

สารบัญ

บทที่	หน้า
1 บทนำ	1
ความเป็นมาของปัญหา.....	1
จุดมุ่งหมายของการศึกษา	2
ขอบเขตของงานวิจัย	3
นิยามศัพท์เฉพาะ.....	4
2 เอกสารและงานวิจัยที่เกี่ยวข้อง	5
การพยากรณ์.....	5
สถิติ.....	8
การทำเหมืองข้อมูล	9
กระบวนการสำหรับการทำเหมืองข้อมูล (CRISP-DM: Cross Industry Standard Process for Data Mining)	11
เทคนิคเคมีน (K-means)	12
เทคนิคการคัดเลือกตัวแปร (Attribute Selection)	17
การแปลงค่าข้อมูล (Data Transformation)	20
เทคนิคเคเนียร์สเนเบอร์ (K-Nearest Neighbor: K-NN)	
เทคนิคเนอ์ฟเบย์ (Naïve Bayes)	21
เทคนิคต้นไม้ตัดสินใจ (Decision Tree)	22
โครงข่ายประสาทเทียม (ANN: Artificial Neural Network)	24
k-fold Cross Validation	39
การวัดประสิทธิภาพของการพยากรณ์	40
ค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Root Mean Square Error: RMSE)	41
การพัฒนานาฬิกสูตร.....	42

สารบัญ (ต่อ)

บทที่	หน้า
ผลสัมฤทธิ์ทางการเรียน.....	46
งานวิจัยที่เกี่ยวข้อง.....	47
3 วิธีดำเนินการวิจัย	51
เครื่องมือที่ใช้ในงานวิจัย.....	51
ขั้นตอนการดำเนินงานวิจัย	51
4 ผลการวิจัย.....	79
ผลการทดลองที่ 1	79
ผลการทดลองที่ 2	96
5 บทสรุป.....	112
สรุปผลการวิจัย.....	112
อภิปรายผล	113
ข้อเสนอแนะ	114
บรรณานุกรม.....	116
ภาคผนวก.....	120
ประวัติผู้วิจัย.....	150

สารบัญตาราง

ตาราง	หน้า
1 แสดงตัวอย่างข้อมูลสำหรับการแบ่งกลุ่มเคมี.....	13
2 แสดง Confusion Matrix	40
3 แสดงโครงสร้างหลักสูตร ปรับปรุง พ.ศ. 2548 หลักสูตรปริญญาตรีวิทยาศาสตร์ บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์.....	53
5 แสดงตัวแปรของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ ช่วงชั้นปีที่ 1 และ 2.....	54
6 แสดงตัวแปรของนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ ช่วงชั้นปีที่ 1 และ 2.....	55
7 แสดงผลการคำนวณหาความคล้ายคลึงระหว่างข้อมูลทุกตัวกับจุดศูนย์กลางวัด ระยะทางแบบยูคลิเดียน	58
8 แสดงข้อมูลนิสิตที่เรียนในสาขาวิชาวิทยาการคอมพิวเตอร์ ช่วงชั้นปีที่ 1 และ 2	60
9 แสดงข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ ช่วงชั้นปีที่ 1 และ 2	65
10 แสดงข้อมูลภูมิหลังของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ ช่วงชั้นปีที่ 1 และ 2.....	66
11 แสดงการคัดเลือกตัวแปรของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ ด้วยสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน.....	67
12 แสดงการคัดเลือกตัวแปรของนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ด้วย สัมประสิทธิ์อันดับ	68
13 แสดงการคัดเลือกตัวแปรของนิสิตสาขาวิชาเทคโนโลยีสารสนเทศด้วย สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน	68
14 แสดงการคัดเลือกตัวแปรของนิสิตสาขาวิชาเทคโนโลยีสารสนเทศ ด้วยสัมประสิทธิ์อันดับ	69
15 แสดงข้อมูลผลการเรียนนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์.....	71
16 แสดงการแปลงค่าข้อมูลผลการเรียนนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์	72
17 แสดงข้อมูลนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์.....	73
18 แสดงข้อมูลผลการเรียนนิสิตสาขาวิชาวิทยาการคอมพิวเตอร์ ช่วงชั้นปีที่ 1 และ 2.....	75
19 แสดงผลทดลองของโมเดลด้วยเทคนิคเคเนียร์เรสเนเบอร์	80

สารบัญตาราง (ต่อ)

ตาราง		หน้า
20	แสดงผลทดลองของโมเดลด้วยเทคนิคเน็ฟเฟีย.....	81
21	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.1	82
22	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.2	83
23	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ สาขาวิชาวิทยาการคอมพิวเตอร์ กำหนดอัตราการเรียนรู้ = 0.3	84
24	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.1.....	85
25	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.2.....	86
26	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดสาขาวิชาเทคโนโลยีสารสนเทศ อัตราการเรียนรู้ = 0.3.....	87
27	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ ของสองสาขาวิชา จำนวน 4 ตัวแปรต้น.....	88
28	แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ.....	89
29	แสดงการทดสอบประสิทธิภาพของการจำแนกข้อมูลชุดฝึกสอน	93
30	แสดงการทดสอบประสิทธิภาพของการจำแนกข้อมูลชุดฝึกสอนที่รวมสาขา.....	94
31	แสดงการเปรียบเทียบการทดสอบประสิทธิภาพของข้อมูลชุดฝึกสอน และข้อมูลชุดทดสอบ	95
32	แสดงการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ของข้อมูลชุดฝึกสอนและข้อมูลชุดทดสอบ	95
33	แสดงการเปรียบเทียบประสิทธิภาพของข้อมูลชุดตรวจสอบ.....	96
34	แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคเคเนียร์โรสเนเบอร์	97
35	แสดงการทดสอบประสิทธิภาพของโมเดลด้วยเทคนิคเน็ฟเฟีย.....	98

สารบัญตาราง (ต่อ)

ตาราง		หน้า
36	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.1 สาขาวิชาวิทยาการคอมพิวเตอร์.....	99
37	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.2 สาขาวิชาวิทยาการคอมพิวเตอร์.....	100
38	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.3 สาขาวิชาวิทยาการคอมพิวเตอร์.....	101
39	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.1 สาขาวิชาเทคโนโลยีสารสนเทศ.....	102
40	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.2 สาขาวิชาเทคโนโลยีสารสนเทศ.....	103
41	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ กำหนดอัตราการเรียนรู้จาก 0.3 สาขาวิชาเทคโนโลยีสารสนเทศ.....	104
42	แสดงผลทดลองของโมเดลด้วยเทคนิคโครงข่ายประสาทเทียมแบบแพร่กลับ	105
43	แสดงผลทดลองของโมเดลด้วยเทคนิคต้นไม้ตัดสินใจ	106
44	แสดงค่าความถูกต้องของการจำแนกข้อมูลชุดฝึกสอน	109
45	แสดงค่าความถูกต้องของการจำแนกข้อมูลชุดฝึกสอนที่รวมสาขา	110
46	แสดงการเปรียบเทียบประสิทธิภาพของข้อมูลชุดฝึกสอน และข้อมูล ชุดทดสอบ	110
47	แสดงการเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ของข้อมูล	111
48	แสดงการเปรียบเทียบประสิทธิภาพของตัวแบบของข้อมูลชุดตรวจสอบ	111

สารบัญภาพ

ภาพ		หน้า
1	แสดงรูปแบบกระบวนการพัฒนาคุณภาพอย่างต่อเนื่อง (CPI)	49
1	แสดงกระบวนการค้นหาความรู้	10
2	แสดงขั้นตอนของกระบวนการสำหรับการทำเหมืองข้อมูล	11
3	แสดงการสุ่มข้อมูลเพื่อทำการแบ่งกลุ่มเคมีน	13
4	แสดงระยะทางของข้อมูลกับข้อมูลเริ่มต้นของกลุ่ม	14
5	แสดงข้อมูลที่ได้รับการแบ่งโดยอาศัย K	15
6	แสดงจุดศูนย์กลางของข้อมูลในแต่ละกลุ่มข้อมูล	16
7	แสดงการวัดระยะทางของข้อมูลกับจุดศูนย์กลาง	16
8	แสดงกลุ่มข้อมูลใหม่	16
9	แสดงโครงสร้างต้นไม้ตัดสินใจ	29
10	แสดงแบบจำลองข่ายประสาทในสมองมนุษย์	25
11	แสดงโครงสร้างของโครงข่ายประสาทเทียม	26
12	แสดงแบบจำลองโครงข่ายประสาทในคอมพิวเตอร์	27
13	แสดงฟังก์ชันแบบเชิงเส้น	28
14	แสดงซิกมอยด์ฟังก์ชัน	29
15	แสดงไฮเปอร์โบลิคแทนเจนต์ฟังก์ชัน	29
16	แสดงการเรียนรู้แบบมีผู้สอน	30
17	แสดงการเรียนรู้แบบไม่มีผู้สอน	31
18	แสดงสถาปัตยกรรมของ Feedforward Network	31
19	แสดงสถาปัตยกรรมของ Feedback Network	32
20	แสดงโครงข่ายประสาทเทียมแบบหลายชั้น	34
21	แสดงโครงข่ายประสาทเทียมแบบแพร่กระจายย้อนกลับ	36
22	แสดงวิธีการแบบ 10-fold Cross Validation	39
23	แสดงขั้นตอนการดำเนินการวิจัย	52

สารบัญภาพ (ต่อ)

ภาพ	หน้า
24 แสดงการคัดเลือกตัวแปรของนิสิตสาขาวิทยาการคอมพิวเตอร์ ด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกณฑ์ร่วมกับวิธีการ ค้นหาแบบจัดลำดับ	62
25 แสดงการคัดเลือกตัวแปรของนิสิตสาขาเทคโนโลยีสารสนเทศ ด้วยเทคนิคการคัดเลือกชุดตัวแปรแบบอัตราส่วนเกณฑ์ร่วมกับ วิธีการค้นหาแบบจัดลำดับ	63
26 แสดงขั้นตอนการคัดเลือกตัวแปรด้วยเทคนิคการคัดเลือกชุดตัวแปร แบบอัตราส่วนเกณฑ์ร่วมกับวิธีการค้นหาแบบจัดลำดับ	64
27 แสดงขั้นตอนการคัดเลือกตัวแปรด้วยสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และสัมประสิทธิ์อันดับ	70
28 แสดงใช้ค่ามาตรฐานอัตราส่วนเกณฑ์สูงสุดเป็นตัวแปรเริ่มต้น.....	77
29 แสดงต้นไม้ตัดสินใจ จำนวน 22 ตัวแปรต้น สาขาวิทยาการคอมพิวเตอร์	90
30 แสดงต้นไม้ตัดสินใจ จำนวน 11 ตัวแปรต้น สาขาวิทยาการคอมพิวเตอร์	90
31 แสดงต้นไม้ตัดสินใจ จำนวน 22 ตัวแปรต้น สาขาเทคโนโลยีสารสนเทศ	91
32 แสดงต้นไม้ตัดสินใจ จำนวน 8 ตัวแปรต้น สาขาเทคโนโลยีสารสนเทศ	91
33 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น สาขาวิทยาการคอมพิวเตอร์	92
34 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น สาขาเทคโนโลยีสารสนเทศ	92
35 แสดงต้นไม้ตัดสินใจของสองสาขาวิชา จำนวน 4 ตัวแปรต้น	93
36 แสดงต้นไม้ตัดสินใจสาขาวิทยาการคอมพิวเตอร์ จำนวน 13 ตัวแปรต้น	107
37 แสดงต้นไม้ตัดสินใจสาขาเทคโนโลยีสารสนเทศ จำนวน 13 ตัวแปรต้น	107
38 แสดงต้นไม้ตัดสินใจสาขาวิทยาการคอมพิวเตอร์ จำนวน 4 ตัวแปรต้น	108
39 แสดงต้นไม้ตัดสินใจสาขาเทคโนโลยีสารสนเทศ จำนวน 4 ตัวแปรต้น	108
40 แสดงต้นไม้ตัดสินใจ จำนวน 4 ตัวแปรต้น ของสองสาขาวิชา	109