chronic kidney)	nominal	0 = ผู้ป่วยไม่เป็นโรคไตเรื้อรัง 1 = ผู้ป่วยเป็นโรคไตเรื้อรัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 1-40

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
1	2	?	1.01	3	0	normal	abnormal	notpresent	notpresent	?	?	?	?
2	3	?	1.01	2	0	normal	normal	notpresent	notpresent	?	22	0.7	?
3	4	?	1.02	1	0	?	normal	notpresent	notpresent	99	23	0.6	138
4	5	?	1.015	1	0	?	normal	notpresent	notpresent	?	16	0.7	138
5	5	50	1.01	0	0	?	normal	notpresent	notpresent	?	25	0.6	?
6	7	50	1.02	4	0	?	normal	notpresent	notpresent	?	18	0.8	?
7	8	60	1.025	3	0	normal	normal	notpresent	notpresent	78	27	0.9	?
8	8	?	?	?	?	?	?	notpresent	notpresent	80	66	2.5	142
9	12	60	1.015	3	0	abnormal	abnormal	present	notpresent	?	51	1.8	?
10	12	80	1.02	0	0	normal	normal	notpresent	notpresent	100	26	0.6	137
11	14	?	1.015	0	0	?	?	notpresent	notpresent	192	15	0.8	137
12	15	60	1.02	3	0	?	normal	notpresent	notpresent	86	15	0.6	138
13	15	80	1.025	0	0	normal	normal	notpresent	notpresent	93	17	0.9	136
14	17	60	1.01	0	0	?	normal	notpresent	notpresent	92	32	2.1	141
15	17	70	1.015	1	0	abnormal	normal	notpresent	notpresent	22	1.5	7.3	145
16	17	60	1.025	0	0	normal	normal	notpresent	notpresent	114	50	1	135
17	19	80	1.02	0	0	normal	normal	notpresent	notpresent	107	23	0.7	141
18	20	60	1.025	0	0	normal	normal	notpresent	notpresent	?	?	?	137
19	20	70	1.02	0	0	normal	normal	notpresent	notpresent	123	44	1	135
20	21	70	1.01	0	0	?	normal	notpresent	notpresent	?	?	?	?
21	21	90	1.01	4	0	normal	abnormal	present	present	107	40	1.7	125
22	23	80	1.025	0	0	normal	normal	notpresent	notpresent	111	34	1.1	145
23	24	?	1.015	2	4	normal	abnormal	notpresent	notpresent	410	31	1.1	?
24	24	70	1.025	0	0	normal	normal	notpresent	notpresent	140	23	0.6	140
25	24	80	1.025	0	0	normal	normal	notpresent	notpresent	125	?	?	136
26	24	80	1.02	0	0	normal	normal	notpresent	notpresent	93	46	1	145
27	25	60	1.02	0	0	normal	normal	notpresent	notpresent	119	27	0.5	?
28	25	70	1.02	0	0	normal	normal	notpresent	notpresent	88	42	0.5	136
29	28	70	1.02	0	0	normal	normal	?	?	131	29	0.6	145
30	28	60	1.025	0	0	normal	normal	notpresent	notpresent	79	50	0.5	145
31	29	70	1.02	0	0	normal	normal	notpresent	notpresent	127	44	1.2	145
32	29	80	1.02	0	0	normal	normal	notpresent	notpresent	83	49	0.9	139
33	29	80	1.02	0	0	normal	normal	notpresent	notpresent	70	16	0.7	138
34	30	70	1.015	0	0	?	normal	notpresent	notpresent	101	106	6.5	135
35	30	80	1.02	0	0	normal	normal	notpresent	notpresent	131	38	1	147
36	30	80	1.025	0	0	normal	normal	notpresent	notpresent	96	25	0.5	144
37	30	60	1.02	0	0	normal	normal	notpresent	notpresent	138	15	1.1	135
38	32	90	1.025	1	0	abnormal	abnormal	notpresent	notpresent	?	223	18.1	113
39	33	80	1.025	0	0	normal	normal	notpresent	notpresent	100	37	1.2	142
40	33	80	1.025	0	0	normal	normal	notpresent	notpresent	128	38	0.6	135

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 1-40

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
1	?	?	?	?	?	no	no	no	good	yes	no	ckd
2	?	10.7	34	12300	?	no	no	no	good	no	no	ckd
3	4.4	12	34	?	?	no	no	no	good	no	no	ckd
4	3.2	8.1	?	?	?	no	no	no	good	no	yes	ckd
5	?	11.8	36	12400	?	no	no	no	good	no	no	ckd
6	?	11.3	38	6000	?	no	no	no	good	no	no	ckd
7	?	12.3	41	6700	?	no	no	no	poor	yes	no	ckd
8	3.6	12.2	38	?	?	no	no	no	good	no	no	ckd
9	?	12.1	?	10300	?	no	no	no	good	no	no	ckd
10	4.4	15.8	49	6600	5.4	no	no	no	good	no	no	notckd
11	4.2	14.3	40	9500	5.4	no	yes	no	poor	yes	no	ckd
12	4	11	33	7700	3.8	yes	yes	no	good	no	no	ckd
13	3.9	16.7	50	6200	5.2	no	no	no	good	no	no	notckd
14	4.2	13.9	52	7000	?	no	no	no	good	no	no	ckd
15	2.8	13.1	41	11200	?	no	no	no	good	no	no	ckd
16	4.9	14.2	51	7200	5.9	no	no	no	good	no	no	notckd
17	4.2	14.4	44	?	?	no	no	no	good	no	no	notckd
18	4.7	14	41	4500	5.5	no	no	no	good	no	no	notckd
19	3.8	14.6	44	5500	4.8	no	no	no	good	no	no	notckd
20	?	?	?	?	?	no	no	no	poor	no	yes	ckd
21	3.5	8.3	23	12400	3.9	no	no	no	good	no	yes	ckd
22	4	14.3	41	7200	?	no	no	no	good	no	no	notckd
23	?	12.4	44	6900	5	no	yes	no	good	yes	no	ckd
24	4.7	16.3	48	5800	5.6	no	no	no	good	no	no	notckd
25	3.5	15.4	43	5600	4.5	no	no	no	good	no	no	notckd
26	3.5	?	?	10700	6.3	no	no	no	good	no	no	notckd
27	?	15.2	40	9200	5.2	no	no	no	good	no	no	notckd
28	3.5	13.3	48	7000	4.9	no	no	no	good	no	no	notckd
29	4.9	?	45	8600	6.5	no	no	no	good	no	no	notckd
30	5	17.6	51	6500	5	no	no	no	good	no	no	notckd
31	5	14.8	48	?	?	no	no	no	good	no	no	notckd
32	3.3	17.5	40	9900	4.7	no	no	no	good	no	no	notckd
33	3.5	13.7	54	5400	5.8	no	no	no	good	no	no	notckd
34	4.3	?	?	?	?	no	no	no	poor	no	no	ckd
35	3.8	14.1	45	9400	5.3	no	no	no	good	no	no	notckd
36	4.8	13.8	42	9000	4.5	no	no	no	good	no	no	notckd
37	4.4	?	?	?	?	no	no	no	good	no	no	notckd
38	6.5	5.5	15	2600	2.8	yes	yes	no	poor	yes	yes	ckd
39	4	16.9	52	6700	6	no	no	no	good	no	no	notckd
40	3.9	13.1	45	6200	4.5	no	no	no	good	no	no	notckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 41-80

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
41	34	70	?	?	?	?	?	notpresent	notpresent	?	219	12.2	130
42	34	70	1.02	0	0	abnormal	normal	notpresent	notpresent	139	19	0.9	?
43	34	90	1.015	2	0	normal	normal	notpresent	notpresent	104	50	1.6	137
44	34	70	1.025	0	0	normal	normal	notpresent	notpresent	?	33	1	150
45	34	70	1.025	0	0	normal	normal	notpresent	notpresent	87	38	0.5	144
46	35	80	1.01	1	0	abnormal	?	notpresent	notpresent	79	202	10.8	134
47	35	80	1.005	3	0	abnormal	normal	notpresent	notpresent	?	?	?	?
48	35	?	1.02	0	0	normal	normal	?	?	99	30	0.5	135
49	35	60	1.025	0	0	normal	normal	notpresent	notpresent	105	39	0.5	135
50	36	80	1.01	0	0	?	normal	notpresent	notpresent	103	?	?	?
51	36	80	1.025	0	0	normal	normal	notpresent	notpresent	85	16	1.1	142
52	37	100	1.01	0	0	abnormal	normal	notpresent	notpresent	?	19	1.3	?
53	37	60	1.025	0	0	normal	normal	notpresent	notpresent	111	35	0.8	135
54	37	60	1.02	0	0	normal	normal	notpresent	notpresent	109	47	1.1	141
55	38	70	?	?	?	?	?	notpresent	notpresent	104	77	1.9	140
56	38	60	1.02	0	0	normal	normal	notpresent	notpresent	91	36	0.7	135
57	39	70	1.01	0	0	?	normal	notpresent	notpresent	121	20	0.8	133
58	39	60	1.02	0	0	normal	normal	notpresent	notpresent	86	37	0.6	150
59	40	80	1.015	3	0	?	normal	notpresent	notpresent	76	162	9.6	141
60	40	70	1.015	3	4	normal	normal	notpresent	notpresent	253	150	11.9	132
61	40	80	1.025	0	0	normal	normal	notpresent	notpresent	140	10	1.2	135
62	40	80	1.02	0	0	normal	normal	notpresent	notpresent	119	15	0.7	150
63	41	70	1.015	2	0	?	abnormal	notpresent	present	?	68	2.8	132
64	41	80	1.015	1	4	abnormal	normal	notpresent	notpresent	210	165	18	135
65	41	70	1.02	0	0	normal	normal	notpresent	notpresent	125	38	0.6	140
66	41	80	1.025	0	0	normal	normal	notpresent	notpresent	112	48	0.7	140
67	42	100	1.015	4	0	normal	abnormal	notpresent	present	?	50	1.4	129
68	42	90	1.02	2	0	abnormal	abnormal	present	notpresent	93	153	2.7	139
69	42	80	1.02	0	0	normal	normal	notpresent	notpresent	132	24	0.7	140
70	42	70	1.02	0	0	normal	normal	notpresent	notpresent	93	32	0.9	143
71	43	80	1.015	2	3	?	abnormal	present	present	?	30	1.1	?
72	44	90	1.01	1	0	?	normal	notpresent	notpresent	?	20	1.1	?
73	44	60	1.025	0	0	normal	normal	notpresent	notpresent	96	33	0.9	147
74	44	70	1.025	0	0	normal	normal	notpresent	notpresent	?	?	?	?
75	44	70	1.025	0	0	normal	normal	notpresent	notpresent	92	40	0.9	141
76	44	60	1.02	0	0	normal	normal	notpresent	notpresent	95	46	0.5	138
77	45	80	1.02	3	0	normal	abnormal	notpresent	notpresent	425	?	?	?
78	45	90	?	?	?	?	?	notpresent	notpresent	360	45	2.4	128
79	45	60	1.01	2	0	normal	abnormal	present	notpresent	268	86	4	134
80	45	70	1.025	2	0	normal	abnormal	present	notpresent	117	52	2.2	136

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 41-80

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class.variable
41	3.8	6	?	?	?	yes	no	no	good	no	yes	ckd
42	?	12.7	42	2200	?	no	no	no	poor	no	no	ckd
43	4.1	11.9	39	?	?	no	no	no	good	no	no	ckd
44	5	15.3	44	10500	6.1	no	no	no	good	no	no	notckd
45	4.8	17.1	47	7400	6.1	no	no	no	good	no	no	notckd
46	3.4	7.9	24	7900	3.1	no	yes	no	good	no	no	ckd
47	?	9.5	28	?	?	no	no	no	good	yes	no	ckd
48	4.9	15.4	48	5000	5.2	no	no	no	good	no	no	notckd
49	3.9	14.7	43	5800	6.2	no	no	no	good	no	no	notckd
50	?	11.9	36	8800	?	no	no	no	good	no	no	ckd
51	4.1	15.6	44	5800	6.3	no	no	no	good	no	no	notckd
52	?	15	44	4100	5.2	yes	no	no	good	no	no	ckd
53	4.1	16.2	50	5500	5.7	no	no	no	good	no	no	notckd
54	4.9	15	48	7000	5.2	no	no	no	good	no	no	notckd
55	3.9	?	?	?	?	yes	no	no	poor	yes	no	ckd
56	3.7	14	46	9100	5.8	no	no	no	good	no	no	notckd
57	3.5	10.9	32	?	?	no	yes	no	good	no	no	ckd
58	5	13.6	51	5800	4.5	no	no	no	good	no	no	notckd
59	4.9	7.6	24	3800	2.8	yes	no	no	good	no	yes	ckd
60	5.6	10.9	31	8800	3.4	yes	yes	no	poor	yes	no	ckd
61	5	15	48	10400	4.5	no	no	no	good	no	no	notckd
62	4.9	?	?	?	?	no	no	no	good	no	no	notckd
63	4.1	11.1	33	?	?	yes	no	no	good	yes	yes	ckd
64	4.7	?	?	?	?	no	yes	no	good	no	no	ckd
65	5	16.8	41	6300	5.9	no	no	no	good	no	no	notckd
66	5	17	52	7200	5.8	no	no	no	good	no	no	notckd
67	4	11.1	39	8300	4.6	yes	no	no	poor	no	no	ckd
68	4.3	9.8	34	9800	?	no	no	no	poor	yes	yes	ckd
69	4.1	14.4	50	5000	4.5	no	no	no	good	no	no	notckd
70	4.7	16.6	43	7100	5.3	no	no	no	good	no	no	notckd
71	?	14	42	14900	?	no	no	no	good	no	no	ckd
72	?	15	48	?	?	no	no	no	good	no	no	ckd
73	4.5	16.9	41	7200	5	no	no	no	good	no	no	notckd
74	?	13.8	48	7800	4.4	no	no	no	good	no	no	notckd
75	4.9	14	52	7500	6.2	no	no	no	good	no	no	notckd
76	4.2	15	50	7700	6.3	no	no	no	good	no	no	notckd
77	?	?	?	?	?	no	no	no	poor	no	no	ckd
78	4.4	8.3	29	5500	3.7	yes	yes	no	good	no	no	ckd
79	5.1	10	29	9200	?	yes	yes	no	good	no	no	ckd
80	3.8	10	30	19100	3.7	no	no	no	good	no	no	ckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 81-120

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
81	45	80	1.025	0	0	normal	normal	notpresent	notpresent	82	49	0.6	147
82	45	60	1.02	0	0	normal	normal	?	?	114	26	0.7	141
83	46	70	1.015	1	0	abnormal	normal	notpresent	notpresent	150	111	6.1	131
84	46	90	1.02	?	?	?	normal	notpresent	notpresent	213	68	2.8	146
85	46	80	1.01	0	0	?	normal	notpresent	notpresent	160	40	2	140
86	46	110	1.015	0	0	?	normal	notpresent	notpresent	130	16	0.9	?
87	46	60	1.025	0	0	normal	normal	notpresent	notpresent	123	46	1	135
88	46	60	1.02	0	0	normal	normal	notpresent	notpresent	102	27	0.7	142
89	47	100	1.01	0	0	?	normal	notpresent	notpresent	204	29	1	139
90	47	80	1.01	0	0	?	abnormal	notpresent	notpresent	?	28	0.9	?
91	47	100	1.01	?	?	normal	?	notpresent	notpresent	122	?	16.9	138
92	47	80	1.02	0	0	normal	normal	notpresent	notpresent	95	35	0.9	140
93	47	60	1.02	0	0	normal	normal	notpresent	notpresent	137	17	0.5	150
94	47	60	1.02	0	0	normal	normal	notpresent	notpresent	109	25	1.1	141
95	47	60	1.02	0	0	normal	normal	notpresent	notpresent	117	22	1.2	138
96	48	80	1.02	1	0	?	normal	notpresent	notpresent	121	36	1.2	?
97	48	70	1.005	4	0	normal	abnormal	present	notpresent	117	56	3.8	111
98	48	70	1.015	0	0	?	normal	notpresent	notpresent	124	24	1.2	142
99	48	80	1.005	4	0	abnormal	abnormal	notpresent	present	133	139	8.5	132
100	48	80	1.015	0	2	?	normal	notpresent	notpresent	214	24	1.3	140
101	48	100	?	?	?	?	?	notpresent	notpresent	103	79	5.3	135
102	48	110	1.015	3	0	abnormal	normal	present	notpresent	106	215	15.2	120
103	48	80	1.025	0	0	normal	normal	notpresent	notpresent	122	33	0.9	146
104	48	60	1.02	0	0	normal	normal	notpresent	notpresent	112	44	1.2	142
105	48	80	1.025	0	0	normal	normal	notpresent	notpresent	75	22	0.8	137
106	50	60	1.01	2	4	?	abnormal	present	notpresent	490	55	4	?
107	50	90	?	?	?	?	?	notpresent	notpresent	89	118	6.1	127
108	50	?	?	?	?	normal	?	notpresent	notpresent	219	176	13.8	136
109	50	70	1.02	3	0	abnormal	normal	present	present	123	40	1.8	?
110	50	70	1.01	0	0	?	normal	notpresent	notpresent	230	50	2.2	?
111	50	90	1.015	1	0	abnormal	abnormal	notpresent	notpresent	?	?	?	?
112	50	80	1.02	0	0	normal	normal	notpresent	notpresent	97	40	0.6	150
113	50	?	1.02	0	0	normal	normal	notpresent	notpresent	92	19	1.2	150
114	50	80	1.02	0	0	normal	normal	notpresent	notpresent	137	46	0.8	139
115	51	80	1.01	2	0	normal	normal	notpresent	notpresent	106	26	1.4	?
116	51	60	1.025	0	0	normal	normal	notpresent	notpresent	99	38	0.8	135
117	52	90	1.015	4	3	normal	abnormal	notpresent	notpresent	224	166	5.6	133
118	52	80	1.02	0	0	normal	normal	notpresent	notpresent	125	22	1.2	139
119	52	80	1.02	0	0	normal	normal	notpresent	notpresent	128	30	1.2	140
120	52	80	1.025	0	0	normal	normal	notpresent	notpresent	99	25	0.8	135

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 81-120

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
81	4.4	15.9	46	9100	4.7	no	no	no	good	no	no	notckd
82	4.2	15	43	9200	5.8	no	no	no	good	no	no	notckd
83	3.7	7.5	27	?	?	no	no	no	good	no	yes	ckd
84	6.3	9.3	?	?	?	yes	yes	no	good	no	no	ckd
85	4.1	9	27	8100	3.2	yes	no	no	poor	no	yes	ckd
86	?	?	?	?	?	no	no	no	good	no	no	ckd
87	5	15.7	50	6300	4.8	no	no	no	good	no	no	notckd
88	4.9	13.2	44	11000	5.4	no	no	no	good	no	no	notckd
89	4.2	9.7	33	9200	4.5	yes	no	no	good	no	yes	ckd
90	?	12.4	44	5600	4.3	no	no	no	good	no	yes	ckd
91	5.2	10.8	33	10200	3.8	no	yes	no	good	no	no	ckd
92	4.1	?	?	?	?	no	no	no	good	no	no	notckd
93	3.5	13.6	44	7900	4.5	no	no	no	good	no	no	notckd
94	4.7	15.8	41	8300	5.2	no	no	no	good	no	no	notckd
95	3.5	13	45	5200	5.6	no	no	no	good	no	no	notckd
96	?	15.4	44	7800	5.2	yes	yes	no	good	no	no	ckd
97	2.5	11.2	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
98	4.2	12.4	37	6400	4.7	no	yes	no	good	no	no	ckd
99	5.5	10.3	36	6200	4	no	yes	no	good	yes	no	ckd
100	4	13.2	39	?	?	no	yes	no	poor	no	no	ckd
101	6.3	6.3	19	7200	2.6	yes	no	yes	poor	no	no	ckd
102	5.7	8.6	26	5000	2.5	yes	no	yes	good	no	yes	ckd
103	3.9	13.9	48	9500	4.8	no	no	no	good	no	no	notckd
104	4.9	14.5	44	9400	6.4	no	no	no	good	no	no	notckd
105	5	16.8	51	6000	6.5	no	no	no	good	no	no	notckd
106	?	9.4	28	?	?	yes	yes	no	good	no	yes	ckd
107	4.4	6	17	6500	?	yes	yes	no	good	yes	yes	ckd
108	4.5	8.6	24	13200	2.7	yes	no	no	good	yes	yes	ckd
109	?	11.1	36	4700	?	no	no	no	good	no	no	ckd
110	?	12	41	10400	4.6	yes	yes	no	good	no	no	ckd
111	?	?	?	?	?	no	no	no	good	yes	no	ckd
112	4.5	14.2	48	10500	5	no	no	no	good	no	no	notckd
113	4.8	14.9	48	4700	5.4	no	no	no	good	no	no	notckd
114	5	14.1	45	9500	4.6	no	no	no	good	no	no	notckd
115	?	11.6	35	7300	4.6	no	no	no	good	no	no	ckd
116	3.7	13	49	8300	5.2	no	no	no	good	no	no	notckd
117	4.7	8.1	23	5000	2.9	yes	yes	no	good	no	yes	ckd
118	4.6	16.5	43	4700	4.6	no	no	no	good	no	no	notckd
119	4.5	15.2	52	4300	5.7	no	no	no	good	no	no	notckd
120	3.7	15	52	6300	5.3	no	no	no	good	no	no	notckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 121-160

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
121	53	90	1.02	2	0	abnormal	abnormal	present	notpresent	70	107	7.2	114
122	53	90	1.015	0	0	?	normal	notpresent	notpresent	?	38	2.2	?
123	53	100	1.01	1	3	abnormal	normal	notpresent	notpresent	213	23	1	139
124	53	60	1.025	0	0	normal	normal	notpresent	notpresent	116	26	1	146
125	54	80	1.01	3	0	abnormal	abnormal	notpresent	notpresent	207	77	6.3	134
126	54	100	1.015	3	0	?	normal	present	notpresent	162	66	1.6	136
127	54	60	?	3	?	?	?	notpresent	notpresent	125	21	1.3	137
128	54	70	?	?	?	?	?	notpresent	notpresent	111	146	7.5	141
129	54	60	1.015	3	2	?	abnormal	notpresent	notpresent	352	137	3.3	133
130	54	90	1.025	1	0	normal	abnormal	notpresent	notpresent	150	18	1.2	140
131	54	70	1.02	0	0	?	?	?	?	76	28	0.6	146
132	55	90	?	?	?	?	?	notpresent	notpresent	143	88	2	?
133	55	70	1.01	3	0	?	normal	notpresent	notpresent	99	25	1.2	?
134	55	90	1.01	2	1	abnormal	abnormal	notpresent	notpresent	273	235	14.2	132
135	55	80	1.01	3	1	normal	abnormal	present	present	214	73	3.9	137
136	55	80	1.02	0	0	normal	normal	notpresent	notpresent	133	17	1.2	135
137	55	80	1.025	0	0	normal	normal	notpresent	notpresent	130	50	1.2	147
138	55	70	1.02	0	0	normal	normal	notpresent	notpresent	107	26	1.1	?
139	55	80	1.02	0	0	normal	normal	notpresent	notpresent	104	28	0.9	142
140	55	80	1.02	0	0	normal	normal	notpresent	notpresent	140	49	0.5	150
141	56	80	1.01	1	0	?	normal	notpresent	notpresent	165	55	1.8	?
142	56	80	?	?	?	?	?	notpresent	notpresent	415	37	1.9	?
143	56	180	?	0	4	?	abnormal	notpresent	notpresent	298	24	1.2	139
144	56	80	1.025	0	0	normal	normal	notpresent	notpresent	139	15	1.2	135
145	56	70	1.025	0	0	normal	normal	notpresent	notpresent	70	46	1.2	135
146	56	60	1.025	0	0	normal	normal	notpresent	notpresent	132	18	1.1	147
147	57	90	1.015	5	0	abnormal	abnormal	notpresent	present	?	322	13	126
148	57	80	1.015	0	0	?	normal	notpresent	notpresent	120	48	1.6	?
149	57	70	1.015	1	0	?	abnormal	notpresent	notpresent	165	45	1.5	140
150	57	80	1.025	0	0	normal	normal	notpresent	notpresent	119	17	1.2	135
151	57	60	1.02	0	0	normal	normal	notpresent	notpresent	105	49	1.2	150
152	57	60	1.02	0	0	normal	normal	notpresent	notpresent	132	18	1.1	150
153	57	80	1.02	0	0	normal	normal	notpresent	notpresent	133	48	1.2	147
154	58	110	1.01	4	0	?	normal	notpresent	notpresent	251	52	2.2	?
155	58	70	1.02	0	0	normal	normal	notpresent	notpresent	102	48	1.2	139
156	58	80	1.02	0	0	normal	normal	notpresent	notpresent	100	50	1.2	140
157	58	70	1.025	0	0	normal	normal	notpresent	notpresent	88	16	1.1	147
158	58	80	1.025	0	0	normal	normal	notpresent	notpresent	131	18	1.1	141
159	59	70	1.01	3	0	normal	abnormal	notpresent	notpresent	76	186	15	135
160	59	100	1.02	4	2	normal	normal	notpresent	notpresent	252	40	3.2	137

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 121-160

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
121	3.7	9.5	29	12100	3.7	yes	yes	no	poor	no	yes	ckd
122	?	10.9	34	4300	3.7	no	no	no	poor	no	yes	ckd
123	4	?	?	?	?	no	yes	no	good	no	no	ckd
124	4.9	15.8	45	7700	5.2	?	?	?	good	no	no	notckd
125	4.8	9.7	28	?	?	yes	yes	no	poor	yes	no	ckd
126	4.4	10.3	33	?	?	yes	yes	no	poor	yes	no	ckd
127	3.4	15	46	?	?	yes	yes	no	good	yes	no	ckd
128	4.7	11	35	8600	4.6	no	no	no	good	no	no	ckd
129	4.5	11.3	31	5800	3.6	yes	yes	yes	poor	yes	no	ckd
130	4.2	?	?	?	?	no	no	no	poor	yes	yes	ckd
131	3.5	14.8	52	8400	5.9	no	no	no	good	no	no	notckd
132	?	?	?	?	?	yes	yes	no	poor	yes	no	ckd
133	?	11.4	?	?	?	no	no	no	poor	yes	no	ckd
134	3.4	8.3	22	14600	2.9	yes	yes	no	poor	yes	yes	ckd
135	4.9	10.9	34	7400	3.7	yes	yes	no	good	yes	no	ckd
136	4.8	13.2	41	6800	5.3	no	no	no	good	no	no	notckd
137	5	15.5	41	9100	6	no	no	no	good	no	no	notckd
138	?	17	50	6700	6.1	no	no	no	good	no	no	notckd
139	4.8	17.3	52	8200	4.8	no	no	no	good	no	no	notckd
140	4.9	15.7	47	6700	4.9	no	no	no	good	no	no	notckd
141	?	13.5	40	11800	5	yes	yes	no	poor	yes	no	ckd
142	?	?	?	?	?	no	yes	no	good	no	no	ckd
143	3.9	11.2	32	10400	4.2	yes	yes	no	poor	yes	no	ckd
144	5	14.8	42	5600	5.5	no	no	no	good	no	no	notckd
145	4.9	15.9	50	11000	5.1	?	?	?	good	no	no	notckd
146	4.7	13.7	45	7500	5.6	no	no	no	good	no	no	notckd
147	4.8	8	24	4200	3.3	yes	yes	yes	poor	yes	yes	ckd
148	?	11.3	36	7200	3.8	yes	yes	no	good	no	no	ckd
149	3.3	10.4	31	4200	3.9	no	no	no	good	no	no	ckd
150	4.7	15.4	42	6200	6.2	no	no	no	good	no	no	notckd
151	4.7	15.7	44	10400	6.2	no	no	no	good	no	no	notckd
152	4.7	15.4	42	11000	4.5	no	no	no	good	no	no	notckd
153	4.3	14.8	46	6600	5.5	no	no	no	good	no	no	notckd
154	?	?	?	13200	4.7	yes	yes	no	good	no	no	ckd
155	4.3	15	40	8100	4.9	no	no	no	good	no	no	notckd
156	3.5	14	50	6700	6.5	no	no	no	good	no	no	notckd
157	3.5	16.4	53	9100	5.2	no	no	no	good	no	no	notckd
158	3.5	15.8	53	6800	6.1	no	no	no	good	no	no	notckd
159	7.6	7.1	22	3800	2.1	yes	no	no	poor	yes	yes	ckd
160	4.7	11.2	30	26400	3.9	yes	yes	no	poor	yes	no	ckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 161-200

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
161	59	100	1.015	4	2	normal	normal	notpresent	notpresent	255	132	12.8	135
162	59	70	1.01	1	3	abnormal	abnormal	notpresent	notpresent	424	55	1.7	138
163	59	70	1.025	0	0	normal	normal	notpresent	notpresent	130	39	0.7	147
164	59	60	1.02	0	0	normal	normal	notpresent	notpresent	113	23	1.1	139
165	60	100	1.025	0	3	?	normal	notpresent	notpresent	263	27	1.3	135
166	60	70	1.01	2	0	normal	abnormal	present	notpresent	144	72	3	?
167	60	80	1.01	1	1	?	normal	notpresent	notpresent	172	32	2.7	?
168	60	70	1.01	0	0	?	normal	notpresent	notpresent	140	27	1.2	?
169	60	90	1.01	2	0	abnormal	normal	notpresent	notpresent	105	53	2.3	136
170	60	70	1.01	1	0	?	normal	notpresent	notpresent	109	96	3.9	135
171	60	90	1.01	3	5	abnormal	normal	notpresent	present	490	95	2.7	131
172	60	70	?	?	?	?	?	notpresent	notpresent	124	52	2.5	?
173	60	90	?	?	?	?	?	notpresent	notpresent	269	51	2.8	138
174	60	80	1.025	0	0	normal	normal	notpresent	notpresent	131	10	0.5	146
175	60	70	1.02	0	0	normal	normal	notpresent	notpresent	?	?	?	?
176	60	60	1.02	0	0	normal	normal	notpresent	notpresent	134	45	0.5	139
177	60	80	1.025	0	0	normal	normal	notpresent	notpresent	81	15	0.5	141
178	61	80	1.015	2	0	abnormal	abnormal	notpresent	notpresent	173	148	3.9	135
179	61	60	1.025	0	0	?	normal	notpresent	notpresent	108	75	1.9	141
180	61	90	1.015	0	2	?	normal	notpresent	notpresent	?	?	?	?
181	61	80	1.02	0	0	?	normal	notpresent	notpresent	131	23	0.8	140
182	61	70	1.025	0	0	normal	normal	notpresent	notpresent	120	29	0.7	137
183	62	70	1.025	3	0	normal	abnormal	notpresent	notpresent	122	42	1.7	136
184	62	80	1.01	1	2	?	?	notpresent	notpresent	309	113	2.9	130
185	62	90	1.02	2	1	?	normal	notpresent	notpresent	169	48	2.4	138
186	62	80	1.02	0	0	normal	normal	notpresent	notpresent	132	34	0.8	147
187	62	80	1.025	0	0	normal	normal	notpresent	notpresent	78	45	0.6	138
188	63	70	1.01	3	0	abnormal	abnormal	present	notpresent	380	60	2.7	131
189	63	80	1.01	2	2	normal	?	notpresent	notpresent	?	?	3.4	136
190	63	100	1.01	2	2	normal	normal	notpresent	present	280	35	3.2	143
191	63	90	1.015	0	0	?	normal	notpresent	notpresent	123	19	2	142
192	63	70	1.025	0	0	normal	normal	notpresent	notpresent	130	37	0.9	150
193	63	70	1.02	0	0	normal	normal	notpresent	notpresent	113	25	0.6	146
194	64	90	1.01	3	3	?	abnormal	present	notpresent	?	35	1.3	?
195	64	60	1.01	4	1	abnormal	abnormal	notpresent	present	239	58	4.3	137
196	64	70	?	?	?	?	?	notpresent	notpresent	113	94	7.3	137
197	64	70	1.01	0	0	?	normal	notpresent	notpresent	107	15	?	?
198	64	100	1.015	4	2	abnormal	abnormal	notpresent	present	163	54	7.2	140
199	64	60	1.02	0	0	normal	normal	notpresent	notpresent	106	27	0.7	150
200	65	90	1.02	2	1	abnormal	normal	notpresent	notpresent	270	39	2	?

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 161-200

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
161	5.7	7.3	20	9800	3.9	yes	yes	yes	good	no	yes	ckd
162	4.5	12.6	37	10200	4.1	yes	yes	yes	good	no	no	ckd
163	4.7	13.5	46	6700	4.5	no	no	no	good	no	no	notckd
164	3.5	15.3	54	6500	4.9	no	no	no	good	no	no	notckd
165	4.3	12.7	37	11400	4.3	yes	yes	yes	good	no	no	ckd
166	?	9.7	29	21600	3.5	yes	yes	no	poor	no	yes	ckd
167	?	11.2	36	?	?	no	yes	yes	poor	no	no	ckd
168	?	?	?	?	?	no	no	no	good	no	no	ckd
169	5.2	11.1	33	10500	4.1	no	no	no	good	no	no	ckd
170	4	13.8	41	?	?	yes	no	no	good	no	no	ckd
171	3.8	11.5	35	12000	4.5	yes	yes	no	good	no	no	ckd
172	?	?	?	?	?	yes	no	no	good	no	no	ckd
173	3.7	11.5	35	?	?	yes	yes	yes	good	yes	no	ckd
174	5	14.5	41	10700	5.1	no	no	no	good	no	no	notckd
175	?	16.4	43	10800	5.7	no	no	no	good	no	no	notckd
176	4.8	14.2	48	10700	5.6	no	no	no	good	no	no	notckd
177	3.6	15	46	10500	5.3	no	no	no	good	no	no	notckd
178	5.2	7.7	24	9200	3.2	yes	yes	yes	poor	yes	yes	ckd
179	5.2	9.9	29	8400	3.7	yes	yes	no	good	no	yes	ckd
180	?	?	?	9800	?	no	yes	no	poor	no	yes	ckd
181	4.1	11.3	35	?	?	no	no	no	good	no	no	ckd
182	3.5	17.4	52	7000	5.3	no	no	no	good	no	no	notckd
183	4.7	12.6	39	7900	3.9	yes	yes	no	good	no	no	ckd
184	2.5	10.6	34	12800	4.9	no	no	no	good	no	no	ckd
185	2.9	13.4	47	11000	6.1	yes	no	no	good	no	no	ckd
186	3.5	17.8	44	4700	4.5	no	no	no	good	no	no	notckd
187	3.5	16.1	50	5400	5.7	no	no	no	good	no	no	notckd
188	4.2	10.8	32	4500	3.8	yes	yes	no	poor	yes	no	ckd
189	4.2	13	40	9800	4.2	yes	no	yes	good	no	no	ckd
190	3.5	13	40	9800	4.2	yes	no	yes	good	no	no	ckd
191	3.8	11.7	34	11400	4.7	no	no	no	good	no	no	ckd
192	5	13.4	41	7300	4.7	no	no	no	good	no	no	notckd
193	4.9	16.5	52	8000	5.1	no	no	no	good	no	no	notckd
194	?	10.3	?	?	?	yes	yes	no	good	yes	no	ckd
195	5.4	9.5	29	7500	3.4	yes	yes	no	poor	yes	no	ckd
196	4.3	7.9	21	?	?	yes	yes	yes	good	yes	yes	ckd
197	?	12.8	38	?	?	no	no	no	good	no	no	ckd
198	4.6	7.9	26	7500	3.4	yes	yes	no	good	yes	no	ckd
199	3.3	14.4	42	8100	4.7	no	no	no	good	no	no	notckd
200	?	12	36	9800	4.9	yes	yes	no	poor	no	yes	ckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 201-240

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
201	65	70	1.01	0	0	?	normal	notpresent	notpresent	93	66	1.6	137
202	65	60	1.015	1	0	?	normal	notpresent	notpresent	91	51	2.2	132
203	65	80	1.015	0	0	?	normal	notpresent	notpresent	115	32	11.5	139
204	65	80	1.01	3	3	?	normal	notpresent	notpresent	294	71	4.4	128
205	65	70	1.015	4	4	?	normal	present	notpresent	307	28	1.5	?
206	65	90	1.01	4	2	normal	normal	notpresent	notpresent	172	82	13.5	145
207	65	60	1.01	2	0	normal	abnormal	present	notpresent	192	17	1.7	130
208	65	60	1.02	0	0	normal	normal	notpresent	notpresent	109	39	1	144
209	65	70	1.025	0	0	?	?	notpresent	notpresent	85	20	1	142
210	66	90	1.015	2	0	?	normal	notpresent	present	153	76	3.3	?
211	66	70	1.015	2	5	?	normal	notpresent	notpresent	447	41	1.7	131
212	66	70	1.02	0	0	normal	normal	notpresent	notpresent	94	19	0.7	135
213	66	70	1.025	0	0	normal	normal	notpresent	notpresent	107	16	1.1	140
214	67	90	1.02	1	0	?	abnormal	present	notpresent	141	66	3.2	138
215	67	80	1.01	1	3	normal	abnormal	notpresent	notpresent	182	391	32	163
216	67	70	1.02	2	0	abnormal	normal	notpresent	notpresent	150	55	1.6	131
217	67	70	1.01	1	0	?	normal	notpresent	notpresent	102	48	3.2	137
218	67	70	1.01	1	0	normal	normal	notpresent	notpresent	?	106	6	137
219	67	80	?	?	?	?	?	notpresent	notpresent	341	37	1.5	?
220	67	80	1.025	0	0	normal	normal	notpresent	notpresent	99	40	0.5	?
221	68	70	1.015	3	1	?	normal	present	notpresent	208	72	2.1	138
222	68	70	?	?	?	?	?	notpresent	notpresent	98	86	4.6	135
223	68	70	1.005	1	0	abnormal	abnormal	present	notpresent	?	28	1.4	?
224	68	90	1.01	0	0	?	normal	notpresent	notpresent	238	57	2.5	?
225	68	60	1.025	0	0	normal	normal	notpresent	notpresent	125	41	1.1	139
226	69	80	1.02	3	0	abnormal	normal	notpresent	notpresent	?	103	4.1	132
227	69	70	1.01	0	4	?	normal	notpresent	notpresent	256	40	1.2	142
228	69	60	?	?	?	?	?	notpresent	notpresent	171	26	48.1	?
229	69	70	1.01	4	3	normal	abnormal	present	present	214	96	6.3	120
230	69	70	1.02	0	0	normal	normal	notpresent	notpresent	83	42	1.2	139
231	70	80	?	?	?	?	?	notpresent	notpresent	158	85	3.2	141
232	70	70	1.015	2	?	?	?	notpresent	notpresent	?	46	1.5	?
233	70	100	1.005	1	0	normal	abnormal	present	notpresent	169	47	2.9	?
234	70	90	1.015	0	0	?	normal	notpresent	notpresent	144	125	4	136
235	70	100	1.015	4	0	normal	normal	notpresent	notpresent	118	125	5.3	136
236	70	90	1.02	2	1	abnormal	abnormal	notpresent	present	184	98.6	3.3	138
237	70	60	?	?	?	?	?	notpresent	notpresent	120	40	0.5	140
238	70	80	1.02	0	0	normal	normal	notpresent	notpresent	74	41	0.5	143
239	71	70	1.01	3	0	normal	abnormal	present	present	219	82	3.6	133
240	71	90	1.015	2	0	?	abnormal	present	present	88	80	4.4	139

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 201-240

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
201	4.5	11.6	36	11900	3.9	no	yes	no	good	no	no	ckd
202	3.8	10	32	9100	4	yes	yes	no	poor	yes	no	ckd
203	4	14.1	42	6800	5.2	no	no	no	good	no	no	ckd
204	5.4	10	32	9000	3.9	yes	yes	yes	good	no	no	ckd
205	?	11	39	6700	?	yes	yes	no	good	no	no	ckd
206	6.3	8.8	31	?	?	yes	yes	no	good	yes	yes	ckd
207	4.3	?	?	9500	?	yes	yes	no	poor	no	no	ckd
208	3.5	13.9	48	9600	4.8	no	no	no	good	no	no	notckd
209	4.8	16.1	43	9600	4.5	no	no	no	good	no	no	notckd
210	?	?	?	?	?	no	no	no	poor	no	no	ckd
211	3.9	12.5	33	9600	4.4	yes	yes	no	good	no	no	ckd
212	3.9	16	41	5300	5.9	no	no	no	good	no	no	notckd
213	3.6	13.6	42	11000	4.9	no	no	no	good	no	no	notckd
214	6.6	?	?	?	?	yes	no	no	good	no	no	ckd
215	3.9	?	?	?	?	no	no	no	good	yes	no	ckd
216	4.8	?	?	?	?	yes	yes	no	good	yes	no	ckd
217	5	11.9	34	7100	3.7	yes	yes	no	good	yes	no	ckd
218	4.9	6.1	19	6500	?	yes	no	no	good	no	yes	ckd
219	?	12.3	41	6900	4.9	yes	yes	no	good	no	yes	ckd
220	?	17.8	44	5900	5.2	no	no	no	good	no	no	notckd
221	5.8	9.7	28	12200	3.4	yes	yes	yes	poor	yes	no	ckd
222	3.4	9.8	?	?	?	yes	yes	yes	poor	yes	no	ckd
223	?	12.9	38	?	?	no	no	yes	good	no	no	ckd
224	?	9.8	28	8000	3.3	yes	yes	no	poor	no	no	ckd
225	3.8	17.4	50	6700	6.1	no	no	no	good	no	no	notckd
226	5.9	12.5	?	?	?	yes	no	no	good	no	no	ckd
227	5.6	?	?	?	?	no	no	no	good	no	no	ckd
228	?	?	?	?	?	yes	no	no	poor	no	no	ckd
229	3.9	9.4	28	11500	3.3	yes	yes	yes	good	yes	yes	ckd
230	3.7	16.2	50	9300	5.4	no	no	no	good	no	no	notckd
231	3.5	10.1	30	?	?	yes	no	no	good	yes	no	ckd
232	?	9.9	?	?	?	no	yes	no	poor	yes	no	ckd
233	?	11.1	32	5800	5	yes	yes	no	poor	no	no	ckd
234	4.6	12	37	8200	4.5	yes	yes	no	poor	yes	no	ckd
235	4.9	12	37	8400	8	yes	no	no	good	no	no	ckd
236	3.9	5.8	?	?	?	yes	yes	yes	poor	no	no	ckd
237	4.6	16	43	4500	4.9	no	no	no	good	no	no	notckd
238	4.5	15.1	48	9700	5.6	no	no	no	good	no	no	notckd
239	4.4	10.4	33	5600	3.6	yes	yes	yes	good	no	no	ckd
240	5.7	11.3	33	10700	3.9	no	no	no	good	no	no	ckd

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ ระเบียบที่ 241-280

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
241	71	60	1.015	4	0	normal	normal	notpresent	notpresent	118	125	5.3	136
242	71	60	1.02	3	2	normal	normal	present	notpresent	424	48	1.5	132
243	71	90	1.01	0	3	?	normal	notpresent	notpresent	303	30	1.3	136
244	71	70	1.02	0	0	normal	normal	notpresent	notpresent	81	18	0.8	145
245	71	60	1.025	0	0	normal	normal	notpresent	notpresent	?	?	0.9	140
246	71	70	1.025	0	0	?	?	notpresent	notpresent	79	47	0.5	142
247	72	80	?	?	?	?	?	notpresent	notpresent	137	65	3.4	141
248	72	90	1.025	1	3	?	normal	notpresent	notpresent	323	40	2.2	137
249	72	90	1.01	2	0	?	abnormal	present	notpresent	124	53	2.3	?
250	72	60	1.02	0	0	normal	normal	notpresent	notpresent	109	26	0.9	150
251	73	90	1.015	3	0	?	abnormal	present	notpresent	107	33	1.5	141
252	73	80	1.02	2	0	abnormal	abnormal	notpresent	notpresent	253	142	4.6	138
253	73	90	1.01	1	4	abnormal	abnormal	present	notpresent	234	56	1.9	?
254	73	60	1.02	0	0	normal	normal	notpresent	notpresent	127	48	0.5	150
255	74	80	1.01	0	0	?	normal	notpresent	notpresent	132	98	2.8	133
256	74	60	?	?	?	?	?	notpresent	notpresent	108	68	1.8	?
257	74	60	1.02	0	0	normal	normal	notpresent	notpresent	88	50	0.6	147
258	75	80	1.015	0	0	?	normal	notpresent	notpresent	156	45	2.4	140
259	75	70	?	1	3	?	?	notpresent	notpresent	123	31	1.4	?
260	75	70	1.025	1	0	?	normal	notpresent	notpresent	158	49	1.4	135
261	75	60	1.02	0	0	normal	normal	notpresent	notpresent	110	50	0.7	135
262	75	70	1.02	0	0	normal	normal	notpresent	notpresent	107	48	0.8	144
263	76	70	1.015	1	0	normal	normal	notpresent	notpresent	92	29	1.8	133
264	76	70	1.015	2	0	normal	abnormal	present	notpresent	226	217	10.2	?
265	78	60	?	?	?	?	?	notpresent	notpresent	114	74	2.9	135
266	79	80	1.025	0	0	normal	normal	notpresent	notpresent	111	44	1.2	146
267	80	70	1.01	2	?	?	abnormal	notpresent	notpresent	?	49	1.2	?
268	80	70	1.015	2	2	?	normal	notpresent	notpresent	141	53	2.2	?
269	80	80	1.025	0	0	normal	normal	notpresent	notpresent	119	46	0.7	141
270	81	60	?	?	?	?	?	notpresent	notpresent	148	39	2.1	147
271	83	70	1.02	3	0	normal	normal	notpresent	notpresent	102	60	2.6	115
272	90	90	1.025	1	0	?	normal	notpresent	notpresent	139	89	3	140
273	?	70	?	?	?	?	?	notpresent	notpresent	93	155	7.3	132
274	?	100	1.015	2	0	abnormal	abnormal	notpresent	notpresent	129	107	6.7	132
275	?	60	1.015	3	0	abnormal	abnormal	notpresent	notpresent	?	34	1.2	?
276	?	70	1.015	4	0	abnormal	normal	notpresent	notpresent	104	16	0.5	?
277	?	70	1.02	0	0	?	?	notpresent	notpresent	219	36	1.3	139
278	?	70	1.01	3	0	normal	normal	notpresent	notpresent	110	115	6	134
279	?	90	?	?	?	?	?	notpresent	notpresent	207	80	6.8	142
280	?	80	?	?	?	?	?	notpresent	notpresent	100	49	1	140

ตารางที่ ก-2 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการสร้างตัวแบบ(ต่อ) ระเบียบที่ 241-280

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
241	4.9	11.4	35	15200	4.3	yes	yes	no	poor	yes	no	ckd
242	4	10.9	31	?	?	yes	yes	yes	good	no	no	ckd
243	4.1	13	38	9200	4.6	yes	yes	no	good	no	no	ckd
244	5	14.7	44	9800	6	no	no	no	good	no	no	notckd
245	4.8	15.2	42	7700	5.5	no	no	no	good	no	no	notckd
246	4.8	16.6	40	5800	5.9	no	no	no	good	no	no	notckd
247	4.7	9.7	28	6900	2.5	yes	yes	no	poor	no	yes	ckd
248	5.3	12.6	?	?	?	no	yes	yes	poor	no	no	ckd
249	?	11.9	39	?	?	no	no	no	good	no	no	ckd
250	4.9	15	52	10500	5.5	no	no	no	good	no	no	notckd
251	4.6	10.1	30	7800	4	no	no	no	poor	no	no	ckd
252	5.8	10.5	33	7200	4.3	yes	yes	yes	good	no	no	ckd
253	?	10.3	28	?	?	no	yes	no	good	no	no	ckd
254	3.5	15.1	52	11000	4.7	no	no	no	good	no	no	notckd
255	5	10.8	31	9400	3.8	yes	yes	no	good	no	no	ckd
256	?	?	?	?	?	yes	yes	no	good	no	no	ckd
257	3.7	17.2	53	6000	4.5	no	no	no	good	no	no	notckd
258	3.4	11.6	35	10300	4	yes	yes	no	poor	no	no	ckd
259	?	?	?	?	?	no	yes	no	good	no	no	ckd
260	4.7	11.1	?	?	?	yes	no	no	poor	yes	no	ckd
261	5	14.3	40	8300	5.8	no	no	no	?	?	?	notckd
262	3.5	13.6	46	10300	4.8	no	no	no	good	no	no	notckd
263	3.9	10.3	32	?	?	yes	no	no	good	no	no	ckd
264	?	10.2	36	12700	4.2	yes	no	no	poor	yes	yes	ckd
265	5.9	8	24	?	?	no	yes	no	good	no	yes	ckd
266	3.6	16.3	40	8000	6.4	no	no	no	good	no	no	notckd
267	?	?	?	?	?	yes	yes	no	good	no	no	ckd
268	?	12.7	40	9600	?	yes	yes	no	poor	yes	no	ckd
269	4.9	13.9	49	5100	5	no	no	no	good	no	no	notckd
270	4.2	10.9	35	9400	2.4	yes	yes	yes	poor	yes	no	ckd
271	5.7	8.7	26	12800	3.1	yes	no	no	poor	no	yes	ckd
272	4.1	12	37	7900	3.9	yes	yes	no	good	no	no	ckd
273	4.9	?	?	?	?	yes	yes	no	good	no	no	ckd
274	4.4	4.8	14	6300	?	yes	no	no	good	yes	yes	ckd
275	?	10.8	33	?	?	no	no	no	good	no	no	ckd
276	?	?	?	?	?	no	no	no	good	yes	no	ckd
277	3.7	12.5	37	9800	4.4	no	no	no	good	no	no	ckd
278	2.7	9.1	26	9200	3.4	yes	yes	no	poor	no	no	ckd
279	5.5	8.5	?	?	?	yes	yes	no	good	no	yes	ckd
280	5	16.3	53	8500	4.9	no	no	no	good	no	no	notckd

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ ระเบียบที่ 1-40

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
1	6	60	1.01	4	0	abnormal	abnormal	notpresent	present	94	67	1	135
2	8	50	1.02	4	0	normal	normal	notpresent	notpresent	?	46	1	135
3	11	80	1.01	3	0	?	normal	notpresent	notpresent	?	17	0.8	?
4	19	70	1.02	0	0	?	normal	notpresent	notpresent	?	?	?	?
5	22	60	1.025	0	0	normal	normal	notpresent	notpresent	97	18	1.2	138
6	23	80	1.025	0	0	normal	normal	notpresent	notpresent	70	36	1	150
7	23	80	1.02	0	0	normal	normal	notpresent	notpresent	99	46	1.2	142
8	23	60	1.02	0	0	normal	normal	notpresent	notpresent	95	24	0.8	145
9	25	80	1.025	0	0	normal	normal	notpresent	notpresent	121	19	1.2	142
10	26	70	1.015	0	4	?	normal	notpresent	notpresent	250	20	1.1	?
11	27	60	?	?	?	?	?	notpresent	notpresent	76	44	3.9	127
12	30	80	1.02	0	0	normal	normal	notpresent	notpresent	89	42	0.5	139
13	30	60	1.025	0	0	normal	normal	notpresent	notpresent	120	31	0.8	150
14	30	80	1.025	0	0	normal	normal	notpresent	notpresent	82	42	0.7	146
15	32	70	1.025	0	0	normal	normal	notpresent	notpresent	100	29	1.1	142
16	32	60	1.025	0	0	normal	normal	notpresent	notpresent	102	17	0.4	147
17	33	90	1.015	0	0	?	normal	notpresent	notpresent	92	19	0.8	?
18	33	60	1.025	0	0	normal	normal	notpresent	notpresent	80	25	0.9	146
19	33	60	?	?	?	normal	normal	notpresent	notpresent	130	41	0.9	141
20	33	80	1.025	0	0	normal	normal	notpresent	notpresent	89	19	1.1	144
21	34	70	1.015	4	0	abnormal	abnormal	notpresent	notpresent	153	22	0.9	133
22	34	60	1.02	0	0	?	normal	notpresent	notpresent	117	28	2.2	138
23	34	80	1.025	0	0	normal	normal	notpresent	notpresent	121	27	1.2	144
24	34	60	1.02	0	0	normal	normal	notpresent	notpresent	91	49	1.2	135
25	35	80	1.02	0	0	normal	normal	notpresent	notpresent	104	31	1.2	135
26	35	70	1.025	0	0	?	?	notpresent	notpresent	82	36	1.1	150
27	38	80	1.02	0	0	normal	normal	notpresent	notpresent	99	19	0.5	147
28	39	70	1.025	0	0	normal	normal	notpresent	notpresent	124	22	0.6	137
29	39	70	1.025	0	0	normal	normal	notpresent	notpresent	131	46	0.6	145
30	41	80	1.02	0	0	normal	normal	notpresent	notpresent	122	25	0.8	138
31	42	80	1.02	0	0	normal	normal	notpresent	notpresent	98	20	0.5	140
32	42	70	1.025	0	0	normal	normal	notpresent	notpresent	75	31	1.2	141
33	43	80	1.025	0	0	normal	normal	notpresent	notpresent	81	46	0.6	135
34	43	80	1.025	0	0	normal	normal	notpresent	notpresent	130	30	1.1	143
35	43	80	1.02	0	0	normal	normal	notpresent	notpresent	114	32	1.1	135
36	43	60	1.025	0	0	normal	normal	notpresent	notpresent	108	25	1	144
37	43	60	1.025	0	0	normal	normal	notpresent	notpresent	117	45	0.7	141
38	44	70	?	?	?	?	?	notpresent	notpresent	106	25	0.9	150
39	45	70	1.01	0	0	?	normal	notpresent	notpresent	?	20	0.7	?
40	45	80	1.015	0	0	?	abnormal	notpresent	notpresent	107	15	1	141

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ(ต่อ) ระเบียบที่ 1-40

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	@lass-variable
1	4.9	9.9	30	16700	4.8	no	no	no	poor	no	no	ckd
2	3.8	?	?	?	?	no	no	no	good	yes	no	ckd
3	?	15	45	8600	?	no	no	no	good	no	no	ckd
4	?	11.5	?	6900	?	no	no	no	good	no	no	ckd
5	4.3	13.5	42	7900	6.4	no	no	no	good	no	no	notckd
6	4.6	17	52	9800	5	no	no	no	good	no	no	notckd
7	4	17.7	46	4300	5.5	no	no	no	good	no	no	notckd
8	5	15	52	6300	4.6	no	no	no	good	no	no	notckd
9	4.9	15	48	6900	5.3	no	no	no	good	no	no	notckd
10	?	15.6	52	6900	6	no	yes	no	good	no	no	ckd
11	4.3	?	?	?	?	no	no	no	poor	yes	yes	ckd
12	5	16.7	52	10200	5	no	no	no	good	no	no	notckd
13	4.6	13.4	44	10700	5.8	no	no	no	good	no	no	notckd
14	5	14.9	45	9400	5.9	no	no	no	good	no	no	notckd
15	4.5	14.3	43	6700	5.9	no	no	no	good	no	no	notckd
16	4.7	14.6	41	6800	5.1	no	no	no	good	no	no	notckd
17	?	11.8	34	7000	?	no	no	no	good	no	no	ckd
18	3.5	14.1	48	7800	5.1	no	no	no	good	no	no	notckd
19	4.4	15.5	52	4300	5.8	no	no	no	good	no	no	notckd
20	5	15	40	10300	4.8	no	no	no	good	no	no	notckd
21	3.8	?	?	?	?	no	no	no	good	yes	no	ckd
22	3.8	?	?	?	?	no	no	no	good	yes	no	ckd
23	3.9	13.6	52	9200	6.3	no	no	no	good	no	no	notckd
24	4.5	13.5	48	8600	4.9	no	no	no	good	no	no	notckd
25	5	16.1	45	4300	5.2	no	no	no	good	no	no	notckd
26	3.5	14.5	52	9400	6.1	no	no	no	good	no	no	notckd
27	3.5	13.6	44	7300	6.4	no	no	no	good	no	no	notckd
28	3.8	13.4	43	?	?	no	no	no	good	no	no	notckd
29	5	15.6	41	9400	4.7	no	no	no	good	no	no	notckd
30	5	17.1	41	9100	5.2	no	no	no	good	no	no	notckd
31	3.5	13.9	44	8400	5.5	no	no	no	good	no	no	notckd
32	3.5	16.5	54	7800	6.2	no	no	no	good	no	no	notckd
33	4.9	13.9	48	6900	4.9	no	no	no	good	no	no	notckd
34	5	15.9	45	7800	4.5	no	no	no	good	no	no	notckd
35	3.9	?	42	?	?	no	no	no	good	no	no	notckd
36	5	17.8	43	7200	5.5	no	no	no	good	no	no	notckd
37	4.4	13	54	7400	5.4	no	no	no	good	no	no	notckd
38	3.6	15	50	9600	6.5	no	no	no	good	no	no	notckd
39	?	?	?	?	?	no	no	no	good	yes	no	ckd
40	4.2	11.8	37	10200	4.2	no	no	no	good	no	no	ckd

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ ระเบียบที่ 41-80

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
41	45	70	1.01	2	0	?	normal	notpresent	notpresent	113	93	2.3	?
42	45	80	1.02	0	0	normal	normal	notpresent	notpresent	117	46	1.2	137
43	46	90	1.01	2	0	normal	abnormal	notpresent	notpresent	99	80	2.1	?
44	46	60	1.01	1	0	normal	normal	notpresent	notpresent	163	92	3.3	141
45	46	70	1.025	0	0	normal	normal	notpresent	notpresent	100	47	0.5	142
46	47	70	1.015	2	0	?	normal	notpresent	notpresent	99	46	2.2	138
47	47	80	?	?	?	?	?	notpresent	notpresent	114	87	5.2	139
48	47	80	?	?	?	?	?	notpresent	notpresent	93	33	0.9	144
49	47	80	1.025	0	0	normal	normal	notpresent	notpresent	124	44	1	140
50	48	80	1.025	4	0	normal	abnormal	notpresent	notpresent	95	163	7.7	136
51	48	70	1.015	1	0	normal	normal	notpresent	notpresent	127	19	1	134
52	49	100	1.01	3	0	abnormal	abnormal	notpresent	notpresent	129	158	11.8	122
53	49	80	1.02	0	0	normal	normal	notpresent	notpresent	122	32	1.2	139
54	50	70	1.02	0	0	?	normal	notpresent	notpresent	109	32	1.4	139
55	50	140	?	?	?	?	?	notpresent	notpresent	101	106	6.5	135
56	50	90	1.01	2	0	normal	abnormal	present	present	128	208	9.2	134
57	51	100	1.015	2	0	normal	normal	notpresent	present	93	20	1.6	146
58	51	60	1.02	0	0	?	?	notpresent	notpresent	129	25	1.2	139
59	51	80	1.02	0	0	normal	normal	notpresent	notpresent	94	15	1.2	144
60	52	100	1.015	3	0	normal	abnormal	present	notpresent	138	60	1.9	?
61	53	60	?	?	?	?	?	notpresent	notpresent	91	114	3.25	142
62	54	80	1.02	3	0	?	abnormal	notpresent	notpresent	208	89	5.9	130
63	54	70	?	?	?	?	?	notpresent	notpresent	233	50.1	1.9	?
64	54	120	1.015	0	0	?	normal	notpresent	notpresent	103	18	1.2	?
65	55	80	1.01	0	0	?	normal	notpresent	notpresent	146	?	?	?
66	55	100	1.015	1	4	normal	?	notpresent	notpresent	297	53	2.8	139
67	55	80	1.02	0	0	normal	normal	notpresent	notpresent	118	18	0.9	135
68	56	90	1.015	2	0	abnormal	abnormal	notpresent	notpresent	129	107	6.7	131
69	56	70	1.015	4	1	abnormal	normal	notpresent	notpresent	210	26	1.7	136
70	56	90	1.005	4	3	abnormal	abnormal	notpresent	notpresent	242	132	16.4	140
71	56	90	1.01	4	1	normal	abnormal	present	notpresent	176	309	13.3	124
72	57	80	?	?	?	?	?	notpresent	notpresent	?	111	9.3	124
73	59	100	?	?	?	?	?	notpresent	notpresent	?	96	6.4	?
74	59	80	1.01	1	0	abnormal	normal	notpresent	notpresent	303	35	1.3	122
75	59	70	?	?	?	?	?	notpresent	notpresent	204	34	1.5	124
76	59	50	1.01	3	0	normal	abnormal	notpresent	notpresent	241	191	12	114
77	60	90	1.015	3	0	?	?	notpresent	notpresent	74	25	1.1	142
78	60	90	?	?	?	?	?	notpresent	notpresent	?	180	76	4.5
79	60	100	1.02	2	0	abnormal	abnormal	notpresent	notpresent	140	55	2.5	?
80	60	60	1.01	3	1	normal	abnormal	present	notpresent	288	36	1.7	130

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ(ต่อ) ระเบียบที่ 41-80

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
41	?	7.9	26	5700	?	no	no	yes	good	no	yes	ckd
42	5	16.2	45	8600	5.2	no	no	no	good	no	no	notckd
43	?	11.1	32	9100	4.1	yes	no	no	good	no	no	ckd
44	4	9.8	28	14600	3.2	yes	yes	no	good	no	no	ckd
45	3.5	16.4	43	5700	6.5	no	no	no	good	no	no	notckd
46	4.1	12.6	?	?	?	no	no	no	good	no	no	ckd
47	3.7	12.1	?	?	?	yes	no	no	poor	no	no	ckd
48	4.5	13.3	52	8100	5.2	no	no	no	good	no	no	notckd
49	4.9	14.9	41	7000	5.7	no	no	no	good	no	no	notckd
50	3.8	9.8	32	6900	3.4	yes	no	no	good	no	yes	ckd
51	3.6	?	?	?	?	yes	yes	no	good	no	no	ckd
52	3.2	8.1	24	9600	3.5	yes	yes	no	poor	yes	yes	ckd
53	3.9	17	41	5600	4.9	no	no	no	good	no	no	notckd
54	4.7	?	?	?	?	no	no	no	poor	no	no	ckd
55	4.3	6.2	18	5800	2.3	yes	yes	no	poor	no	yes	ckd
56	4.8	8.2	22	16300	2.7	no	no	no	poor	yes	yes	ckd
57	4.5	?	?	?	?	no	no	no	poor	no	no	ckd
58	5	17.2	40	8100	5.9	no	no	no	good	no	no	notckd
59	3.7	15.5	46	9500	6.4	no	no	no	good	no	no	notckd
60	?	10.8	33	9600	4	yes	yes	no	good	no	yes	ckd
61	4.3	8.6	28	11000	3.8	yes	yes	no	poor	yes	yes	ckd
62	4.9	9.3	?	?	?	yes	yes	no	poor	yes	no	ckd
63	?	11.7	?	?	?	no	yes	no	good	no	no	ckd
64	?	?	?	?	?	no	no	no	good	no	no	ckd
65	?	9.8	?	?	?	no	no	no	good	no	no	ckd
66	4.5	11.2	34	13600	4.4	yes	yes	no	good	no	no	ckd
67	3.6	15.5	43	7200	5.4	no	no	no	good	no	no	notckd
68	4.8	9.1	29	6400	3.4	yes	no	no	good	no	no	ckd
69	3.8	16.1	52	12500	5.6	no	no	no	good	no	no	ckd
70	4.2	8.4	26	?	3	yes	yes	no	poor	yes	yes	ckd
71	6.5	3.1	9	5400	2.1	yes	yes	no	poor	yes	yes	ckd
72	5.3	6.8	?	4300	3	yes	yes	no	good	no	yes	ckd
73	?	6.6	?	?	?	yes	yes	no	good	no	yes	ckd
74	3.5	10.4	35	10900	4.3	no	yes	no	poor	no	no	ckd
75	4.1	9.8	37	6000	?	no	yes	no	good	no	no	ckd
76	2.9	9.6	31	15700	3.8	no	yes	no	good	yes	no	ckd
77	3.2	12.2	39	7800	4.4	yes	yes	no	good	yes	no	ckd
78	?	10.9	32	6200	3.6	yes	yes	yes	good	no	no	ckd
79	?	10.1	29	?	?	yes	no	no	poor	no	no	ckd
80	3	7.9	25	15200	3	yes	no	no	poor	no	yes	ckd

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ ระเบียบที่ 81-120

Rank	age	blood pressure	specific gravity	albumin	sugar	red blood cells	pus cell	pus cell clumps	bacteria	blood glucose random	blood urea nitrogen	serum creatinine	sodium
81	60	80	1.02	0	2	?	?	notpresent	notpresent	?	?	?	?
82	60	50	1.01	0	0	?	normal	notpresent	notpresent	261	58	2.2	113
83	61	90	1.01	1	1	?	normal	notpresent	notpresent	159	39	1.5	133
84	61	80	1.015	0	4	?	normal	notpresent	notpresent	360	19	0.7	137
85	61	70	?	?	?	?	?	notpresent	notpresent	100	28	2.1	?
86	61	70	1.025	0	0	normal	normal	notpresent	notpresent	133	38	1	142
87	62	80	1.01	2	3	normal	normal	notpresent	notpresent	423	53	1.8	?
88	62	60	1.015	1	0	?	abnormal	present	notpresent	100	31	1.6	?
89	62	80	1.015	0	5	?	?	notpresent	notpresent	246	24	1	?
90	62	90	1.015	1	0	?	normal	notpresent	notpresent	94	25	1.1	131
91	62	?	1.015	3	0	abnormal	?	notpresent	notpresent	?	?	?	?
92	63	100	1.01	1	0	?	normal	notpresent	notpresent	78	61	1.8	141
93	64	90	1.015	3	2	?	abnormal	present	notpresent	463	64	2.8	135
94	64	70	1.02	0	0	normal	normal	notpresent	notpresent	97	27	0.7	145
95	65	70	1.01	2	0	?	normal	present	notpresent	112	73	3.3	?
96	65	100	1.015	0	0	?	normal	notpresent	notpresent	90	98	2.5	?
97	65	70	1.02	1	0	abnormal	abnormal	notpresent	notpresent	139	29	1	?
98	65	80	1.015	2	1	normal	normal	present	notpresent	215	133	2.5	?
99	65	80	1.015	0	0	?	normal	notpresent	notpresent	92	37	1.5	140
100	65	80	?	?	?	?	?	notpresent	notpresent	74	66	2	136
101	65	70	1.015	1	0	?	normal	notpresent	notpresent	203	46	1.4	?
102	66	70	1.02	1	0	normal	?	notpresent	notpresent	248	30	1.7	138
103	68	70	1.01	0	0	?	normal	notpresent	notpresent	100	54	24	104
104	68	80	1.01	3	2	normal	abnormal	present	present	157	90	4.1	130
105	68	80	1.015	0	0	?	abnormal	notpresent	notpresent	171	30	1	?
106	69	70	1.01	3	4	normal	abnormal	notpresent	notpresent	264	87	2.7	130
107	70	70	1.01	1	0	normal	?	present	present	171	153	5.2	?
108	72	90	?	?	?	?	?	notpresent	notpresent	308	36	2.5	131
109	72	90	?	?	?	?	?	notpresent	notpresent	84	145	7.1	135
110	72	100	?	?	?	?	?	notpresent	notpresent	201	241	13.4	127
111	73	70	1.005	0	0	normal	normal	notpresent	notpresent	70	32	0.9	125
112	73	100	1.01	3	2	abnormal	abnormal	present	notpresent	295	90	5.6	140
113	73	?	1.01	1	0	?	?	notpresent	notpresent	95	51	1.6	142
114	73	80	1.025	0	0	normal	normal	notpresent	notpresent	118	44	0.7	137
115	76	70	1.015	3	4	normal	abnormal	present	notpresent	?	164	9.7	131
116	76	90	?	?	?	?	normal	notpresent	notpresent	93	155	7.3	132
117	76	90	?	?	?	?	?	notpresent	notpresent	172	46	1.7	141
118	80	70	1.02	0	0	normal	normal	notpresent	notpresent	?	?	?	135
119	82	80	1.01	2	2	normal	?	notpresent	notpresent	140	70	3.4	136
120	?	70	1.01	0	2	?	normal	notpresent	notpresent	220	68	2.8	?

ตารางที่ ก-3 ตัวอย่างข้อมูลผู้ป่วยโรคไตเรื้อรังสำหรับการทดสอบตัวแบบ(ต่อ) ระเบียบที่ 81-120

Rank	potassium	hemoglobin	packed cell volume	white blood cell count	red blood cell count	hypertension	diabetes mellitus	coronary artery disease	appetite	pedal edema	anemia	Class variable
81	?	?	?	?	?	no	yes	no	good	no	no	ckd
82	3	?	?	4200	3.4	yes	no	no	good	no	no	ckd
83	4.9	11.3	34	9600	4	yes	yes	no	poor	no	no	ckd
84	4.4	15.2	44	8300	5.2	yes	yes	no	good	no	no	ckd
85	?	12.6	43	?	?	yes	yes	no	good	no	no	ckd
86	3.6	13.7	47	9200	4.9	no	no	no	good	no	no	notckd
87	?	9.6	31	7500	?	no	yes	no	poor	no	yes	ckd
88	?	10.3	30	5300	3.7	yes	no	yes	good	no	no	ckd
89	?	13.6	40	8500	4.7	yes	yes	no	good	no	no	ckd
90	3.7	?	?	?	?	yes	no	no	good	yes	yes	ckd
91	?	14.3	42	10200	4.8	yes	yes	no	good	no	no	ckd
92	4.4	12.2	36	10500	4.3	no	yes	no	good	no	no	ckd
93	4.1	12.2	40	9800	4.6	yes	yes	no	good	no	yes	ckd
94	4.8	13.8	49	6400	4.8	no	no	no	good	no	no	notckd
95	?	10.9	37	?	?	no	no	no	good	no	no	ckd
96	?	9.1	28	5500	3.6	yes	no	no	good	no	no	ckd
97	?	10.5	32	?	?	yes	no	no	good	yes	no	ckd
98	?	13.2	41	?	?	no	yes	no	good	no	no	ckd
99	5.2	8.8	25	10700	3.2	yes	no	yes	good	yes	no	ckd
100	5.4	9.1	25	?	?	yes	yes	yes	good	yes	no	ckd
101	?	11.4	36	5000	4.1	yes	yes	no	poor	yes	no	ckd
102	5.3	?	?	?	?	yes	yes	no	good	no	no	ckd
103	4	12.4	36	?	?	no	no	no	good	no	no	ckd
104	6.4	5.6	16	11000	2.6	yes	yes	yes	poor	yes	no	ckd
105	?	13.7	43	4900	5.2	no	yes	no	good	no	no	ckd
106	4	12.5	37	9600	4.1	yes	yes	yes	good	yes	no	ckd
107	?	?	?	?	?	no	yes	no	poor	no	no	ckd
108	4.3	?	?	?	?	yes	yes	no	poor	no	no	ckd
109	5.3	?	?	?	?	no	yes	no	good	no	no	ckd
110	4.8	9.4	28	?	?	yes	yes	no	good	no	yes	ckd
111	4	10	29	18900	3.5	yes	yes	no	good	yes	no	ckd
112	2.9	9.2	30	7000	3.2	yes	yes	yes	poor	no	no	ckd
113	3.5	?	?	?	?	no	no	no	good	no	no	ckd
114	3.5	14.8	45	9300	4.7	no	no	no	good	no	no	notckd
115	4.4	10.2	30	11300	3.4	yes	yes	yes	poor	yes	no	ckd
116	4.9	?	?	?	?	yes	yes	yes	poor	no	no	ckd
117	5.5	9.6	30	?	?	yes	yes	no	good	no	yes	ckd
118	4.1	15.3	48	6300	6.1	no	no	no	good	no	no	notckd
119	4.2	13	40	9800	4.2	yes	yes	no	good	no	no	ckd
120	?	8.7	27	?	?	yes	yes	no	good	no	yes	ckd



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลผู้ป่วยโรคไตเรื้อรังเรื้อรัง (Chronic Kidney Disease)

1. วิธีความใกล้เคียงกันมากที่สุด

Classifier

Choose |Bk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last|"

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:50:56 - lazy.IBK

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.02 seconds

=== Summary ===

Correctly Classified Instances	269	96.0714 %
Incorrectly Classified Instances	11	3.9286 %
Kappa statistic	0.9183	
Mean absolute error	0.0426	
Root mean squared error	0.1975	
Relative absolute error	2.9462 %	
Root relative squared error	40.5124 %	
Coverage of cases (0.95 level)	96.0714 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.947	0.018	0.988	0.947	0.967	0.920	0.965	0.968	ckd
	0.982	0.053	0.922	0.982	0.951	0.920	0.965	0.913	notckd
Weighted Avg.	0.961	0.032	0.962	0.961	0.961	0.920	0.965	0.946	

=== Confusion Matrix ===

a b <-- classified as

162	9	a = ckd
2	107	b = notckd

รูปที่ ข-1 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)

Classifier

Choose |Bk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last|"

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:48:46 - lazy.IBK

17:50:52 - msc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.06 seconds

=== Summary ===

Correctly Classified Instances	113	94.1667 %
Incorrectly Classified Instances	7	5.8333 %
Kappa statistic	0.8726	
Mean absolute error	0.0615	
Root mean squared error	0.2407	
Relative absolute error	13.2131 %	
Root relative squared error	50.4277 %	
Coverage of cases (0.95 level)	94.1667 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.937	0.049	0.974	0.937	0.955	0.874	0.944	0.954	ckd
	0.951	0.063	0.886	0.951	0.918	0.874	0.944	0.860	notckd
Weighted Avg.	0.942	0.054	0.944	0.942	0.942	0.874	0.944	0.922	

=== Confusion Matrix ===

a b <-- classified as

74	5	a = ckd
2	39	b = notckd

รูปที่ ข-2 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ในการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose IBK -K 2 -W 0 -A "weka.core.neighboursearch.LinearNNSearch-A" ["weka.core.EuclideanDistance -R first-last"]

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:52:33 - lazy.IBK

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.08 seconds

=== Summary ===

Correctly Classified Instances	269	96.0714 %
Incorrectly Classified Instances	11	3.9286 %
Kappa statistic	0.9183	
Mean absolute error	0.048	
Root mean squared error	0.2067	
Relative absolute error	10.1003 %	
Root relative squared error	42.3901 %	
Coverage of cases (0.95 level)	96.0714 %	
Mean rel. region size (0.95 level)	50.7143 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.947	0.018	0.988	0.947	0.967	0.920	0.964	0.968	ckd
	0.982	0.053	0.922	0.982	0.951	0.920	0.964	0.913	notckd
Weighted Avg.	0.961	0.032	0.962	0.961	0.961	0.920	0.964	0.946	

=== Confusion Matrix ===

a	b	<-- classified as	
162	9	a = ckd	
2	107	b = notckd	

รูปที่ ข-3 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=2)

Classifier

Choose IBK -K 2 -W 0 -A "weka.core.neighboursearch.LinearNNSearch-A" ["weka.core.EuclideanDistance -R first-last"]

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:52:33 - lazy.IBK

17:53:55 - miscInputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	113	94.1667 %
Incorrectly Classified Instances	7	5.8333 %
Kappa statistic	0.8726	
Mean absolute error	0.0724	
Root mean squared error	0.2537	
Relative absolute error	15.5551 %	
Root relative squared error	53.2227 %	
Coverage of cases (0.95 level)	94.1667 %	
Mean rel. region size (0.95 level)	51.25 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.937	0.049	0.974	0.937	0.955	0.874	0.943	0.953	ckd
	0.951	0.063	0.886	0.951	0.918	0.874	0.943	0.860	notckd
Weighted Avg.	0.942	0.054	0.944	0.942	0.942	0.874	0.943	0.921	

=== Confusion Matrix ===

a	b	<-- classified as	
74	5	a = ckd	
2	39	b = notckd	

รูปที่ ข-4 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=2)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose [JBk-K3-W0-A "weka.core.neighboursearch.LinearNNSearch-A {"weka.core.EuclideanDistance-R first-last}"]

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:55:33 - lazy.IBk

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.08 seconds

=== Summary ===

Correctly Classified Instances	266	95	%
Incorrectly Classified Instances	14	5	%
Kappa statistic	0.8966		
Mean absolute error	0.0606		
Root mean squared error	0.2162		
Relative absolute error	12.7339	%	
Root relative squared error	44.3422	%	
Coverage of cases (0.95 level)	96.4286	%	
Mean rel. region size (0.95 level)	52.8571	%	
Total Number of Instances	280		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.930	0.018	0.988	0.930	0.958	0.899	0.967	0.969	ckd
	0.982	0.070	0.899	0.982	0.939	0.899	0.967	0.921	notckd
Weighted Avg.	0.950	0.039	0.993	0.950	0.950	0.899	0.967	0.950	

=== Confusion Matrix ===

a b <-- classified as

159	12	a = ckd
2	107	b = notckd

รูปที่ ข-5 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=3)

Classifier

Choose [JBk-K3-W0-A "weka.core.neighboursearch.LinearNNSearch-A {"weka.core.EuclideanDistance-R first-last}"]

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

17:55:33 - lazy.IBk

17:56:37 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	110	91.6667	%
Incorrectly Classified Instances	10	8.3333	%
Kappa statistic	0.8211		
Mean absolute error	0.0899		
Root mean squared error	0.265		
Relative absolute error	19.318	%	
Root relative squared error	55.5961	%	
Coverage of cases (0.95 level)	95	%	
Mean rel. region size (0.95 level)	54.1667	%	
Total Number of Instances	120		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.899	0.049	0.973	0.899	0.934	0.826	0.947	0.955	ckd
	0.951	0.101	0.830	0.951	0.886	0.826	0.947	0.879	notckd
Weighted Avg.	0.917	0.067	0.924	0.917	0.912	0.826	0.947	0.929	

=== Confusion Matrix ===

a b <-- classified as

71	8	a = ckd
2	39	b = notckd

รูปที่ ข-6 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A {"weka.core.EuclideanDistance -R first-last"}"

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 65

(Nom) class

Result list (right-click for options)

17:58:03 - lazy.IBk

17:58:06 - misc.InputMappedClassifier

Classifier output

----- Predictions on test set -----

inst#	actual	predicted	error	prediction
1	1:ckd	1:ckd	, 0.996	
2	1:ckd	2:notckd	+, 0.996	
3	1:ckd	2:notckd	+, 0.996	
4	1:ckd	1:ckd	, 0.996	
5	2:notckd	2:notckd	, 0.996	
6	2:notckd	2:notckd	, 0.996	
7	2:notckd	2:notckd	, 0.996	
8	2:notckd	2:notckd	, 0.996	
9	2:notckd	2:notckd	, 0.996	
10	1:ckd	1:ckd	, 0.996	
11	1:ckd	1:ckd	, 0.996	
12	2:notckd	2:notckd	, 0.996	
13	2:notckd	2:notckd	, 0.996	
14	2:notckd	2:notckd	, 0.996	
15	2:notckd	2:notckd	, 0.996	
16	2:notckd	2:notckd	, 0.996	
17	1:ckd	2:notckd	+, 0.996	
18	2:notckd	2:notckd	, 0.996	
19	2:notckd	2:notckd	, 0.996	
20	2:notckd	2:notckd	, 0.996	
21	1:ckd	1:ckd	, 0.996	
22	1:ckd	1:ckd	, 0.996	
23	2:notckd	2:notckd	, 0.996	
24	2:notckd	2:notckd	, 0.996	
25	2:notckd	2:notckd	, 0.996	
26	2:notckd	2:notckd	, 0.996	
27	2:notckd	2:notckd	, 0.996	
28	2:notckd	2:notckd	, 0.996	

รูปที่ ข-7 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (k=1)

Classifier

Choose IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A {"weka.core.EuclideanDistance -R first-last"}"

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) class

Result list (right-click for options)

17:58:03 - lazy.IBk

17:58:06 - misc.InputMappedClassifier

Classifier output

29	2:notckd	2:notckd	, 0.996	
30	2:notckd	2:notckd	, 0.996	
31	2:notckd	2:notckd	, 0.996	
32	2:notckd	2:notckd	, 0.996	
33	2:notckd	2:notckd	, 0.996	
34	2:notckd	2:notckd	, 0.996	
35	2:notckd	2:notckd	, 0.996	
36	2:notckd	2:notckd	, 0.996	
37	2:notckd	2:notckd	, 0.996	
38	2:notckd	1:ckd	+, 0.996	
39	1:ckd	1:ckd	, 0.996	
40	1:ckd	1:ckd	, 0.996	
41	1:ckd	1:ckd	, 0.996	
42	2:notckd	2:notckd	, 0.996	
43	1:ckd	1:ckd	, 0.996	
44	1:ckd	1:ckd	, 0.996	
45	2:notckd	2:notckd	, 0.996	
46	1:ckd	1:ckd	, 0.996	
47	1:ckd	1:ckd	, 0.996	
48	2:notckd	1:ckd	+, 0.996	
49	2:notckd	2:notckd	, 0.996	
50	1:ckd	1:ckd	, 0.996	
51	1:ckd	1:ckd	, 0.996	
52	1:ckd	1:ckd	, 0.996	
53	2:notckd	2:notckd	, 0.996	
54	1:ckd	1:ckd	, 0.996	
55	1:ckd	1:ckd	, 0.996	
56	1:ckd	1:ckd	, 0.996	
57	1:ckd	2:notckd	+, 0.996	
58	2:notckd	2:notckd	, 0.996	
59	2:notckd	2:notckd	, 0.996	

รูปที่ ข-7 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose `IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"`

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) class

Result list (right-click for options)

17:58:03 - lazy.IBk

17:58:06 - misc.InputMappedClassifier

Classifier output

```

60, 1:ckd, 1:ckd, , 0.996
61, 1:ckd, 1:ckd, , 0.996
62, 1:ckd, 1:ckd, , 0.996
63, 1:ckd, 1:ckd, , 0.996
64, 1:ckd, 1:ckd, , 0.996
65, 1:ckd, 1:ckd, , 0.996
66, 1:ckd, 1:ckd, , 0.996
67, 2:notckd, 2:notckd, , 0.996
68, 1:ckd, 1:ckd, , 0.996
69, 1:ckd, 1:ckd, , 0.996
70, 1:ckd, 1:ckd, , 0.996
71, 1:ckd, 1:ckd, , 0.996
72, 1:ckd, 1:ckd, , 0.996
73, 1:ckd, 1:ckd, , 0.996
74, 1:ckd, 1:ckd, , 0.996
75, 1:ckd, 1:ckd, , 0.996
76, 1:ckd, 1:ckd, , 0.996
77, 1:ckd, 1:ckd, , 0.996
78, 1:ckd, 1:ckd, , 0.996
79, 1:ckd, 1:ckd, , 0.996
80, 1:ckd, 1:ckd, , 0.996
81, 1:ckd, 1:ckd, , 0.996
82, 1:ckd, 1:ckd, , 0.996
83, 1:ckd, 1:ckd, , 0.996
84, 1:ckd, 1:ckd, , 0.996
85, 1:ckd, 1:ckd, , 0.996
86, 2:notckd, 2:notckd, , 0.996
87, 1:ckd, 1:ckd, , 0.996
88, 1:ckd, 1:ckd, , 0.996
89, 1:ckd, 1:ckd, , 0.996
90, 1:ckd, 1:ckd, , 0.996

```

รูปที่ ข-7 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (ต่อ)

Classifier

Choose `IBk -K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"`

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) class

Result list (right-click for options)

17:58:03 - lazy.IBk

17:58:06 - misc.InputMappedClassifier

Classifier output

```

91, 1:ckd, 1:ckd, , 0.996
92, 1:ckd, 1:ckd, , 0.996
93, 1:ckd, 1:ckd, , 0.996
94, 2:notckd, 2:notckd, , 0.996
95, 1:ckd, 1:ckd, , 0.996
96, 1:ckd, 1:ckd, , 0.996
97, 1:ckd, 1:ckd, , 0.996
98, 1:ckd, 2:notckd, +, 0.996
99, 1:ckd, 1:ckd, , 0.996
100, 1:ckd, 1:ckd, , 0.996
101, 1:ckd, 1:ckd, , 0.996
102, 1:ckd, 1:ckd, , 0.996
103, 1:ckd, 1:ckd, , 0.996
104, 1:ckd, 1:ckd, , 0.996
105, 1:ckd, 1:ckd, , 0.996
106, 1:ckd, 1:ckd, , 0.996
107, 1:ckd, 1:ckd, , 0.996
108, 1:ckd, 1:ckd, , 0.996
109, 1:ckd, 1:ckd, , 0.996
110, 1:ckd, 1:ckd, , 0.996
111, 1:ckd, 1:ckd, , 0.996
112, 1:ckd, 1:ckd, , 0.996
113, 1:ckd, 1:ckd, , 0.996
114, 2:notckd, 2:notckd, , 0.996
115, 1:ckd, 1:ckd, , 0.996
116, 1:ckd, 1:ckd, , 0.996
117, 1:ckd, 1:ckd, , 0.996
118, 2:notckd, 2:notckd, , 0.996
119, 1:ckd, 1:ckd, , 0.996
120, 1:ckd, 1:ckd, , 0.996

```

รูปที่ ข-7 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างที่ 1 การคำนวณค่าความถูกต้อง (Accuracy) และค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) ของการวิเคราะห์ข้อมูลโรคไตเรื้อรัง สำหรับการทำนายตัวแบบด้วยวิธีความใกล้เคียงกันมากที่สุด

จากรูปที่ ข-2

$$\begin{aligned} \text{ค่าความถูกต้อง (Accuracy)} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100 \\ &= \frac{75+39}{75+39+2+5} \times 100 \\ &= 94.17\% \end{aligned}$$

จากรูปที่ ข-7

- 1) กำหนดให้ค่า $y_i = 1$ ได้จากในกรณีที่ค่าจริง (actual) ใน class attribute ของข้อมูลโรคไตเรื้อรังในระเบียนนั้นตรงกันกับค่าทำนาย (predicted)
เช่น ระเบียนที่ 1 ค่าจริง (actual) = 1 : เป็นโรคไตเรื้อรัง
ค่าทำนาย (predicted) = 1 : เป็นโรคไตเรื้อรัง
จะได้ $y_i = y_1 = 1$
- 2) กำหนดให้ค่า $y_i = 1$ ได้จากในกรณีที่ค่าจริง (actual) ใน class attribute ของข้อมูลโรคไตเรื้อรังในระเบียนนั้นตรงกันกับค่าทำนาย (predicted)
เช่น ระเบียนที่ 5 ค่าจริง (actual) = 2 : ไม่เป็นโรคไตเรื้อรัง
ค่าทำนาย (predicted) = 2 : ไม่เป็นโรคไตเรื้อรัง
จะได้ $y_i = y_5 = 1$
- 3) กำหนดให้ค่า $y_i = 0$ ได้จากในกรณีที่ค่าจริง (actual) ใน class attribute ของข้อมูลโรคไตเรื้อรังในระเบียนนั้นไม่ตรงกันกับค่าทำนาย (predicted)
เช่น ระเบียนที่ 17 ค่าจริง (actual) = 1 : เป็นโรคไตเรื้อรัง
ค่าทำนาย (predicted) = 2 : ไม่เป็นโรคไตเรื้อรัง
จะได้ $y_i = y_{17} = 0$
- 4) กำหนดให้ค่า $y_i = 0$ ได้จากในกรณีที่ค่าจริง (actual) ใน class attribute ของข้อมูลโรคไตเรื้อรังในระเบียนนั้นไม่ตรงกันกับค่าทำนาย (predicted)
เช่น ระเบียนที่ 38 ค่าจริง (actual) = 2 : ไม่เป็นโรคไตเรื้อรัง
ค่าทำนาย (predicted) = 1 : เป็นโรคไตเรื้อรัง
จะได้ $y_i = y_{38} = 0$

กำหนดให้ค่า \hat{y}_i ได้จาก ค่าการทำนาย (prediction) ซึ่งอยู่ที่คอลัมน์ขวาสุดของระเบียนนั้น

เอกสารนี้เป็นช่อง Classifier output การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ลำดับที่	y_i	\hat{y}_i	e_i^2
1	1	0.996	0.000016
2	0	0.996	0.992016
3	0	0.996	0.992016
4	1	0.996	0.000016
5	1	0.996	0.000016
6	1	0.996	0.000016
7	1	0.996	0.000016
8	1	0.996	0.000016
9	1	0.996	0.000016
10	1	0.996	0.000016
11	1	0.996	0.000016
12	1	0.996	0.000016
13	1	0.996	0.000016
14	1	0.996	0.000016
15	1	0.996	0.000016
16	1	0.996	0.000016
17	0	0.996	0.992016
18	1	0.996	0.000016
19	1	0.996	0.000016
20	1	0.996	0.000016
21	1	0.996	0.000016
22	1	0.996	0.000016
23	1	0.996	0.000016
24	1	0.996	0.000016
25	1	0.996	0.000016
26	1	0.996	0.000016
27	1	0.996	0.000016
28	1	0.996	0.000016
29	1	0.996	0.000016
30	1	0.996	0.000016
31	1	0.996	0.000016
32	1	0.996	0.000016
33	1	0.996	0.000016
34	1	0.996	0.000016
35	1	0.996	0.000016
36	1	0.996	0.000016
37	1	0.996	0.000016
38	0	0.996	0.992016
39	1	0.996	0.000016
40	1	0.996	0.000016
41	1	0.996	0.000016
42	1	0.996	0.000016
43	1	0.996	0.000016
44	1	0.996	0.000016
45	1	0.996	0.000016
46	1	0.996	0.000016
47	1	0.996	0.000016

ลำดับที่	y_i	\hat{y}_i	e_i^2
48	0	0.996	0.992016
49	1	0.996	0.000016
50	1	0.996	0.000016
51	1	0.996	0.000016
52	1	0.996	0.000016
53	1	0.996	0.000016
54	1	0.996	0.000016
55	1	0.996	0.000016
56	1	0.996	0.000016
57	0	0.996	0.992016
58	1	0.996	0.000016
59	1	0.996	0.000016
60	1	0.996	0.000016
61	1	0.996	0.000016
62	1	0.996	0.000016
63	1	0.996	0.000016
64	1	0.996	0.000016
65	1	0.996	0.000016
66	1	0.996	0.000016
67	1	0.996	0.000016
68	1	0.996	0.000016
69	1	0.996	0.000016
70	1	0.996	0.000016
71	1	0.996	0.000016
72	1	0.996	0.000016
73	1	0.996	0.000016
74	1	0.996	0.000016
75	1	0.996	0.000016
76	1	0.996	0.000016
77	1	0.996	0.000016
78	1	0.996	0.000016
79	1	0.996	0.000016
80	1	0.996	0.000016
81	1	0.996	0.000016
82	1	0.996	0.000016
83	1	0.996	0.000016
84	1	0.996	0.000016
85	1	0.996	0.000016
86	1	0.996	0.000016
87	1	0.996	0.000016
88	1	0.996	0.000016
89	1	0.996	0.000016
90	1	0.996	0.000016
91	1	0.996	0.000016
92	1	0.996	0.000016
93	1	0.996	0.000016
94	1	0.996	0.000016

ลำดับที่	y_i	\hat{y}_i	e_i^2
95	1	0.996	0.000016
96	1	0.996	0.000016
97	1	0.996	0.000016
98	0	0.996	0.992016
99	1	0.996	0.000016
100	1	0.996	0.000016
101	1	0.996	0.000016
102	1	0.996	0.000016
103	1	0.996	0.000016
104	1	0.996	0.000016
105	1	0.996	0.000016
106	1	0.996	0.000016
107	1	0.996	0.000016

108	1	0.996	0.000016
109	1	0.996	0.000016
110	1	0.996	0.000016
111	1	0.996	0.000016
112	1	0.996	0.000016
113	1	0.996	0.000016
114	1	0.996	0.000016
115	1	0.996	0.000016
116	1	0.996	0.000016
117	1	0.996	0.000016
118	1	0.996	0.000016
119	1	0.996	0.000016
120	1	0.996	0.000016
รวม			6.9459

$$\begin{aligned} \text{ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE)} &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \\ &= \frac{6.9459}{120} = 0.0579 \end{aligned}$$

$$\begin{aligned} \text{รากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE)} &= \sqrt{\text{MSE}} \\ &= \sqrt{0.0579} = 0.2406 \end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. วิธีต้นไม้ตัดสินใจ

Classifier

Choose: J48 -C 0.25 -M 2

Test options

Use training set

Supplied test set Set...

Cross-validation Folds: 10

Percentage split %: 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

18:12:56 - trees.J48

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.03 seconds

=== Summary ===

Correctly Classified Instances	280	100 %
Incorrectly Classified Instances	0	0 %
Kappa statistic	1	
Mean absolute error	0.0179	
Root mean squared error	0.0631	
Relative absolute error	3.7679 %	
Root relative squared error	12.9417 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	51.0714 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	ckd
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	notckd
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

=== Confusion Matrix ===

a b <-- classified as

171	0	1	a = ckd
0	109	1	b = notckd

รูปที่ ข-8 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีต้นไม้ตัดสินใจ

Classifier

Choose: J48 -C 0.25 -M 2

Test options

Use training set

Supplied test set Set...

Cross-validation Folds: 10

Percentage split %: 66

More options...

(Nom) class

Start Stop

Result list (right-click for options)

18:12:56 - trees.J48

18:14:20 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.02 seconds

=== Summary ===

Correctly Classified Instances	120	100 %
Incorrectly Classified Instances	0	0 %
Kappa statistic	1	
Mean absolute error	0.0254	
Root mean squared error	0.0771	
Relative absolute error	5.4554 %	
Root relative squared error	16.1636 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	52.9167 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	ckd
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	notckd
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

=== Confusion Matrix ===

a b <-- classified as

79	0	1	a = ckd
0	41	1	b = notckd

รูปที่ ข-9 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีต้นไม้ตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose J48 -C 0.25 -M 2

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) class
Start Stop
Result list (right-click for options)
 18:12:56 - trees.J48
 18:14:20 - misc.InputMappedClassifier

Classifier output
 == Predictions on test set ==
 inst#,actual,predicted,error,prediction
 1,1:ckd,1:ckd,,0.986
 2,1:ckd,1:ckd,,0.986
 3,1:ckd,1:ckd,,0.986
 4,1:ckd,1:ckd,,0.727
 5,2:notckd,2:notckd,,1
 6,2:notckd,2:notckd,,1
 7,2:notckd,2:notckd,,1
 8,2:notckd,2:notckd,,1
 9,2:notckd,2:notckd,,1
 10,1:ckd,1:ckd,,0.986
 11,1:ckd,1:ckd,,0.986
 12,2:notckd,2:notckd,,1
 13,2:notckd,2:notckd,,1
 14,2:notckd,2:notckd,,1
 15,2:notckd,2:notckd,,1
 16,2:notckd,2:notckd,,1
 17,1:ckd,1:ckd,,0.986
 18,2:notckd,2:notckd,,1
 19,2:notckd,2:notckd,,0.79
 20,2:notckd,2:notckd,,1
 21,1:ckd,1:ckd,,0.986
 22,1:ckd,1:ckd,,0.986
 23,2:notckd,2:notckd,,1
 24,2:notckd,2:notckd,,1
 25,2:notckd,2:notckd,,1
 26,2:notckd,2:notckd,,1
 27,2:notckd,2:notckd,,1
 28,2:notckd,2:notckd,,1

รูปที่ ข-10 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีต้นไม้ตัดสินใจ

Classifier
Choose J48 -C 0.25 -M 2

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) class
Start Stop
Result list (right-click for options)
 18:12:56 - trees.J48
 18:14:20 - misc.InputMappedClassifier

Classifier output
 29,2:notckd,2:notckd,,1
 30,2:notckd,2:notckd,,1
 31,2:notckd,2:notckd,,1
 32,2:notckd,2:notckd,,1
 33,2:notckd,2:notckd,,1
 34,2:notckd,2:notckd,,1
 35,2:notckd,2:notckd,,0.946
 36,2:notckd,2:notckd,,1
 37,2:notckd,2:notckd,,1
 38,2:notckd,2:notckd,,0.79
 39,1:ckd,1:ckd,,0.986
 40,1:ckd,1:ckd,,0.986
 41,1:ckd,1:ckd,,0.986
 42,2:notckd,2:notckd,,1
 43,1:ckd,1:ckd,,0.986
 44,1:ckd,1:ckd,,0.986
 45,2:notckd,2:notckd,,1
 46,1:ckd,1:ckd,,0.986
 47,1:ckd,1:ckd,,0.986
 48,2:notckd,2:notckd,,0.79
 49,2:notckd,2:notckd,,1
 50,1:ckd,1:ckd,,0.986
 51,1:ckd,1:ckd,,0.986
 52,1:ckd,1:ckd,,0.986
 53,2:notckd,2:notckd,,1
 54,1:ckd,1:ckd,,0.986
 55,1:ckd,1:ckd,,0.986
 56,1:ckd,1:ckd,,0.986
 57,1:ckd,1:ckd,,0.986
 58,2:notckd,2:notckd,,1
 59,2:notckd,2:notckd,,1

รูปที่ ข-10 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีต้นไม้ตัดสินใจ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose J48 -C 0.25 -M 2

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) class

Result list (right-click for options)

18:12:56 - trees.J48

18:14:20 - misc.InputMappedClassifier

Classifier output

```

60,1:ckd,1:ckd,,0.986
61,1:ckd,1:ckd,,0.986
62,1:ckd,1:ckd,,0.986
63,1:ckd,1:ckd,,0.986
64,1:ckd,1:ckd,,0.986
65,1:ckd,1:ckd,,0.986
66,1:ckd,1:ckd,,0.986
67,2:notckd,2:notckd,,1
68,1:ckd,1:ckd,,0.986
69,1:ckd,1:ckd,,0.986
70,1:ckd,1:ckd,,0.986
71,1:ckd,1:ckd,,0.986
72,1:ckd,1:ckd,,0.986
73,1:ckd,1:ckd,,0.986
74,1:ckd,1:ckd,,0.986
75,1:ckd,1:ckd,,0.986
76,1:ckd,1:ckd,,0.986
77,1:ckd,1:ckd,,0.986
78,1:ckd,1:ckd,,0.986
79,1:ckd,1:ckd,,0.986
80,1:ckd,1:ckd,,0.986
81,1:ckd,1:ckd,,0.509
82,1:ckd,1:ckd,,0.986
83,1:ckd,1:ckd,,0.986
84,1:ckd,1:ckd,,0.986
85,1:ckd,1:ckd,,0.986
86,2:notckd,2:notckd,,1
87,1:ckd,1:ckd,,0.986
88,1:ckd,1:ckd,,0.986
89,1:ckd,1:ckd,,0.986
90,1:ckd,1:ckd,,0.986

```

รูปที่ ข-10 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีต้นไม้ตัดสินใจ (ต่อ)

Classifier

Choose J48 -C 0.25 -M 2

Test options

Use training set

Supplied test set

Cross-validation Folds 10

Percentage split % 66

(Nom) class

Result list (right-click for options)

18:12:56 - trees.J48

18:14:20 - misc.InputMappedClassifier

Classifier output

```

91,1:ckd,1:ckd,,0.986
92,1:ckd,1:ckd,,0.986
93,1:ckd,1:ckd,,0.986
94,2:notckd,2:notckd,,1
95,1:ckd,1:ckd,,0.986
96,1:ckd,1:ckd,,0.986
97,1:ckd,1:ckd,,0.986
98,1:ckd,1:ckd,,0.986
99,1:ckd,1:ckd,,0.986
100,1:ckd,1:ckd,,0.986
101,1:ckd,1:ckd,,0.986
102,1:ckd,1:ckd,,0.986
103,1:ckd,1:ckd,,0.986
104,1:ckd,1:ckd,,0.986
105,1:ckd,1:ckd,,0.986
106,1:ckd,1:ckd,,0.986
107,1:ckd,1:ckd,,0.986
108,1:ckd,1:ckd,,0.986
109,1:ckd,1:ckd,,0.986
110,1:ckd,1:ckd,,0.986
111,1:ckd,1:ckd,,0.986
112,1:ckd,1:ckd,,0.986
113,1:ckd,1:ckd,,0.986
114,2:notckd,2:notckd,,1
115,1:ckd,1:ckd,,0.986
116,1:ckd,1:ckd,,0.986
117,1:ckd,1:ckd,,0.986
118,2:notckd,2:notckd,,0.503
119,1:ckd,1:ckd,,0.986
120,1:ckd,1:ckd,,0.986

```

รูปที่ ข-10 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีต้นไม้ตัดสินใจ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. วิธีโครงข่ายประสาทเทียม

Classifier

Choose: MultilayerPerceptron-1.0.1-M.0.9-N.20000-V.0-S.0-E.20-H.a

Test options

Use training set

Supplied test set Set...

Cross-validation Folds: 10

Percentage split %: 66

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

18:58:35 - functions.MultilayerPerceptron

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.03 seconds

=== Summary ===

Correctly Classified Instances	280	100 %
Incorrectly Classified Instances	0	0 %
Kappa statistic	1	
Mean absolute error	0.0001	
Root mean squared error	0.0005	
Relative absolute error	0.0301 %	
Root relative squared error	0.097 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	ckd
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	notckd
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

=== Confusion Matrix ===

a b <-- classified as

171 0 | a = ckd

0 109 | b = notckd

รูปที่ ข-11 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีโครงข่ายประสาทเทียม

Classifier

Choose: MultilayerPerceptron-1.0.1-M.0.9-N.20000-V.0-S.0-E.20-H.a

Test options

Use training set

Supplied test set Set...

Cross-validation Folds: 10

Percentage split %: 66

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:00:30 - misc.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	118	98.3333 %
Incorrectly Classified Instances	2	1.6667 %
Kappa statistic	0.9634	
Mean absolute error	0.0138	
Root mean squared error	0.1046	
Relative absolute error	2.9592 %	
Root relative squared error	21.9386 %	
Coverage of cases (0.95 level)	99.1667 %	
Mean rel. region size (0.95 level)	50.8333 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.975	0.000	1.000	0.975	0.987	0.964	0.999	1.000	ckd
	1.000	0.025	0.953	1.000	0.976	0.964	0.999	0.998	notckd
Weighted Avg.	0.983	0.009	0.984	0.983	0.983	0.964	0.999	0.999	

=== Confusion Matrix ===

a b <-- classified as

77 2 | a = ckd

0 41 | b = notckd

รูปที่ ข-12 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีโครงข่ายประสาทเทียม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds

Percentage split %

More options...

(Nom) Class variable

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:00:50 - misc.InputMappedClassifier

Classifier output

--- Predictions on test set ---

inst#	actual	predicted	error	prediction
1,1	ckd	1:ckd,,1		
2,1	ckd	1:ckd,,1		
3,1	ckd	1:ckd,,1		
4,1	ckd	1:ckd,,0.998		
5,2	notckd	2:notckd,,1		
6,2	notckd	2:notckd,,1		
7,2	notckd	2:notckd,,1		
8,2	notckd	2:notckd,,1		
9,2	notckd	2:notckd,,1		
10,1	ckd	2:notckd,+,0.999		
11,1	ckd	1:ckd,,1		
12,2	notckd	2:notckd,,1		
13,2	notckd	2:notckd,,1		
14,2	notckd	2:notckd,,1		
15,2	notckd	2:notckd,,1		
16,2	notckd	2:notckd,,1		
17,1	ckd	1:ckd,,1		
18,2	notckd	2:notckd,,1		
19,2	notckd	2:notckd,,1		
20,2	notckd	2:notckd,,1		
21,1	ckd	1:ckd,,1		
22,1	ckd	1:ckd,,1		
23,2	notckd	2:notckd,,1		
24,2	notckd	2:notckd,,1		
25,2	notckd	2:notckd,,1		
26,2	notckd	2:notckd,,1		
27,2	notckd	2:notckd,,1		
28,2	notckd	2:notckd,,1		

รูปที่ ข-13 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีโครงข่ายประสาทเทียม

Classifier

Choose **MultilayerPerceptron -L D.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a**

Test options

Use training set

Supplied test set **Set...**

Cross-validation Folds

Percentage split %

More options...

(Nom) Class variable

Start **Stop**

Result list (right-click for options)

18:58:45 - functions.MultilayerPerceptron

19:00:30 - misc.InputMappedClassifier

Classifier output

29,2	notckd	2:notckd,,1		
30,2	notckd	2:notckd,,1		
31,2	notckd	2:notckd,,1		
32,2	notckd	2:notckd,,1		
33,2	notckd	2:notckd,,1		
34,2	notckd	2:notckd,,1		
35,2	notckd	2:notckd,,0.998		
36,2	notckd	2:notckd,,1		
37,2	notckd	2:notckd,,1		
38,2	notckd	2:notckd,,0.999		
39,1	ckd	1:ckd,,1		
40,1	ckd	1:ckd,,1		
41,1	ckd	1:ckd,,1		
42,2	notckd	2:notckd,,1		
43,1	ckd	1:ckd,,1		
44,1	ckd	1:ckd,,1		
45,2	notckd	2:notckd,,1		
46,1	ckd	1:ckd,,1		
47,1	ckd	1:ckd,,1		
48,2	notckd	2:notckd,,0.927		
49,2	notckd	2:notckd,,1		
50,1	ckd	2:notckd,+,0.556		
51,1	ckd	1:ckd,,1		
52,1	ckd	1:ckd,,1		
53,2	notckd	2:notckd,,1		
54,1	ckd	1:ckd,,1		
55,1	ckd	1:ckd,,1		
56,1	ckd	1:ckd,,1		
57,1	ckd	1:ckd,,1		
58,2	notckd	2:notckd,,0.999		
59,2	notckd	2:notckd,,1		

รูปที่ ข-13 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีโครงข่ายประสาทเทียม (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
 Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

Classifier output

```

60,1:ckd,1:ckd,,1
61,1:ckd,1:ckd,,1
62,1:ckd,1:ckd,,1
63,1:ckd,1:ckd,,1
64,1:ckd,1:ckd,,1
65,1:ckd,1:ckd,,1
66,1:ckd,1:ckd,,1
67,2:noteckd,2:noteckd,,1
68,1:ckd,1:ckd,,1
69,1:ckd,1:ckd,,1
70,1:ckd,1:ckd,,1
71,1:ckd,1:ckd,,1
72,1:ckd,1:ckd,,1
73,1:ckd,1:ckd,,1
74,1:ckd,1:ckd,,1
75,1:ckd,1:ckd,,1
76,1:ckd,1:ckd,,1
77,1:ckd,1:ckd,,1
78,1:ckd,1:ckd,,1
79,1:ckd,1:ckd,,1
80,1:ckd,1:ckd,,1
81,1:ckd,1:ckd,,0.988
82,1:ckd,1:ckd,,1
83,1:ckd,1:ckd,,1
84,1:ckd,1:ckd,,0.999
85,1:ckd,1:ckd,,1
86,2:noteckd,2:noteckd,,1
87,1:ckd,1:ckd,,1
88,1:ckd,1:ckd,,1
89,1:ckd,1:ckd,,1
90,1:ckd,1:ckd,,1
    
```

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 18:58:45 - functions.MultilayerPerceptron
 19:00:30 - misc.InputMappedClassifier

รูปที่ ข-13 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีโครงข่ายประสาทเทียม (ต่อ)

Classifier
 Choose **MultilayerPerceptron -L 0.1 -M 0.9 -N 20000 -V 0 -S 0 -E 20 -H a**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

Classifier output

```

90,1:ckd,1:ckd,,1
91,1:ckd,1:ckd,,1
92,1:ckd,1:ckd,,1
93,1:ckd,1:ckd,,1
94,2:noteckd,2:noteckd,,1
95,1:ckd,1:ckd,,1
96,1:ckd,1:ckd,,1
97,1:ckd,1:ckd,,1
98,1:ckd,1:ckd,,0.997
99,1:ckd,1:ckd,,1
100,1:ckd,1:ckd,,1
101,1:ckd,1:ckd,,1
102,1:ckd,1:ckd,,1
103,1:ckd,1:ckd,,1
104,1:ckd,1:ckd,,1
105,1:ckd,1:ckd,,1
106,1:ckd,1:ckd,,1
107,1:ckd,1:ckd,,1
108,1:ckd,1:ckd,,1
109,1:ckd,1:ckd,,1
110,1:ckd,1:ckd,,1
111,1:ckd,1:ckd,,1
112,1:ckd,1:ckd,,1
113,1:ckd,1:ckd,,1
114,2:noteckd,2:noteckd,,1
115,1:ckd,1:ckd,,1
116,1:ckd,1:ckd,,1
117,1:ckd,1:ckd,,1
118,2:noteckd,2:noteckd,,1
119,1:ckd,1:ckd,,1
120,1:ckd,1:ckd,,1
    
```

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 18:58:45 - functions.MultilayerPerceptron
 19:00:30 - misc.InputMappedClassifier

รูปที่ ข-13 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีโครงข่ายประสาทเทียม (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. วิธีใช้พอร์ตเวกเตอร์แมชชีน

Classifier

Choose `SMD-C 1.0-L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel-E 1.0 -C 250007"`

Test options

Use training set

Supplied test set Set...

Cross-validation Folds

Percentage split %

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

18:45:33 - functions.SMO

Classifier output

=== Evaluation on training set ===

Time taken to test model on training data: 0.06 seconds

=== Summary ===

Correctly Classified Instances	277	98.9286 %
Incorrectly Classified Instances	3	1.0714 %
Kappa statistic	0.9776	
Mean absolute error	0.0107	
Root mean squared error	0.1035	
Relative absolute error	2.2525 %	
Root relative squared error	21.2289 %	
Coverage of cases (0.95 level)	98.9286 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.982	0.000	1.000	0.982	0.991	0.978	0.991	0.993	ckd
	1.000	0.018	0.973	1.000	0.986	0.978	0.991	0.973	notckd
Weighted Avg.	0.989	0.007	0.990	0.989	0.989	0.978	0.991	0.985	

=== Confusion Matrix ===

```

a b <-- classified as
168 3 | a = ckd
0 109 | b = notckd
    
```

รูปที่ ข-14 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีพอร์ตเวกเตอร์แมชชีน

Classifier

Choose `SMD-C 1.0-L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel-E 1.0 -C 250007"`

Test options

Use training set

Supplied test set Set...

Cross-validation Folds

Percentage split %

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

18:45:33 - functions.SMO

18:47:31 - misc.ipputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	118	98.3333 %
Incorrectly Classified Instances	2	1.6667 %
Kappa statistic	0.9634	
Mean absolute error	0.0167	
Root mean squared error	0.1291	
Relative absolute error	3.5828 %	
Root relative squared error	27.0801 %	
Coverage of cases (0.95 level)	98.3333 %	
Mean rel. region size (0.95 level)	50 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.975	0.000	1.000	0.975	0.987	0.964	0.987	0.991	ckd
	1.000	0.025	0.953	1.000	0.976	0.964	0.987	0.953	notckd
Weighted Avg.	0.983	0.009	0.984	0.983	0.983	0.964	0.987	0.978	

=== Confusion Matrix ===

```

a b <-- classified as
77 2 | a = ckd
0 41 | b = notckd
    
```

รูปที่ ข-15 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีพอร์ตเวกเตอร์แมชชีน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **SMD -C 1.0-L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 2**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:45:33 - functions.SMO

18:47:31 - misc.InputMappedClassifier

Classifier output

===== Predictions on test set =====

inst#	actual	predicted	error	prediction
1	1	ckd	1:ckd,,1	
2	1	ckd	1:ckd,,1	
3	1	ckd	1:ckd,,1	
4	1	ckd	2:notckd,+1	
5	2	notckd	2:notckd,,1	
6	2	notckd	2:notckd,,1	
7	2	notckd	2:notckd,,1	
8	2	notckd	2:notckd,,1	
9	2	notckd	2:notckd,,1	
10	1	ckd	1:ckd,,1	
11	1	ckd	1:ckd,,1	
12	2	notckd	2:notckd,,1	
13	2	notckd	2:notckd,,1	
14	2	notckd	2:notckd,,1	
15	2	notckd	2:notckd,,1	
16	2	notckd	2:notckd,,1	
17	1	ckd	1:ckd,,1	
18	2	notckd	2:notckd,,1	
19	2	notckd	2:notckd,,1	
20	2	notckd	2:notckd,,1	
21	1	ckd	1:ckd,,1	
22	1	ckd	2:notckd,+1	
23	2	notckd	2:notckd,,1	
24	2	notckd	2:notckd,,1	
25	2	notckd	2:notckd,,1	
26	2	notckd	2:notckd,,1	
27	2	notckd	2:notckd,,1	
28	2	notckd	2:notckd,,1	

รูปที่ ข-16 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน

Classifier

Choose **SMD -C 1.0-L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 2**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:45:33 - functions.SMO

18:47:31 - misc.InputMappedClassifier

Classifier output

29	2	notckd	2:notckd,,1	
30	2	notckd	2:notckd,,1	
31	2	notckd	2:notckd,,1	
32	2	notckd	2:notckd,,1	
33	2	notckd	2:notckd,,1	
34	2	notckd	2:notckd,,1	
35	2	notckd	2:notckd,,1	
36	2	notckd	2:notckd,,1	
37	2	notckd	2:notckd,,1	
38	2	notckd	2:notckd,,1	
39	1	ckd	1:ckd,,1	
40	1	ckd	1:ckd,,1	
41	1	ckd	1:ckd,,1	
42	2	notckd	2:notckd,,1	
43	1	ckd	1:ckd,,1	
44	1	ckd	1:ckd,,1	
45	2	notckd	2:notckd,,1	
46	1	ckd	1:ckd,,1	
47	1	ckd	1:ckd,,1	
48	2	notckd	2:notckd,,1	
49	2	notckd	2:notckd,,1	
50	1	ckd	1:ckd,,1	
51	1	ckd	1:ckd,,1	
52	1	ckd	1:ckd,,1	
53	2	notckd	2:notckd,,1	
54	1	ckd	1:ckd,,1	
55	1	ckd	1:ckd,,1	
56	1	ckd	1:ckd,,1	
57	1	ckd	1:ckd,,1	
58	2	notckd	2:notckd,,1	
59	2	notckd	2:notckd,,1	

รูปที่ ข-16 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier

Choose **SMD -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 25000**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:45:33 - functions.SMO

18:47:31 - misc.InputMappedClassifier

Classifier output

```

60,1:ckd,1:ckd,,1
61,1:ckd,1:ckd,,1
62,1:ckd,1:ckd,,1
63,1:ckd,1:ckd,,1
64,1:ckd,1:ckd,,1
65,1:ckd,1:ckd,,1
66,1:ckd,1:ckd,,1
67,2:notckd,2:notckd,,1
68,1:ckd,1:ckd,,1
69,1:ckd,1:ckd,,1
70,1:ckd,1:ckd,,1
71,1:ckd,1:ckd,,1
72,1:ckd,1:ckd,,1
73,1:ckd,1:ckd,,1
74,1:ckd,1:ckd,,1
75,1:ckd,1:ckd,,1
76,1:ckd,1:ckd,,1
77,1:ckd,1:ckd,,1
78,1:ckd,1:ckd,,1
79,1:ckd,1:ckd,,1
80,1:ckd,1:ckd,,1
81,1:ckd,1:ckd,,1
82,1:ckd,1:ckd,,1
83,1:ckd,1:ckd,,1
84,1:ckd,1:ckd,,1
85,1:ckd,1:ckd,,1
86,2:notckd,2:notckd,,1
87,1:ckd,1:ckd,,1
88,1:ckd,1:ckd,,1
89,1:ckd,1:ckd,,1
90,1:ckd,1:ckd,,1
<
    
```

รูปที่ ข-16 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน (ต่อ)

Classifier

Choose **SMD -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 2**

Test options

Use training set

Supplied test set

Cross-validation Folds

Percentage split %

(Nom) Class variable

Result list (right-click for options)

18:45:33 - functions.SMO

18:47:31 - misc.InputMappedClassifier

Classifier output

```

91,1:ckd,1:ckd,,1
92,1:ckd,1:ckd,,1
93,1:ckd,1:ckd,,1
94,2:notckd,2:notckd,,1
95,1:ckd,1:ckd,,1
96,1:ckd,1:ckd,,1
97,1:ckd,1:ckd,,1
98,1:ckd,1:ckd,,1
99,1:ckd,1:ckd,,1
100,1:ckd,1:ckd,,1
101,1:ckd,1:ckd,,1
102,1:ckd,1:ckd,,1
103,1:ckd,1:ckd,,1
104,1:ckd,1:ckd,,1
105,1:ckd,1:ckd,,1
106,1:ckd,1:ckd,,1
107,1:ckd,1:ckd,,1
108,1:ckd,1:ckd,,1
109,1:ckd,1:ckd,,1
110,1:ckd,1:ckd,,1
111,1:ckd,1:ckd,,1
112,1:ckd,1:ckd,,1
113,1:ckd,1:ckd,,1
114,2:notckd,2:notckd,,1
115,1:ckd,1:ckd,,1
116,1:ckd,1:ckd,,1
117,1:ckd,1:ckd,,1
118,2:notckd,2:notckd,,1
119,1:ckd,1:ckd,,1
120,1:ckd,1:ckd,,1
<
    
```

รูปที่ ข-16 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. วิชฐานกฎ

Classifier

Choose DecisionTable -X1-S "weka.attributeSelection.BestFirst-D1-N5"

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

19:02:58 - rules.DecisionTable

Classifier output

Time taken to test model on training data: 0.05 seconds

=== Summary ===

Correctly Classified Instances	277	98.9286 %
Incorrectly Classified Instances	3	1.0714 %
Kappa statistic	0.9774	
Mean absolute error	0.0798	
Root mean squared error	0.1417	
Relative absolute error	16.7778 %	
Root relative squared error	29.0604 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	63.9286 %	
Total Number of Instances	280	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.994	0.018	0.988	0.994	0.991	0.977	0.999	0.999	ckd
	0.982	0.006	0.991	0.982	0.986	0.977	0.999	0.998	notckd
Weighted Avg.	0.989	0.013	0.989	0.989	0.989	0.977	0.999	0.998	

=== Confusion Matrix ===

a b <- classified as

170 1 | a = ckd

2 107 | b = notckd

รูปที่ ข-17 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีฐานกฎ

Classifier

Choose DecisionTable -X1-S "weka.attributeSelection.BestFirst-D1-N5"

Test options

Use training set

Supplied test set Set...

Cross-validation Folds 10

Percentage split % 66

More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)

19:02:58 - rules.DecisionTable

19:04:02 - rules.InputMappedClassifier

Classifier output

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.03 seconds

=== Summary ===

Correctly Classified Instances	114	95 %
Incorrectly Classified Instances	6	5 %
Kappa statistic	0.8889	
Mean absolute error	0.1192	
Root mean squared error	0.213	
Relative absolute error	25.6196 %	
Root relative squared error	44.6786 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	67.0833 %	
Total Number of Instances	120	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.962	0.073	0.962	0.962	0.962	0.889	0.987	0.990	ckd
	0.927	0.038	0.927	0.927	0.927	0.889	0.987	0.973	notckd
Weighted Avg.	0.950	0.061	0.950	0.950	0.950	0.889	0.987	0.984	

=== Confusion Matrix ===

a b <- classified as

76 3 | a = ckd

3 38 | b = notckd

รูปที่ ข-18 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีฐานกฎ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอญญาที่เห็น ไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N5"

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop

Result list (right-click for options)
 19:02:58 - rules.DecisionTable
 19:16:02 - misc.InputMappedClassifier

Classifier output
 Predictions on test set ==
 inst#, actual, predicted, error, prediction
 1, 1:ckd, 1:ckd, , 0.667
 2, 1:ckd, 2:noteckd, +, 0.714
 3, 1:ckd, 1:ckd, , 0.667
 4, 1:ckd, 1:ckd, , 0.667
 5, 2:noteckd, 2:noteckd, , 0.983
 6, 2:noteckd, 2:noteckd, , 0.974
 7, 2:noteckd, 2:noteckd, , 0.974
 8, 2:noteckd, 2:noteckd, , 0.983
 9, 2:noteckd, 2:noteckd, , 0.983
 10, 1:ckd, 2:noteckd, +, 0.8
 11, 1:ckd, 1:ckd, , 0.923
 12, 2:noteckd, 2:noteckd, , 0.974
 13, 2:noteckd, 2:noteckd, , 0.983
 14, 2:noteckd, 2:noteckd, , 0.983
 15, 2:noteckd, 2:noteckd, , 0.983
 16, 2:noteckd, 2:noteckd, , 0.983
 17, 1:ckd, 1:ckd, , 0.923
 18, 2:noteckd, 2:noteckd, , 0.983
 19, 2:noteckd, 2:noteckd, , 0.974
 20, 2:noteckd, 2:noteckd, , 0.983
 21, 1:ckd, 1:ckd, , 0.667
 22, 1:ckd, 1:ckd, , 0.923
 23, 2:noteckd, 2:noteckd, , 0.983
 24, 2:noteckd, 2:noteckd, , 0.983
 25, 2:noteckd, 2:noteckd, , 0.974
 26, 2:noteckd, 1:ckd, +, 0.667
 27, 2:noteckd, 2:noteckd, , 0.983
 28, 2:noteckd, 2:noteckd, , 0.983
 <

รูปที่ ข-19 ผลการวิเคราะห์ข้อมูลสำหรับการทำงานนายตัวแบบด้วยวิธีฐานกฎ

Classifier
Choose DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N5"

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop

Result list (right-click for options)
 19:02:58 - rules.DecisionTable
 19:16:02 - misc.InputMappedClassifier

Classifier output
 29, 2:noteckd, 2:noteckd, , 0.974
 30, 2:noteckd, 2:noteckd, , 0.974
 31, 2:noteckd, 2:noteckd, , 0.983
 32, 2:noteckd, 2:noteckd, , 0.974
 33, 2:noteckd, 2:noteckd, , 0.983
 34, 2:noteckd, 2:noteckd, , 0.974
 35, 2:noteckd, 2:noteckd, , 0.714
 36, 2:noteckd, 2:noteckd, , 0.974
 37, 2:noteckd, 2:noteckd, , 0.983
 38, 2:noteckd, 1:ckd, +, 0.667
 39, 1:ckd, 1:ckd, , 0.8
 40, 1:ckd, 1:ckd, , 0.923
 41, 1:ckd, 1:ckd, , 0.981
 42, 2:noteckd, 2:noteckd, , 0.974
 43, 1:ckd, 1:ckd, , 0.967
 44, 1:ckd, 1:ckd, , 0.967
 45, 2:noteckd, 2:noteckd, , 0.974
 46, 1:ckd, 1:ckd, , 0.981
 47, 1:ckd, 1:ckd, , 0.981
 48, 2:noteckd, 1:ckd, +, 0.667
 49, 2:noteckd, 2:noteckd, , 0.983
 50, 1:ckd, 1:ckd, , 0.967
 51, 1:ckd, 2:noteckd, +, 0.714
 52, 1:ckd, 1:ckd, , 0.963
 53, 2:noteckd, 2:noteckd, , 0.974
 54, 1:ckd, 1:ckd, , 0.923
 55, 1:ckd, 1:ckd, , 0.981
 56, 1:ckd, 1:ckd, , 0.967
 57, 1:ckd, 1:ckd, , 0.667
 58, 2:noteckd, 2:noteckd, , 0.8
 59, 2:noteckd, 2:noteckd, , 0.974
 <

รูปที่ ข-19 ผลการวิเคราะห์ข้อมูลสำหรับการทำงานนายตัวแบบด้วยวิธีฐานกฎ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop

Result list (right-click for options)
 19:02:58 - rules.DecisionTable
 19:16:02 - misc.InputMappedClassifier

Classifier output
 60,1:ckd,1:ckd,,0.967
 61,1:ckd,1:ckd,,0.981
 62,1:ckd,1:ckd,,0.981
 63,1:ckd,1:ckd,,0.981
 64,1:ckd,1:ckd,,0.8
 65,1:ckd,1:ckd,,0.667
 66,1:ckd,1:ckd,,0.967
 67,2:notckd,2:notckd,,0.974
 68,1:ckd,1:ckd,,0.963
 69,1:ckd,1:ckd,,0.61
 70,1:ckd,1:ckd,,0.963
 71,1:ckd,1:ckd,,0.967
 72,1:ckd,1:ckd,,0.981
 73,1:ckd,1:ckd,,0.981
 74,1:ckd,1:ckd,,0.963
 75,1:ckd,1:ckd,,0.981
 76,1:ckd,1:ckd,,0.967
 77,1:ckd,1:ckd,,0.923
 78,1:ckd,1:ckd,,0.981
 79,1:ckd,1:ckd,,0.963
 80,1:ckd,1:ckd,,0.967
 81,1:ckd,1:ckd,,0.667
 82,1:ckd,1:ckd,,0.923
 83,1:ckd,1:ckd,,0.981
 84,1:ckd,1:ckd,,0.667
 85,1:ckd,1:ckd,,0.981
 86,2:notckd,2:notckd,,0.983
 87,1:ckd,1:ckd,,0.967
 88,1:ckd,1:ckd,,0.981
 89,1:ckd,1:ckd,,0.667
 90,1:ckd,1:ckd,,0.8

รูปที่ ข-19 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีฐานกฎ (ต่อ)

Classifier
Choose DecisionTable -X 1 -S "weka.attributeSelection.BestFirst -D 1 -N 5"

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop

Result list (right-click for options)
 19:02:58 - rules.DecisionTable
 19:16:02 - misc.InputMappedClassifier

Classifier output
 91,1:ckd,1:ckd,,0.61
 92,1:ckd,1:ckd,,0.981
 93,1:ckd,1:ckd,,0.981
 94,2:notckd,2:notckd,,0.983
 95,1:ckd,1:ckd,,0.981
 96,1:ckd,1:ckd,,0.981
 97,1:ckd,1:ckd,,0.667
 98,1:ckd,1:ckd,,0.667
 99,1:ckd,1:ckd,,0.981
 100,1:ckd,1:ckd,,0.981
 101,1:ckd,1:ckd,,0.981
 102,1:ckd,1:ckd,,0.667
 103,1:ckd,1:ckd,,0.981
 104,1:ckd,1:ckd,,0.967
 105,1:ckd,1:ckd,,0.667
 106,1:ckd,1:ckd,,0.967
 107,1:ckd,1:ckd,,0.667
 108,1:ckd,1:ckd,,0.923
 109,1:ckd,1:ckd,,0.923
 110,1:ckd,1:ckd,,0.981
 111,1:ckd,1:ckd,,0.75
 112,1:ckd,1:ckd,,0.963
 113,1:ckd,1:ckd,,0.923
 114,2:notckd,2:notckd,,0.983
 115,1:ckd,1:ckd,,0.967
 116,1:ckd,1:ckd,,0.923
 117,1:ckd,1:ckd,,0.981
 118,2:notckd,2:notckd,,0.75
 119,1:ckd,1:ckd,,0.667
 120,1:ckd,1:ckd,,0.981

รูปที่ ข-19 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีฐานกฎ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. วิธีการถดถอยลอจิสติก

Classifier
Choose Logistic-R 1.0E-8-M-1

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop
Result list (right-click for options)
19:24:09 - functions.Logistic

Classifier output
 === Evaluation on training set ===
 Time taken to test model on training data: 0.05 seconds
 === Summary ===
 Correctly Classified Instances 280 100 %
 Incorrectly Classified Instances 0 0 %
 Kappa statistic 1
 Mean absolute error 0
 Root mean squared error 0
 Relative absolute error 0 %
 Root relative squared error 0.0003 %
 Coverage of cases (0.95 level) 100 %
 Mean rel. region size (0.95 level) 50 %
 Total Number of Instances 280
 === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	ckd
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	notckd
Weighted Avg.	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	

 === Confusion Matrix ===
 a b <-- classified as
 171 0 | a = ckd
 0 109 | b = notckd

รูปที่ ข-20 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีการถดถอยลอจิสติก

Classifier
Choose Logistic-R 1.0E-8-M-1

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
Start Stop
Result list (right-click for options)
19:24:09 - functions.Logistic
19:25:15 - misc.InputMappedClassifier

Classifier output
 === Evaluation on test set ===
 Time taken to test model on supplied test set: 0.02 seconds
 === Summary ===
 Correctly Classified Instances 112 93.3333 %
 Incorrectly Classified Instances 8 6.6667 %
 Kappa statistic 0.8535
 Mean absolute error 0.0689
 Root mean squared error 0.257
 Relative absolute error 14.8155 %
 Root relative squared error 53.9184 %
 Coverage of cases (0.95 level) 94.1667 %
 Mean rel. region size (0.95 level) 50.8333 %
 Total Number of Instances 120
 === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.937	0.073	0.961	0.937	0.949	0.854	0.945	0.970	ckd
	0.927	0.063	0.884	0.927	0.905	0.854	0.963	0.878	notckd
Weighted Avg.	0.933	0.070	0.935	0.933	0.934	0.854	0.951	0.939	

 === Confusion Matrix ===
 a b <-- classified as
 74 5 | a = ckd
 3 38 | b = notckd

รูปที่ ข-21 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีการถดถอยลอจิสติก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
 Choose **Logistic -R 1.0E-8 -M -1**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

(Nom) Class variable **v**
 Start Stop
 Result list (right-click for options)
 19:24:09 - functions.Logistic
 19:25:15 - misc.InputMappedClassifier

Classifier output
 == Predictions on test set ==
 inst#, actual, predicted, error, prediction
 1, 1:ckd, 1:ckd, , 1
 2, 1:ckd, 1:ckd, , 1
 3, 1:ckd, 1:ckd, , 1
 4, 1:ckd, 1:ckd, , 1
 5, 2:noteckd, 2:noteckd, , 1
 6, 2:noteckd, 2:noteckd, , 1
 7, 2:noteckd, 2:noteckd, , 1
 8, 2:noteckd, 2:noteckd, , 1
 9, 2:noteckd, 2:noteckd, , 1
 10, 1:ckd, 1:ckd, , 1
 11, 1:ckd, 2:noteckd, +, 1
 12, 2:noteckd, 2:noteckd, , 1
 13, 2:noteckd, 2:noteckd, , 1
 14, 2:noteckd, 2:noteckd, , 1
 15, 2:noteckd, 2:noteckd, , 1
 16, 2:noteckd, 2:noteckd, , 1
 17, 1:ckd, 1:ckd, , 0.984
 18, 2:noteckd, 2:noteckd, , 1
 19, 2:noteckd, 1:ckd, +, 0.903
 20, 2:noteckd, 2:noteckd, , 1
 21, 1:ckd, 1:ckd, , 1
 22, 1:ckd, 2:noteckd, +, 1
 23, 2:noteckd, 2:noteckd, , 1
 24, 2:noteckd, 2:noteckd, , 1
 25, 2:noteckd, 2:noteckd, , 1
 26, 2:noteckd, 2:noteckd, , 1
 27, 2:noteckd, 2:noteckd, , 1
 28, 2:noteckd, 2:noteckd, , 1
 <

รูปที่ ข-22 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีการถดถอยลอจิสติก

Classifier
 Choose **Logistic -R 1.0E-8 -M -1**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

(Nom) Class variable **v**
 Start Stop
 Result list (right-click for options)
 19:24:09 - functions.Logistic
 19:25:15 - misc.InputMappedClassifier

Classifier output
 29, 2:noteckd, 2:noteckd, , 1
 30, 2:noteckd, 2:noteckd, , 1
 31, 2:noteckd, 2:noteckd, , 1
 32, 2:noteckd, 2:noteckd, , 1
 33, 2:noteckd, 2:noteckd, , 1
 34, 2:noteckd, 2:noteckd, , 1
 35, 2:noteckd, 2:noteckd, , 0.985
 36, 2:noteckd, 2:noteckd, , 1
 37, 2:noteckd, 2:noteckd, , 1
 38, 2:noteckd, 1:ckd, +, 1
 39, 1:ckd, 2:noteckd, +, 1
 40, 1:ckd, 1:ckd, , 1
 41, 1:ckd, 1:ckd, , 1
 42, 2:noteckd, 2:noteckd, , 1
 43, 1:ckd, 1:ckd, , 1
 44, 1:ckd, 1:ckd, , 1
 45, 2:noteckd, 2:noteckd, , 1
 46, 1:ckd, 1:ckd, , 1
 47, 1:ckd, 1:ckd, , 1
 48, 2:noteckd, 1:ckd, +, 1
 49, 2:noteckd, 2:noteckd, , 1
 50, 1:ckd, 1:ckd, , 1
 51, 1:ckd, 1:ckd, , 1
 52, 1:ckd, 1:ckd, , 1
 53, 2:noteckd, 2:noteckd, , 1
 54, 1:ckd, 1:ckd, , 1
 55, 1:ckd, 1:ckd, , 1
 56, 1:ckd, 1:ckd, , 1
 57, 1:ckd, 1:ckd, , 1
 58, 2:noteckd, 2:noteckd, , 1
 59, 2:noteckd, 2:noteckd, , 1
 <

รูปที่ ข-22 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีการถดถอยลอจิสติก (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose **Logistic -R 1.0E-8 -M -1**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:24:09 - functions.Logistic
 19:25:15 - misc.InputMappedClassifier

Classifier output
 60,1:ckd,1:ckd,,1
 61,1:ckd,1:ckd,,1
 62,1:ckd,1:ckd,,1
 63,1:ckd,1:ckd,,1
 64,1:ckd,1:ckd,,1
 65,1:ckd,1:ckd,,1
 66,1:ckd,1:ckd,,1
 67,2:notckd,2:notckd,,1
 68,1:ckd,1:ckd,,1
 69,1:ckd,1:ckd,,1
 70,1:ckd,1:ckd,,1
 71,1:ckd,1:ckd,,1
 72,1:ckd,1:ckd,,1
 73,1:ckd,1:ckd,,1
 74,1:ckd,1:ckd,,1
 75,1:ckd,1:ckd,,1
 76,1:ckd,1:ckd,,1
 77,1:ckd,1:ckd,,1
 78,1:ckd,1:ckd,,1
 79,1:ckd,1:ckd,,1
 80,1:ckd,1:ckd,,1
 81,1:ckd,1:ckd,,1
 82,1:ckd,1:ckd,,1
 83,1:ckd,1:ckd,,1
 84,1:ckd,1:ckd,,1
 85,1:ckd,1:ckd,,1
 86,2:notckd,2:notckd,,1
 87,1:ckd,1:ckd,,1
 88,1:ckd,1:ckd,,1
 89,1:ckd,1:ckd,,1
 90,1:ckd,2:notckd,+,1
 <

รูปที่ ข-22 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีการถดถอยลอจิสติก (ต่อ)

Classifier
Choose **Logistic -R 1.0E-8 -M -1**

Test options
 Use training set
 Supplied test set **Set...**
 Cross-validation Folds **10**
 Percentage split % **66**
 More options...

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:24:09 - functions.Logistic
 19:25:15 - misc.InputMappedClassifier

Classifier output
 91,1:ckd,1:ckd,,1
 92,1:ckd,1:ckd,,1
 93,1:ckd,1:ckd,,1
 94,2:notckd,2:notckd,,1
 95,1:ckd,1:ckd,,1
 96,1:ckd,1:ckd,,1
 97,1:ckd,1:ckd,,1
 98,1:ckd,1:ckd,,1
 99,1:ckd,2:notckd,+,1
 100,1:ckd,1:ckd,,0.663
 101,1:ckd,1:ckd,,1
 102,1:ckd,1:ckd,,1
 103,1:ckd,1:ckd,,1
 104,1:ckd,1:ckd,,1
 105,1:ckd,1:ckd,,1
 106,1:ckd,1:ckd,,1
 107,1:ckd,1:ckd,,1
 108,1:ckd,1:ckd,,1
 109,1:ckd,1:ckd,,1
 110,1:ckd,1:ckd,,1
 111,1:ckd,1:ckd,,1
 112,1:ckd,1:ckd,,1
 113,1:ckd,1:ckd,,1
 114,2:notckd,2:notckd,,1
 115,1:ckd,1:ckd,,1
 116,1:ckd,1:ckd,,1
 117,1:ckd,1:ckd,,1
 118,2:notckd,2:notckd,,1
 119,1:ckd,1:ckd,,1
 120,1:ckd,1:ckd,,1
 <

รูปที่ ข-22 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีการถดถอยลอจิสติก (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. วิธีนาอ็ฟ เบย์

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...
 (Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes

Classifier output
 === Evaluation on training set ===
 Time taken to test model on training data: 0.11 seconds
 === Summary ===

Correctly Classified Instances	272	97.1429 %
Incorrectly Classified Instances	8	2.8571 %
Kappa statistic	0.9407	
Mean absolute error	0.028	
Root mean squared error	0.1566	
Relative absolute error	5.8829 %	
Root relative squared error	32.1078 %	
Coverage of cases (0.95 level)	98.5714 %	
Mean rel. region size (0.95 level)	51.4286 %	
Total Number of Instances	280	

 === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.953	0.000	1.000	0.953	0.976	0.942	1.000	1.000	ckd
	1.000	0.047	0.932	1.000	0.965	0.942	1.000	1.000	notckd
Weighted Avg.	0.971	0.018	0.973	0.971	0.972	0.942	1.000	1.000	

 === Confusion Matrix ===
 a b <-- classified as
 163 8 | a = ckd
 0 109 | b = notckd

รูปที่ ข-23 ผลการวิเคราะห์ข้อมูลสำหรับการสร้างตัวแบบด้วยวิธีนาอ็ฟ เบย์

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...
 (Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes
 19:37:58 - misc.InputMappedClassifier

Classifier output
 === Evaluation on test set ===
 Time taken to test model on supplied test set: 0.03 seconds
 === Summary ===

Correctly Classified Instances	114	95 %
Incorrectly Classified Instances	6	5 %
Kappa statistic	0.8926	
Mean absolute error	0.049	
Root mean squared error	0.2023	
Relative absolute error	10.5249 %	
Root relative squared error	42.4244 %	
Coverage of cases (0.95 level)	97.5 %	
Mean rel. region size (0.95 level)	52.0833 %	
Total Number of Instances	120	

 === Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	FRC Area	Class
	0.924	0.000	1.000	0.924	0.961	0.898	1.000	1.000	ckd
	1.000	0.076	0.872	1.000	0.932	0.898	1.000	1.000	notckd
Weighted Avg.	0.950	0.026	0.956	0.950	0.951	0.898	1.000	1.000	

 === Confusion Matrix ===
 a b <-- classified as
 73 6 | a = ckd
 0 41 | b = notckd

รูปที่ ข-24 ผลการวิเคราะห์ข้อมูลสำหรับการทดสอบตัวแบบด้วยวิธีนาอ็ฟ เบย์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes
 19:57:58 - misc.InputMappedClassifier

Classifier output
 === Predictions on test set ===
 inst#,actual,predicted,error,prediction
 1,1:ckd,1:ckd,,1
 2,1:ckd,1:ckd,,1
 3,1:ckd,1:ckd,,1
 4,1:ckd,2:notckd,+,0.993
 5,2:notckd,2:notckd,,1
 6,2:notckd,2:notckd,,1
 7,2:notckd,2:notckd,,1
 8,2:notckd,2:notckd,,1
 9,2:notckd,2:notckd,,1
 10,1:ckd,1:ckd,,1
 11,1:ckd,1:ckd,,1
 12,2:notckd,2:notckd,,1
 13,2:notckd,2:notckd,,1
 14,2:notckd,2:notckd,,1
 15,2:notckd,2:notckd,,1
 16,2:notckd,2:notckd,,1
 17,1:ckd,2:notckd,+,0.66
 18,2:notckd,2:notckd,,1
 19,2:notckd,2:notckd,,1
 20,2:notckd,2:notckd,,1
 21,1:ckd,1:ckd,,1
 22,1:ckd,2:notckd,+,0.999
 23,2:notckd,2:notckd,,1
 24,2:notckd,2:notckd,,1
 25,2:notckd,2:notckd,,1
 26,2:notckd,2:notckd,,1
 27,2:notckd,2:notckd,,1
 28,2:notckd,2:notckd,,1

รูปที่ ข-25 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีนาอิว เบย์

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable

Start Stop

Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes
 19:57:58 - misc.InputMappedClassifier

Classifier output
 29,2:notckd,2:notckd,,1
 30,2:notckd,2:notckd,,1
 31,2:notckd,2:notckd,,1
 32,2:notckd,2:notckd,,1
 33,2:notckd,2:notckd,,1
 34,2:notckd,2:notckd,,1
 35,2:notckd,2:notckd,,1
 36,2:notckd,2:notckd,,1
 37,2:notckd,2:notckd,,1
 38,2:notckd,2:notckd,,1
 39,1:ckd,1:ckd,,0.808
 40,1:ckd,2:notckd,+,0.739
 41,1:ckd,1:ckd,,1
 42,2:notckd,2:notckd,,1
 43,1:ckd,1:ckd,,1
 44,1:ckd,1:ckd,,1
 45,2:notckd,2:notckd,,1
 46,1:ckd,1:ckd,,1
 47,1:ckd,1:ckd,,1
 48,2:notckd,2:notckd,,1
 49,2:notckd,2:notckd,,1
 50,1:ckd,1:ckd,,1
 51,1:ckd,1:ckd,,0.997
 52,1:ckd,1:ckd,,1
 53,2:notckd,2:notckd,,1
 54,1:ckd,2:notckd,+,1
 55,1:ckd,1:ckd,,1
 56,1:ckd,1:ckd,,1
 57,1:ckd,1:ckd,,1
 58,2:notckd,2:notckd,,1
 59,2:notckd,2:notckd,,1

รูปที่ ข-25 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีนาอิว เบย์ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes
 19:37:58 - misc.InputMappedClassifier

Classifier output
 60,1:ckd,1:ckd,,1
 61,1:ckd,1:ckd,,1
 62,1:ckd,1:ckd,,1
 63,1:ckd,1:ckd,,1
 64,1:ckd,2:notckd,+,0.859
 65,1:ckd,1:ckd,,1
 66,1:ckd,1:ckd,,1
 67,2:notckd,2:notckd,,1
 68,1:ckd,1:ckd,,1
 69,1:ckd,1:ckd,,1
 70,1:ckd,1:ckd,,1
 71,1:ckd,1:ckd,,1
 72,1:ckd,1:ckd,,1
 73,1:ckd,1:ckd,,1
 74,1:ckd,1:ckd,,1
 75,1:ckd,1:ckd,,1
 76,1:ckd,1:ckd,,1
 77,1:ckd,1:ckd,,1
 78,1:ckd,1:ckd,,1
 79,1:ckd,1:ckd,,1
 80,1:ckd,1:ckd,,1
 81,1:ckd,1:ckd,,1
 82,1:ckd,1:ckd,,1
 83,1:ckd,1:ckd,,1
 84,1:ckd,1:ckd,,1
 85,1:ckd,1:ckd,,0.996
 86,2:notckd,2:notckd,,1
 87,1:ckd,1:ckd,,1
 88,1:ckd,1:ckd,,1
 89,1:ckd,1:ckd,,1
 90,1:ckd,1:ckd,,1
 <

รูปที่ ข-25 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีนาอ์ฟ เบย์ (ต่อ)

Classifier
Choose NaiveBayes

Test options
 Use training set
 Supplied test set Set...
 Cross-validation Folds 10
 Percentage split % 66
 More options...

(Nom) Class variable
 Start Stop
 Result list (right-click for options)
 19:36:14 - bayes.NaiveBayes
 19:37:58 - misc.InputMappedClassifier

Classifier output
 91,1:ckd,1:ckd,,1
 92,1:ckd,1:ckd,,1
 93,1:ckd,1:ckd,,1
 94,2:notckd,2:notckd,,1
 95,1:ckd,1:ckd,,1
 96,1:ckd,1:ckd,,1
 97,1:ckd,1:ckd,,1
 98,1:ckd,1:ckd,,1
 99,1:ckd,1:ckd,,1
 100,1:ckd,1:ckd,,1
 101,1:ckd,1:ckd,,1
 102,1:ckd,1:ckd,,1
 103,1:ckd,1:ckd,,1
 104,1:ckd,1:ckd,,1
 105,1:ckd,1:ckd,,0.59
 106,1:ckd,1:ckd,,1
 107,1:ckd,1:ckd,,1
 108,1:ckd,1:ckd,,1
 109,1:ckd,1:ckd,,1
 110,1:ckd,1:ckd,,1
 111,1:ckd,1:ckd,,1
 112,1:ckd,1:ckd,,1
 113,1:ckd,1:ckd,,0.983
 114,2:notckd,2:notckd,,1
 115,1:ckd,1:ckd,,1
 116,1:ckd,1:ckd,,1
 117,1:ckd,1:ckd,,1
 118,2:notckd,2:notckd,,1
 119,1:ckd,1:ckd,,1
 120,1:ckd,1:ckd,,1
 <

รูปที่ ข-25 ผลการวิเคราะห์ข้อมูลสำหรับการทำนายตัวแบบด้วยวิธีนาอ์ฟ เบย์ (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Author Biography

Name	Mr. Surawat Sripaoraya
Date of Birth	29 August 1991
Address	87 Soi 5 Chutkeaw Klongjun Bangkapi Bangkok 10240
Education	Bachelor of Science in Statistics GPA 2.69 KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
Work place	Regional Container Lines (RCL)
Department	Office of President (OOP)
Position	Corporate Analyst



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้