



การพัฒนาแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียด
ในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูล

โดย
ปวีณา ชัยวนารมย์

สนับสนุนงบประมาณโดย
มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์
ประจำปีงบประมาณ 2558

DEVELOPING A MODEL FOR MULTILEVEL STRESS
PREDICTION USING DATA MINING TECHNIQUE

By

PAWEENA CHAIWANAROM

Granted by

Rajamangala University of Technology Rattanakosin
Fiscal year 2015

กิตติกรรมประกาศ

ผู้วิจัยขอขอบคุณเจ้าของตำราและผลงานตีพิมพ์ในหัวข้อที่เกี่ยวข้องที่ได้สร้างสรรค์เอกสารอันเป็นประโยชน์ในการค้นคว้าเพื่อต่อยอดความรู้ในเรื่องของความเครียดและการทำเหมืองข้อมูล

งานวิจัยฉบับนี้ สำเร็จสมบูรณ์ได้ด้วยการสนับสนุนงบประมาณแผ่นดิน ประจำปี 2558 จากมหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ ผู้วิจัยขอขอบพระคุณเป็นอย่างสูง

ท้ายที่สุดนี้ ผู้วิจัยขอกราบขอบพระคุณบิดา มารดา ครู อาจารย์ ที่ได้ให้การส่งเสริมสนับสนุนในการศึกษาแก่ผู้วิจัยมาโดยตลอด ผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิจัยนี้จะมีประโยชน์ต่อมหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ ตลอดจนผู้สนใจทั่วไป

ปวีณา ชัยวนารมย์
กันยายน 2558



บทคัดย่อ

รหัสโครงการ : A33 / 2558

ชื่อโครงการ : การพัฒนาแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูล

ชื่อนักวิจัย : ดร.ปวีณา ชัยวนารมย์

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาหาแบบจำลองที่เหมาะสมสำหรับพยากรณ์ความเครียดด้วยเทคนิคการทำเหมืองข้อมูล กระบวนการวิจัยเริ่มต้นด้วยให้กลุ่มตัวอย่างจำนวน 300 คนทำแบบทดสอบประเมินความเครียดเป็นจำนวน 4 รอบ โดยแต่ละรอบการประเมินมีระยะห่างของการทำแบบทดสอบเป็นระยะเวลา 2 เดือน จากนั้น ผลการทดสอบที่ได้ทั้งหมดถูกป้อนเข้าสู่โปรแกรม WEKA เพื่อทำการสร้างแบบจำลองด้วยอัลกอริทึมทางด้านเหมืองข้อมูลจำนวน 6 อัลกอริทึม คือ Bayesian Network, Naïve Bayesian, Decision Tree:4.5, Decision Table, Partial Rules (PART) และ Multilayer Perceptron (MLP) ในการสร้างแบบจำลองและการทดสอบแบบจำลองนั้น วิธี 10-fold cross-validation ได้ถูกนำมาใช้ในการแบ่งข้อมูลออกเป็นสองชุด ได้แก่ ชุดข้อมูลสอนและชุดข้อมูลทดสอบ จากการทดสอบแบบจำลองทั้งหมดพบว่า แบบจำลองที่เหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียดคือ แบบจำลองที่สร้างจากอัลกอริทึม MLP ที่ใช้กับข้อมูลย้อนหลัง 6 เดือน โดยแบบจำลองนี้มีค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกลับ และค่าความเหวี่ยง เท่ากับ 81%, 0.81, 0.81 และ 0.81 ตามลำดับ

คำสำคัญ : ความเครียด, เหมืองข้อมูล, โปรแกรมเวก้า, การจำแนกข้อมูล, การพยากรณ์

E-mail Address : paweena.c@rmutr.ac.th

ระยะเวลาโครงการ : ตุลาคม 2557 – กันยายน 2558

Abstract

Code of project : A33 / 2558

Project name : Developing A Model for Multilevel Stress Prediction using Data Mining Technique

Researcher : Dr.Paweena Chaiwanarom

This research aims to study and find the optimal stress prediction model based on data mining techniques. In the research procedures, we let 300 sample testers do a self-analysis stress test 4 times, while each time has been done every two months. Next, these raw data were used as the input data sets of six data mining algorithms based on WEKA to develop the prediction models. In this research, we studied six algorithms, i.e., Bayesian Network, Naïve Bayesian, Decision Tree:4.5, Decision Table, Partial Rules (PART), and Multilayer Perceptron (MLP). To evaluate these six models, 10-fold cross-validation were utilized to split the data into a training set and a test set. The experimental results show that the MPL algorithm with the data set of 3 test times during the last 6 month-period is the optimal prediction model comparing to the other five models. That is, its accuracy, precision, recall and F-measure are 81%, 0.81, 0.81 and 0.81, respectively

Keywords: Stress, Data Mining, WEKA, Data Classification, Prediction

E-mail Address : paweena.c@rmutr.ac.th

Period of project : October 2014 – September 2015

สารบัญ

	หน้า
กิตติกรรมประกาศ	ก
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
สารบัญ	ง
สารบัญตาราง	ฉ
สารบัญภาพ	ช
บทที่ 1 บทนำ	1
1. ความเป็นมาและความสำคัญของปัญหา	1
2. วัตถุประสงค์การวิจัย	3
3. คำถามการวิจัย	3
4. กรอบแนวคิดการวิจัย	3
5. ขอบเขตการวิจัย	7
6. ประโยชน์ที่คาดว่าจะได้รับ	7
บทที่ 2 ทบทวนวรรณกรรมที่เกี่ยวข้อง / ทฤษฎีที่เกี่ยวข้อง	8
1. การพยากรณ์	8
2. การทำเหมืองข้อมูล	12
3. แนวคิดเกี่ยวกับความเครียด	20
4. โปรแกรม WEKA	29
5. งานวิจัยที่เกี่ยวข้อง	30
บทที่ 3 ระเบียบวิธีการวิจัย	33
1. ประชากรและกลุ่มตัวอย่าง	33
2. เครื่องมือที่ใช้ในการวิจัย	34
3. การเก็บรวบรวมข้อมูล	35

สารบัญ (ต่อ)

	หน้า
4. การเตรียมข้อมูล	36
5. การสร้างแบบจำลองเพื่อพยากรณ์ความเครียด	46
บทที่ 4 ผลการวิจัย	52
1. การประเมินประสิทธิภาพ	52
2. การออกแบบการทดลอง	53
3. ผลการทดลอง	55
บทที่ 5 สรุปผล อภิปรายผลและข้อเสนอแนะ	66
1. สรุปผลการวิจัย	66
2. อภิปรายผล	68
3. ข้อเสนอแนะ	67
บรรณานุกรม	68
ภาคผนวก แบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง	71
ประวัติผู้วิจัย	75



สารบัญตาราง

ตารางที่		หน้า
1	ตัวอย่างข้อมูลประวัติส่วนตัวนักศึกษา	14
2	ตัวอย่างข้อมูลการลงทะเบียนเรียนของนักศึกษา	14
3	จำนวนประชากรและกลุ่มตัวอย่างจำแนกตามชั้นปี	34
4	ผลการวิเคราะห์ค่าของข้อมูลเพศและชั้นปี	39
5	การวิเคราะห์ข้อมูลอาการ พฤติกรรมหรือความรู้สึก ของการเก็บข้อมูลรอบที่ 1	40
6	การวิเคราะห์ข้อมูลอาการ พฤติกรรมหรือความรู้สึก ของการเก็บข้อมูลรอบที่ 2	41
7	การวิเคราะห์ข้อมูลอาการ พฤติกรรมหรือความรู้สึก ของการเก็บข้อมูลรอบที่ 3	42
8	การวิเคราะห์ข้อมูลอาการ พฤติกรรมหรือความรู้สึก ของการเก็บข้อมูลรอบที่ 4	43
9	Confusion matrix	52
10	แอทริบิวต์ที่ใช้ในแต่ละการทดลอง	54
11	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Bayesian Network	55
12	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Naïve Bayesian	55
13	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Multilayer Perceptron (MLP)	56
14	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Decision Table	56
15	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Partial Rules Decision (PART)	56
16	ประสิทธิภาพของแต่ละการทดลองโดยใช้ Decision Tree (C4.5)	58

สารบัญภาพ

ภาพที่		หน้า
1	กระบวนการประเมินความเครียดโดยใช้แบบทดสอบของกรมสุขภาพจิต	4
2	กระบวนการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด	5
3	การออกแบบการทดลองโดยจำแนกตามช่วงเวลา	55
4	ตัวอย่างต้นไม้ตัดสินใจ	18
5	ตัวอย่างการจัดกลุ่มข้อมูล	19
6	ข้อมูลที่นำมาใช้ในการศึกษาในรูปแบบ Excel	36
7	การตัดเอาทริบิวต์และเรคคอร์ดที่ไม่ใช้ในการวิเคราะห์และประเมินความเครียด	37
8	การแก้ไขข้อมูลบางแอทริบิวต์ที่ขาดหายไป	37
9	การเพิ่มข้อมูลคะแนนรวมและระดับความเครียด	38
10	การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 1	44
11	การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 2	44
12	การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 3	45
13	การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 4	45
14	การวิเคราะห์ระดับความเครียด 5 ระดับ จำนวน 4 รอบ	46
15	การบันทึกข้อมูลในรูปแบบนามสกุล *.csv	48
16	เว็บไซต์ที่ใช้ในการแปลงไฟล์ *.csv เป็น *.arff	49
17	การใช้โปรแกรม WordPad เปิดไฟล์ *.arff ขึ้นมาแก้ไข	49
18	การใช้โปรแกรม Weka เพื่อเปิดไฟล์ *.arff ขึ้นมาทำงาน	50
19	การใช้โปรแกรม Weka เพื่อวิเคราะห์ข้อมูลในเบื้องต้น	51
20	การใช้โปรแกรม Weka เพื่อสร้างแบบจำลอง 6 อัลกอริทึม	51
21	วิธีการวัดประสิทธิภาพของแบบจำลอง	53
22	ผลลัพธ์บางส่วนของการทดลองที่ 2 อัลกอริทึม Multilayer Perceptron (MLP)	57
23	ผลการทดลองที่ 1 จำแนกตามอัลกอริทึม	58
24	ผลการทดลองที่ 2 จำแนกตามอัลกอริทึม	59
25	ผลการทดลองที่ 3 จำแนกตามอัลกอริทึม	59
26	ผลการทดลองที่ 4 จำแนกตามอัลกอริทึม	60
27	ผลการทดลองที่ 5 จำแนกตามอัลกอริทึม	60
28	ผลการทดลองที่ 6 จำแนกตามอัลกอริทึม	61

สารบัญภาพ

ภาพที่		หน้า
29	ผลการทดลองที่ 7 จำแนกตามอัลกอริทึม	61
30	ประสิทธิภาพของอัลกอริทึม Bayesian Network โดยใช้ข้อมูลในตารางที่ 11	62
31	ประสิทธิภาพของอัลกอริทึม Naïve Bayesian โดยใช้ข้อมูลในตารางที่ 12	63
32	ประสิทธิภาพของอัลกอริทึม Multilayer Perceptron โดยใช้ข้อมูลในตารางที่ 13	63
33	ประสิทธิภาพของอัลกอริทึม Decision Table โดยใช้ข้อมูลในตารางที่ 14	64
34	ประสิทธิภาพของอัลกอริทึม Partial Rules Decision โดยใช้ข้อมูลในตารางที่ 15	64
35	ประสิทธิภาพของอัลกอริทึม Decision Tree (C4.5) โดยใช้ข้อมูลในตารางที่ 16	65



บทที่ 1

บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

จากการที่ภาครัฐได้ตระหนักถึงความสำคัญของการส่งเสริมการลดปัจจัยเสี่ยงด้านสุขภาพ อย่างเป็นองค์รวม โดยได้มีการกำหนดลงในยุทธศาสตร์การพัฒนาคอนสัสม์แห่งการเรียนรู้ตลอดชีวิต อย่างยั่งยืนของยุทธศาสตร์การพัฒนาประเทศตามแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ฉบับที่ 11 (พ.ศ. 2555-2559) ซึ่งสอดคล้องกับนโยบายและยุทธศาสตร์การวิจัยของชาติฉบับที่ 8 (พ.ศ. 2555-2559) ยุทธศาสตร์การวิจัยที่ 1 ที่มุ่งเน้นการสร้างศักยภาพและความสามารถในการพัฒนาสังคมใน แผนงานวิจัยที่ 3.6 การวิจัยเกี่ยวกับการพึ่งพาตนเองด้านสุขภาพและเสริมสร้างความมั่นคงด้าน สุขภาพอนามัย อีกทั้งประเด็นการส่งเสริมสุขภาพและการป้องกันโรครยังถูกจัดให้เป็นหนึ่งในกลุ่มเรื่อง วิจัย 13 เรื่องที่ควรมุ่งเน้นและหนึ่งใน 15 เรื่องเร่งด่วน ซึ่งสอดคล้องกับยุทธศาสตร์การวิจัยราย ประเด็นสุขภาพและชีวเวชศาสตร์ รวมไปถึงการเป็นหนึ่งในนโยบายระยะการบริหารราชการ 4 ปีของ รัฐบาล ด้านสังคมและคุณภาพชีวิตที่เน้นการพัฒนาสุขภาพของประชาชนโดยการเรียนรู้ป้องกันโรค เพื่อการดูแลรักษาสุขภาพตนเองของประชาชนอย่างมีประสิทธิภาพ จากนโยบายและยุทธศาสตร์ ดังกล่าวสะท้อนให้เห็นว่าประเด็นสุขภาพเป็นประเด็นพื้นฐานที่มีความสำคัญและส่งผลกระทบต่อ รวมทั้งภาคเศรษฐกิจและสังคมในการพัฒนาประเทศ

ปัจจัยสำคัญที่ทุกภาคส่วนต้องหันมาให้ความสำคัญกับประเด็นสุขภาพมากยิ่งขึ้นคือ การที่ ประเทศไทยจะเข้าสู่ประชาคมเศรษฐกิจอาเซียน (ASEAN Economic Community: AEC) ในปี 2558 ซึ่งจะก่อให้เกิดการสร้างความร่วมมือในการพัฒนาในทุกๆ ด้าน โดยเฉพาะด้านเศรษฐกิจ การค้า การลงทุน ซึ่งจะทำให้ภูมิภาคอาเซียนมีความมั่งคั่งและสามารถแข่งขันกับภูมิภาคอื่นๆ ได้ สิ่ง ที่ตามมาคือ จะมีการเปลี่ยนแปลงภายในภูมิภาคอย่างกว้างขวาง ซึ่งมีผลทั้งด้านบวกและด้านลบต่อ ประเทศไทยทั้งทางเศรษฐกิจ สังคม และสิ่งแวดล้อม ส่งผลให้วิถีชีวิตของประชาชนเปลี่ยนแปลงไป เพราะการแข่งขันจะทำให้ประชาชนในประชาคมอาเซียนและคนไทยต้องความเร่งรีบในการใช้ชีวิตจน เกิดภาวะความเครียดทางจิตใจและส่งผลกระทบต่อสุขภาพกายและสุขภาพจิตของตนเอง นอกจากนี้ ด้วยวิถีชีวิตที่ขาดความสมดุลและไม่เหมาะสมทำให้ประชาชนขาดการเอาใจใส่ที่จะป้องกัน และควบคุมปัจจัยเสี่ยงต่างๆ ที่มีผลกระทบต่อสุขภาพ จึงคาดการณ์ได้ว่าหลังจากการเข้าสู่ ประชาคมอาเซียนปัญหาความเครียดของคนไทยจะส่งผลกระทบต่อตนเองและต่อสังคมโดยรวม กล่าวคือ จำนวนประชากรที่ป่วยด้วยโรคไม่ติดต่อเรื้อรังหรือโรควิถีชีวิตที่มาจากความเครียดจะมีเพิ่ม มากขึ้น เช่น โรคเบาหวาน ความดันโลหิตสูง โรคหลอดเลือดสมอง เป็นต้น (ชาญยุทธัน วิหคโต และ นิทยา พันธุ์เวทย์, 2557) นอกจากนี้แล้วยังส่งผลกระทบต่อบุคคลรอบข้าง ไม่ว่าจะเป็นบุคคลใน ครอบครัว ที่ทำงานและในสังคมอีกด้วย หากไม่สามารถสกัดกั้นหรือป้องกันปัญหาความเครียดได้ สิ่งนี้

จะเป็นพื้นฐานที่ทำให้ประชาชนเกิดการเจ็บป่วยทางกายและมีปัญหาสุขภาพจิตเพิ่มมากขึ้น ส่งผลให้ตนเองและภาครัฐมีภาระค่าใช้จ่ายทางด้านสุขภาพและการสูญเสียทางเศรษฐกิจตามมาอย่างมหาศาล

จากที่มาและการคาดการณ์ถึงความสำคัญของปัญหาความเครียดข้างต้น ผู้วิจัยจึงได้ทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับปัญหาสุขภาพจิตและความเครียดจากหลายแหล่งข้อมูล (ปราณี และคณะ, 2552; สุภาพ, 2554; กิตติศักดิ์, 2555; ชาญยุทน์ และนิตยา, 2557; วีระยุทธ และคณะ, 2557) เพื่อหาแนวทางที่ประชาชนจะสามารถตรวจสอบระดับความเครียดด้วยตนเองและศึกษาแนวทางปฏิบัติเพื่อลดความเครียดที่ตนกำลังประสบอยู่ได้ จากการค้นคว้าพบว่า “แบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง” (กรมสุขภาพจิต กระทรวงสาธารณสุข, 2541) ได้ถูกออกมาไว้ในโครงการวิจัยกรมสุขภาพจิต เป็นที่เรียบร้อยแล้วในปี พ.ศ. 2541 แบบประเมินนี้มีข้อคำถามเพื่อใช้ทดสอบความเครียดจำนวน 20 ข้อ ซึ่งปัจจุบันได้ถูกนำมาใช้เป็นมาตรฐานสำหรับทดสอบความเครียดกันอย่างแพร่หลาย โดยผู้ตอบแบบประเมินจะป้อนอาการ, พฤติกรรมหรือความรู้สึกของตนเองในระยะเวลา 2 เดือนที่ผ่านมาโดยการตอบแบบสอบถามจำนวน 20 ข้อ ผลลัพธ์ที่ได้คือระดับความเครียดของผู้รับการประเมินพร้อมทั้งแนวทางในการจัดการกับความเครียดอย่างเหมาะสม แม้ว่าการวิจัยชิ้นนี้จะยังเป็นประโยชน์กับประชาชนในการช่วยลดค่าใช้จ่ายในการเดินทางมาสถานบริการสาธารณสุข รวมถึงช่วยลดภาระงานบริการของสถานบริการลงได้เป็นอย่างดีก็ตามงานวิจัยชิ้นนี้ยังมีจุดที่สามารถนำมาพัฒนาต่อยอดได้ใน 3 ประเด็นดังนี้

1. ยังไม่มีการนำคำตอบจากผลการประเมินและวิเคราะห์ความเครียดจากผู้เข้ามาเก็บรวบรวมอย่างเป็นรูปธรรม
2. ขาดการติดตามผลหลังจากผู้ทำการประเมินตนเองเรียบร้อยแล้ว ส่งผลให้ไม่สามารถทราบได้ว่าผู้ใช้มีภาวะความเครียดเป็นอย่างไรเมื่อเทียบกับอดีต
3. แบบประเมินและวิเคราะห์ความเครียดที่มีอยู่นี้ เป็นเพียงการวิเคราะห์สภาวะของผู้รับการประเมิน ณ ปัจจุบัน กล่าวคือ ไม่สามารถทำนายหรือคาดการณ์โอกาสในการเกิดความเครียดในอนาคตอันใกล้ได้

จากประเด็นต่างๆ ที่กล่าวมาข้างต้น ส่งผลให้ผู้วิจัยเกิดแนวคิดที่จะทำการวิจัยต่อยอดโดยใช้แบบประเมินและวิเคราะห์ความเครียดของกรมสุขภาพจิตเป็นเครื่องมือในการจัดเก็บข้อมูล ก่อนจะรวบรวมข้อมูลเหล่านั้นเข้าไว้ด้วยกันเพื่อใช้ในการสร้างแบบจำลองในการพยากรณ์โอกาสการเกิดความเครียดในอนาคตต่อไป

2. วัตถุประสงค์การวิจัย

1. เพื่อศึกษาทฤษฎีและกระบวนการที่นำมาใช้ในการพัฒนาแบบจำลองการพยากรณ์และจัดกลุ่มความเครียดด้วยเทคนิคการทำเหมืองข้อมูล และสามารถนำขั้นตอนวิธีแต่ละชนิดมาประยุกต์ใช้กับข้อมูลได้อย่างเหมาะสม
2. เพื่อรวบรวมคำตอบจากแบบประเมินและวิเคราะห์ความเครียดมาจัดเก็บไว้อย่างเป็นรูปธรรม

3. เพื่อศึกษาว่าเมื่อประชาชนที่มีความเครียดนำแนวทางในการลดความเครียดไปปฏิบัติแล้ว ส่งผลให้ความเครียดลดลงจริงหรือไม่

4. เพื่อพัฒนาแบบจำลองที่ใช้ในการพยากรณ์โอกาสการเกิดความเครียดในระดับต่างๆ ด้วยเทคนิคการทำเหมืองข้อมูล

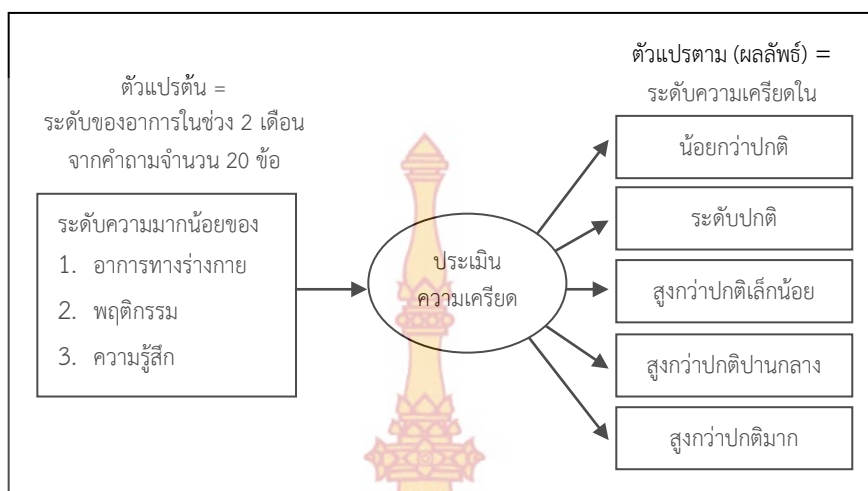
3. คำถามการวิจัย

1. อัลกอริทึมใดในการทำเหมืองข้อมูลที่เหมาะสมสำหรับสร้างแบบจำลองเพื่อพยากรณ์ความเครียด
2. ควรใช้ระยะเวลาเท่าใดในการเก็บข้อมูลจึงจะให้ผลดีที่สุดในการพยากรณ์ความเครียด

4. กรอบแนวคิดการวิจัย

งานวิจัยนี้ใช้แบบประเมินและวิเคราะห์ความเครียดของกรมสุขภาพจิตเป็นเครื่องมือในการจัดเก็บข้อมูล เมื่อผู้เข้ารับการประเมินตอบคำถามครบทั้ง 20 ข้อแล้ว ผู้วิจัยจะแจ้งผลการประเมินว่าผู้เข้ารับการประเมินแต่ละท่านมีความเครียดอยู่ในระดับใดพร้อมทั้งให้คำแนะนำในการปฏิบัติตนเพื่อบำบัดความเครียด ก่อนจะนำคำตอบและผลการประเมินเหล่านี้เก็บรวบรวมไว้อย่างเป็นทางการเป็นรูปธรรม หลังจากนั้นอีก 2 เดือน ผู้วิจัยจะทำการติดตามผลโดยให้ผู้เข้ารับการประเมินกลุ่มเดิมทำแบบประเมินชุดเดิม เพื่อดูว่าผลการประเมินระดับความเครียดเปลี่ยนแปลงไปหรือไม่เมื่อเวลาเปลี่ยนไป ผู้วิจัยจะแจ้งผลการประเมินและให้คำแนะนำแก่ผู้เข้ารับการประเมินพร้อมทั้งนำข้อมูลไปรวบรวมไว้เช่นเดียวกับการเก็บข้อมูลในครั้งแรก หลังจากนั้น ทุก 2 เดือน ผู้วิจัยจะทำการเก็บข้อมูลในลักษณะเดียวกันนี้อีก 2 ครั้ง รวมเป็นการเก็บข้อมูลทั้งสิ้นจำนวน 4 ครั้ง เหตุที่ต้องเก็บข้อมูลถึง 4 ครั้ง เนื่องจากต้องการนำข้อมูลทั้งหมดมาทำการทดลองว่าควรใช้ระยะเวลาในการเก็บข้อมูลนานเท่าไร เพื่อให้การสร้างแบบจำลองการพยากรณ์ความเครียดมีความแม่นยำสูงสุด

จากภาพ 1 ซึ่งแสดงกระบวนการประเมินความเครียดโดยใช้แบบทดสอบของกรมสุขภาพจิต มีรายละเอียดของขั้นตอนต่างๆ เริ่มต้นจากการใช้แบบทดสอบเก็บข้อมูลจากกลุ่มตัวอย่าง โดยให้กลุ่มตัวอย่างทำแบบทดสอบจำนวน 20 ข้อ เพื่อตอบว่าอาการทางพฤติกรรม และความรู้สึกของตนเองในช่วง 2 เดือนย้อนหลังเป็นอย่างไร จากนั้นนำผลที่ได้เข้าสู่กระบวนการประเมินความเครียดเพื่อบอกว่า ผู้ทำแบบประเมินมีความเครียดอยู่ในระดับใดใน 5 ระดับ พร้อมทั้งให้คำแนะนำในการปฏิบัติตัวเพื่อลดความเครียด

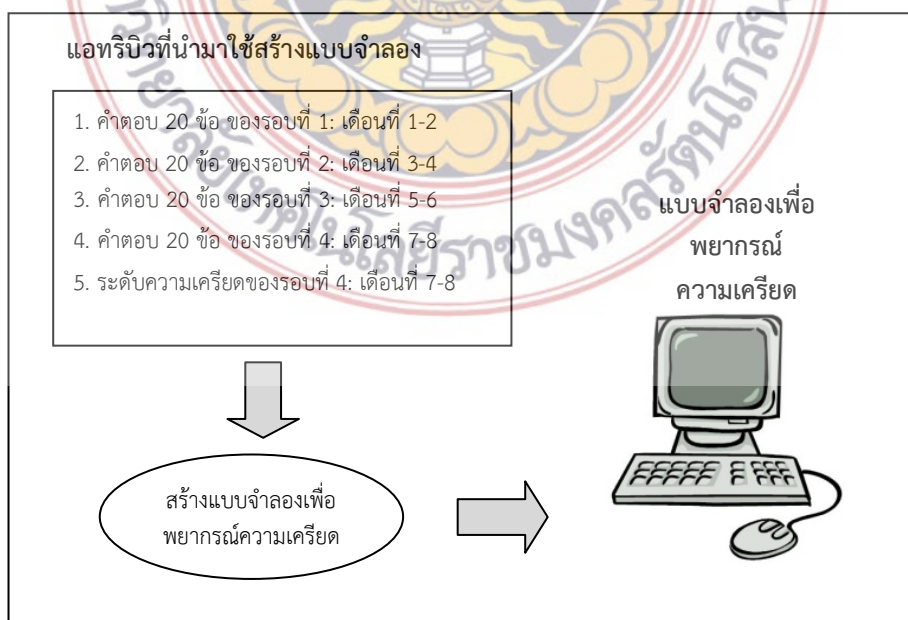


ภาพที่ 1 กระบวนการประเมินความเครียดโดยใช้แบบทดสอบของกรมสุขภาพจิต

กระบวนการดังภาพที่ 1 จะถูกกระทำซ้ำทั้งสิ้น 4 รอบ แต่ละรอบห่างกัน 2 เดือน และเพื่อตอบคำถามการวิจัยที่ตั้งไว้ว่า 1) อัลกอริทึมใดในการทำเหมืองข้อมูลที่เหมาะสมสำหรับสร้างแบบจำลองเพื่อพยากรณ์ความเครียด และ 2) ควรใช้ระยะเวลาเท่าใดในการเก็บข้อมูลจึงจะให้ผลดีที่สุดในการพยากรณ์ความเครียด เช่น 2 เดือน หรือ 6 เดือน เป็นต้น ผู้วิจัยได้ออกแบบกรอบแนวคิดการวิจัยเพื่อตอบโจทย์วิจัยทั้ง 2 ข้อดังต่อไปนี้

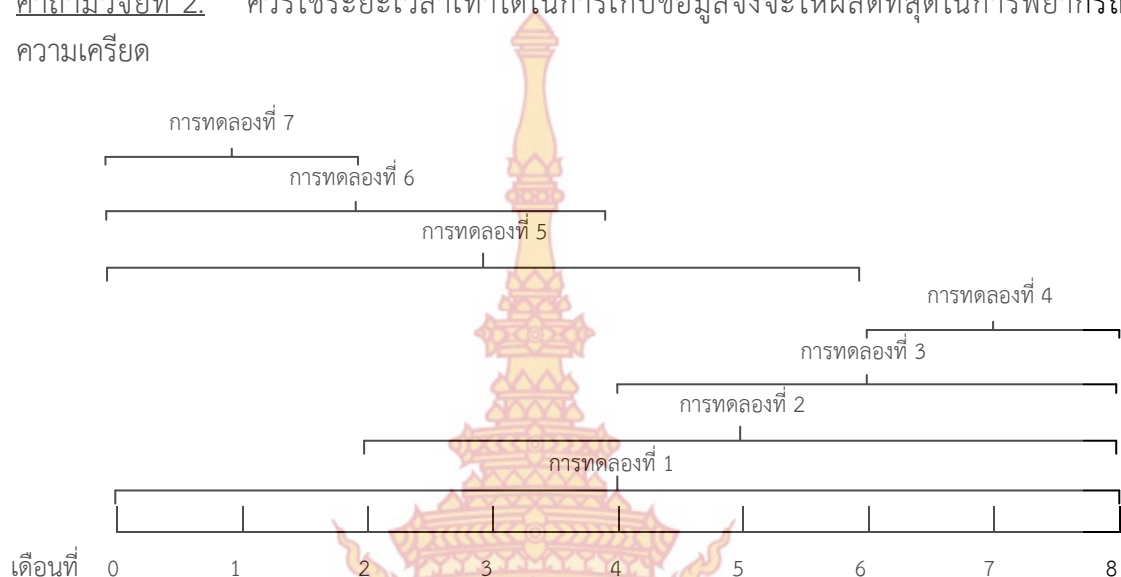
คำถามวิจัยที่ 1. อัลกอริทึมใดที่เหมาะสมในการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด

ในกระบวนการของการสร้างแบบจำลองจะมีการนำอัลกอริทึมทางด้านเหมืองข้อมูลจำนวน 6 วิธี มาใช้ในการสร้างแบบจำลองจำนวน 6 แบบ เพื่อเปรียบเทียบว่าอัลกอริทึมใดให้ผลดีที่สุดในการนำมาใช้เพื่อสร้างแบบจำลอง อัลกอริทึมที่นำมาทดลอง คือ Bayesian Network, Naïve Bayesian, Decision Tree:4.5, Decision Table, Partial Rules (PART), and Multilayer Perceptron (MLP) ภาพที่ 2 แสดงกระบวนการสร้างแบบจำลองข้อมูลเพื่อพยากรณ์ความเครียด



ภาพที่ 2 กระบวนการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด

คำถามวิจัยที่ 2. ควรใช้ระยะเวลาเท่าใดในการเก็บข้อมูลจึงจะให้ผลดีที่สุดในการพยากรณ์ความเครียด



ภาพที่ 3 การออกแบบการทดลองโดยจำแนกตามช่วงเวลา

การเก็บข้อมูลของงานวิจัยนี้ใช้วิธีการให้นักศึกษาตอบแบบประเมินความเครียดจากการสังเกตอาการ, พฤติกรรมหรือความรู้สึกของตนเองในระยะเวลา 2 เดือน โดยกำหนดให้ช่วงเวลาของการเก็บข้อมูลในแต่ละรอบเป็นดังนี้ รอบที่ 1 เก็บสิ้นเดือนที่ 2 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 1 - 2 รอบที่ 2 เก็บสิ้นเดือนที่ 4 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 3 - 4 รอบที่ 3 เก็บสิ้นเดือนที่ 6 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 5 - 6 และรอบที่ 4 เก็บสิ้นเดือนที่ 8 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 7 - 8 จากภาพที่ 3 แขนงอนแสดงระยะเวลาตั้งแต่เดือนที่ 1 ถึงเดือนที่ 8 ความหมายของการทดลองแต่ละแบบมีดังนี้

1. การทดลองที่ 1 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 - 8 รวมระยะเวลาต่อเนื่อง 8 เดือนย้อนหลัง
2. การทดลองที่ 2 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนสิงหาคม โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 3 - 8 รวมระยะเวลาต่อเนื่อง 6 เดือนย้อนหลัง
3. การทดลองที่ 3 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 5 - 8 รวมระยะเวลาต่อเนื่อง 4 เดือนย้อนหลัง
4. การทดลองที่ 4 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 7 - 8 รวมระยะเวลาต่อเนื่อง 2 เดือนย้อนหลัง
5. การทดลองที่ 5 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 - 6 รวมระยะเวลา 6 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้

ในรอบที่ 4 (เดือนที่ 7 – 8) มาใช้ จึงเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 6 เพื่อพยากรณ์ความเครียดตอนปลายเดือนที่ 8

6. การทดลองที่ 6 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 – 4 รวมระยะเวลา 4 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้ในรอบที่ 3 – 4 (เดือนที่ 5 – 8) มาใช้ จึงเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 4 เพื่อพยากรณ์ความเครียดตอนปลายเดือนที่ 8
7. การทดลองที่ 7 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 – 2 รวมระยะเวลา 2 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้ในรอบที่ 2, 3, 4 (เดือนที่ 3 – 8) มาใช้ นับเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 2 เพื่อพยากรณ์ความเครียดตอนปลายเดือนที่ 8

โดยสรุปคือ การทดลองที่ 1, 2, 3, 4 เป็นการทดลองที่ช่วงเวลาของการเก็บข้อมูลกับช่วงเวลาของการพยากรณ์ต่อเนื่องกัน ส่วนการทดลองที่ 5, 6 และ 7 นั้น เป็นการทดลองที่ช่วงเวลาของการเก็บข้อมูลกับช่วงเวลาในการพยากรณ์ไม่ต่อเนื่องกัน ดังนั้น ในแต่ละอัลกอริทึมจะทำการทดลองจำนวน 7 ครั้ง โดยแต่ละครั้งจะมีความแตกต่างกันตรงจำนวนแอทริบิวต์นำมาใช้ในการประมวลผลนั่นเอง

จากกรอบความคิดในการวิจัยข้างต้น นำไปสู่การตั้งสมมติฐานของการวิจัยดังนี้

1. กลุ่มตัวอย่างที่นำมาใช้ทั้ง 4 รอบ ต้องเป็นกลุ่มตัวอย่างเดียวกัน
2. ช่วงเวลาที่เหมาะสมในการเก็บข้อมูลควรเป็นช่วงปลายเดือน โดยตั้งสมมติฐานว่าความเครียดในช่วงปลายเดือน (ที่เงินเดือนยังไม่ออก) น่าจะมีมากกว่าช่วงต้นเดือนหรือช่วงกลางเดือน

5. ขอบเขตการวิจัย

งานวิจัยนี้มุ่งเน้นการศึกษาปัจจัยที่ทำให้เกิดภาวะความเครียดในปัจจุบัน และสร้างแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูลมีขอบเขตการศึกษาและพัฒนา ดังนี้

1. งานวิจัยนี้เก็บข้อมูลจากกลุ่มตัวอย่างจริงเป็นจำนวน 4 ครั้ง เพื่อใช้ทดลองว่าควรเก็บข้อมูลกี่ครั้งเพื่อให้ผลการพยากรณ์ออกมาดีที่สุด
2. ระยะเวลาในการเก็บข้อมูลแต่ละครั้งคือ 2 เดือน ตามหลักการของแบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง (สถาบันกัลยาณ์ราชนครินทร์, 2555)

6. ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ทราบว่าอัลกอริทึมใดที่เหมาะสมในการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด
2. ได้ทราบว่าแอทริบิวต์ (ตัวแปรต้น) ไດที่เหมาะสมในการนำมาใช้สร้างแบบจำลองเพื่อพยากรณ์ความเครียด
3. ได้ทราบจำนวนครั้งของการเก็บข้อมูลที่เหมาะสมเพื่อสร้างแบบจำลองพยากรณ์ความเครียด
4. ได้แบบจำลองเพื่อพยากรณ์ความเครียด

บทที่ 2

ทบทวนวรรณกรรมที่เกี่ยวข้อง / ทฤษฎีที่เกี่ยวข้อง

การวิจัยเรื่อง การพัฒนาแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับ ด้วยเทคนิคการทำเหมืองข้อมูล ผู้วิจัยได้ศึกษาค้นคว้า ตำรา เอกสาร ที่เป็นประโยชน์ และงานวิจัยต่างๆ ที่เกี่ยวข้องนำมารวบรวมและสรุปเป็นสาระสำคัญ มีหัวข้อต่างๆ ดังนี้

1. การพยากรณ์
2. การทำเหมืองข้อมูล (Data Mining)
3. แนวคิดเกี่ยวกับความเครียด (Stress)
4. โปรแกรม WEKA
5. งานวิจัยที่เกี่ยวข้อง

1. การพยากรณ์ (Forecasting)

การพยากรณ์เป็นทั้งศิลป์และวิทยาศาสตร์ เป็นการทำนายเหตุการณ์ที่สามารถพิสูจน์ ได้ด้วยวิธีทางวิทยาศาสตร์แล้ว นอกจากนั้นแล้ว การพยากรณ์ยังจะเป็นจุดเริ่มต้นที่สำคัญ เป็นเส้นทางที่กำหนดการดำเนินการอื่นๆ ทั้งหมด การพยากรณ์เรียกอีกอย่างหนึ่งว่าทำนาย (Prediction) หมายถึง การคาดคะเนหรือทำนายการเกิดเหตุการณ์หรือสภาพการณ์ต่างๆ ที่อาจจะเกิดขึ้นในอนาคต เช่น การพยากรณ์ยอดขายเพื่อประมาณการความต้องการวัตถุดิบ, การพยากรณ์ราคาหุ้นเพื่อเตรียมวางแผนที่จะซื้อหรือขายหุ้น, การพยากรณ์ความเสี่ยงในการเกิดโรคเพื่อหาแนวทางการป้องกันและลดโอกาสในการที่จะเป็นโรค เป็นต้น โดยการพยากรณ์จะทำการศึกษาแนวโน้มและรูปแบบการเกิดเหตุการณ์จากข้อมูลในอดีตและ/หรือใช้ความรู้ ความสามารถ ประสบการณ์ และดุลยพินิจของผู้พยากรณ์ (นิภา นิรุตติกุล, 2549)

1.1 องค์ประกอบของการพยากรณ์

องค์ประกอบของการพยากรณ์ที่คมีดังนี้

- ระบุวัตถุประสงค์ในการนำผลการพยากรณ์ไปใช้และช่วงเวลาที่จะใช้สำหรับพยากรณ์
- รวบรวมข้อมูลอย่างมีระบบ ถูกต้องตามความเป็นจริง
- ควรเลือกเทคนิคหรือวิธีการที่ให้ความเที่ยงตรงมากที่สุดเพื่อให้เกิดความคลาดเคลื่อนน้อยที่สุด
- บอกข้อจำกัดและสมมติฐานที่ตั้งไว้ในการพยากรณ์
- ตรวจสอบความถูกต้องแม่นยำของค่าพยากรณ์ที่ได้กับค่าจริงที่เกิดขึ้นเป็นระยะ

1.2 ระยะเวลาของการพยากรณ์

การพยากรณ์ตามแผนการปฏิบัติงาน เป็นการพยากรณ์ที่ขึ้นอยู่กับระยะเวลาของเวลาที่กำหนดตาม แผนงานโดยแบ่งออกเป็นช่วงๆ เช่นเดียวกับการวางแผนการทำงาน โดยทั่วไป ซึ่งสามารถแบ่งตามระยะเวลาการพยากรณ์ (อัจจรา จันทรฉาย, 2544) สามารถแบ่งออกได้ 4 ประเภท ได้ดังนี้

1. การพยากรณ์ระยะสั้นมาก (Immediate Term Forecasting) โดยปกติจะเป็นการพยากรณ์สั้นกว่า 1 เดือน เช่น การพยากรณ์ยอดขายเป็นรายวันใน 1 เดือน, การพยากรณ์กระแสเงินสดเป็นรายสัปดาห์ เป็นต้น

2. การพยากรณ์ระยะสั้น (Short Term Forecasting) เป็นการพยากรณ์ล่วงหน้าระยะเวลาไม่เกิน 1 ปี แต่ส่วนใหญ่ช่วงเวลาพยากรณ์อยู่ราวๆ 3 เดือน เพื่อให้ข้อมูลแก่ผู้บริหารระดับฝ่ายหรือหัวหน้าสายงาน การพยากรณ์เหมาะสำหรับการวางแผนในการทำบัญชีเงินสด กำหนดตารางการทำงาน การกำหนดงานและอัตราการผลิต เป็นต้น

3. การพยากรณ์ระยะปานกลาง (Medium Term Forecasting) เป็นการพยากรณ์ภายในระยะเวลา 1 ถึง 3 ปีข้างหน้า

4. การพยากรณ์ระยะยาว (Long Term Forecasting) เป็นการพยากรณ์มากกว่า 3 ปีขึ้นไป การพยากรณ์นี้เป็นช่วงระยะเวลาอันยาวเหมาะสำหรับการวางแผนเพื่อผลิตสินค้าใหม่ การขยายโรงงาน การวิจัย และพัฒนาในอนาคต

1.3 รูปแบบข้อมูลที่นำมาใช้ในการพยากรณ์ มีอยู่ 4 ประเภทได้แก่

1. ข้อมูลที่มีลักษณะสม่ำเสมอในแนวนอน (Horizontal Data Pattern) เป็นข้อมูลที่ไม่ค่อยเปลี่ยนแปลง จะมีค่าอยู่ใกล้กับค่าเฉลี่ยของข้อมูล

2. ข้อมูลที่มีลักษณะฤดูกาล (Seasonal Data Pattern) เป็นข้อมูลที่มีลักษณะขึ้นลงตามฤดูกาล เช่น ยอดขายร่มจะขายดีในฤดูฝน เป็นต้น

3. ข้อมูลที่ขึ้นลงตามวัฏจักร (Cyclical Data Pattern) ข้อมูลที่เก็บโดยมากจะเป็นรายปี และเก็บเป็นเวลาที่ยาวจะเห็นลักษณะการขึ้นลงของข้อมูลเป็นวัฏจักร

4. ข้อมูลที่มีลักษณะเป็นแนวโน้ม (Trend Data Pattern) เป็นข้อมูลที่มีลักษณะเป็นแนวโน้มเพิ่มขึ้นหรือลดลงอย่างสม่ำเสมอ

1.4 ประเภทของการพยากรณ์

การพยากรณ์สามารถแบ่งออกได้เป็น 2 ประเภทคือ การพยากรณ์เชิงคุณภาพและการพยากรณ์เชิงปริมาณ

1.4.1 การพยากรณ์เชิงคุณภาพ

การพยากรณ์เชิงคุณภาพ เป็นการพยากรณ์เกี่ยวกับความคิดเห็น ความต้องการ ความพอใจ การพยากรณ์คุณภาพผู้พยากรณ์จะต้องมีความเข้าใจในกระบวนการที่จะพยากรณ์อย่างแท้จริงทั้งหมดไม่เฉพาะส่วนใดส่วนหนึ่งเท่านั้น การพยากรณ์เชิงคุณภาพผู้พยากรณ์อาจจะเข้าไปศึกษาหรือสังเกตการณ์ กระบวนการผลิตสินค้าอย่างมีส่วนร่วมด้วยตนเองหรือส่งทีมงานไปศึกษาแล้วจดบันทึกเหตุการณ์ที่เกิดขึ้นอย่างละเอียดรวมทั้งความคิด ความรู้สึกของผู้พยากรณ์ในขณะนั้น การ

พยากรณ์เชิงคุณภาพมีข้อจำกัดคือ ผลการพยากรณ์ไม่สามารถนำไปอ้างอิงสรุปต่อ เพราะผลที่วิเคราะห์ออกมาใช้ได้กับเฉพาะงานหรือเฉพาะกลุ่มเท่านั้น เช่น การพยากรณ์ความต้องการผู้บริโภคสินค้าของบริษัทแห่งหนึ่งซึ่งผู้พยากรณ์อยากจะทราบความต้องการในสินค้าของบริษัทดังกล่าวในอนาคตเพื่อบริษัทจะได้เตรียมการปรับปรุงรูปแบบ ขนาด ฯลฯ นอกจากนี้ข้อมูลของการพยากรณ์เชิงคุณภาพยังได้ มาจากความคิดเห็นของคณะกรรมการบริหารของบริษัทและความคิดเห็นของกลุ่มพนักงานที่มีประสบการณ์ของบริษัท เป็นต้น การพยากรณ์เชิงคุณภาพสามารถแบ่งออกได้เป็น 3 กลุ่มใหญ่ๆ ดังนี้

1.4.1.1 เทคนิคที่ใช้วิจารณ์ญาณ (Subjective Assessment Methods) ใช้วิจารณ์ญาณประสบการณ์ของผู้พยากรณ์ ในการพยากรณ์

1.4.1.2 วิธีการค้นหา (Exploratory) เริ่มจากการศึกษาสภาพแวดล้อมในปัจจุบันและพยากรณ์ว่าจะอะไรจะเกิดขึ้นในอนาคตและเกิดขึ้นเมื่อไหร่ ตัวอย่างเทคนิค ได้แก่

- Scenario Analysis เป็นเทคนิคที่มีการพัฒนาจินตนาการเกี่ยวกับอนาคตด้านต่างๆ โดยการกำหนดสมมติฐาน และพัฒนาทางเลือก การสร้าง Scenario เป็นประโยชน์ในการวางแผนกลยุทธ์

- Delphi เป็นเทคนิคที่ให้ผู้บริหารแต่ละคนออกความคิดเห็นเป็นอิสระ โดยไม่ได้พบปะปรึกษาหารือแบบเผชิญหน้า แต่ใช้การออกแบบสอบถามและส่งคืนภายหลัง จากนั้นจะรวบรวมและส่งคืนให้สมาชิกประเมินคำตอบใหม่ โดยชี้ให้เห็นว่าคนส่วนมากมีความคิดเห็นอย่างไร ทำเช่นนี้จนกว่าการคาดคะเนเกิดจากความเห็นพ้องกัน

1.4.1.3 เทคนิคด้าน Normative เริ่มจากค้นหาจุดมุ่งหมายและทิศทางแล้วจึงค้นหาวิธีการพัฒนาและเทคโนโลยีเพื่อทำให้บรรลุจุดมุ่งหมาย ตัวอย่างเทคนิคได้แก่

- Relevance Trees คล้ายกับ Decision Trees เทคนิคนี้เป็นการระบุความต้องการในอนาคตและค้นหาว่าอะไรที่ต้องทำเพื่อให้บรรลุเป้าหมาย

- System Dynamic เป็นการวิเคราะห์ระบบโดยมีเป้าหมายในการพิจารณาความสัมพันธ์ซึ่งกันและกันของส่วนต่างๆ ในระบบหรือสถานะแวดล้อมมากกว่าที่จะศึกษาแต่ละส่วนแยกกัน

1.4.2 การพยากรณ์เชิงปริมาณ

การพยากรณ์เชิงปริมาณ เป็นการพยากรณ์ที่ต้องใช้เครื่องมือทางสถิติมาช่วยในการพยากรณ์ ข้อมูลทางสถิติเชิงปริมาณก็ได้มาจากประสบการณ์ การทำงานหรือการทำธุรกิจในชีวิตประจำวันที่ได้มีการจดบันทึกไว้ในอดีต ข้อมูลเชิงปริมาณจะอยู่ในรูปของตัวเลข ได้แก่ ข้อมูลจำนวนการขายสินค้าของบริษัทใน รอบ 1 เดือน หรือข้อมูลการผลิตสินค้าของบริษัท ในรอบ 3 เดือนที่ผ่านมา เป็นต้น ข้อมูลพื้นฐานที่ได้เหล่านี้มีประโยชน์มาก ผู้พยากรณ์สามารถเลือกรูปแบบการคำนวณให้เหมาะสม การพยากรณ์มีหลายรูปแบบตามชนิดของสถิติ สถิติบางชนิดคำนวณยากต้องใช้วิธีการ และขั้นตอนการคำนวณซับซ้อน แต่บางชนิดก็สามารถคำนวณได้ด้วยวิธีอย่างง่าย การพยากรณ์เชิงปริมาณสามารถแบ่งออกได้เป็น 2 กลุ่มหลักๆ คือ

1.4.2.1 เทคนิคอนุกรมเวลา (Time Series Techniques) เป็นเทคนิคที่ใช้ข้อมูลในอดีตเพื่อพยากรณ์ในอนาคต โดยข้อมูลในอดีตจะเก็บรวบรวมเป็นวัน รายสัปดาห์ รายเดือนหรือรายปีอย่างต่อเนื่อง สามารถแบ่งออกได้เป็น 2 กลุ่ม ได้แก่

- กลุ่มค่าเฉลี่ยเคลื่อนที่ (Moving Average) เทคนิคนี้เหมาะกับข้อมูลที่มีลักษณะคงที่ไม่เปลี่ยนแปลงมากในแต่ละงวด เช่น ยอดขายของสินค้าหรือบริการที่ไม่มีอิทธิพลของฤดูกาล (seasonal) เข้ามาเกี่ยวข้อง ในบางกรณีอาจใช้เทคนิคค่าถ่วงเฉลี่ยแบบถ่วงน้ำหนัก ซึ่ง ผลรวมของน้ำหนักที่ได้ต้องมีค่าเท่ากับ 1

- กลุ่มเทคนิคปรับเรียบเส้นโค้ง (Smoothing Technique) เป็นเทคนิคที่เหมาะสมกับข้อมูลที่ค่อนข้างไม่เปลี่ยนแปลง ใช้หลักการเดียวกับค่าถ่วงเฉลี่ยแบบเคลื่อนที่อย่างง่าย คือ ใช้ข้อมูลในอดีตมาถ่วงน้ำหนัก แต่น้ำหนักที่ถ่วงข้อมูลกับข้อมูลในอดีตไม่เท่ากัน

1.4.2.2 เทคนิคความสัมพันธ์ของข้อมูล (Causal Models) เป็นเทคนิคที่เน้นความสัมพันธ์ของตัวแปรในการพยากรณ์ เช่น การวิเคราะห์การถดถอยแบบง่ายหรือการพยากรณ์เชิงเดี่ยว (Simple Regression) เช่น การหาความสัมพันธ์ของยอดขายกับค่าโฆษณา เป็นต้น และการวิเคราะห์การถดถอยแบบพหุ (Multiple Regression) เช่น การหาความสัมพันธ์ของยอดขาย งบโฆษณา และจำนวนพนักงาน เป็นต้น

2. การทำเหมืองข้อมูล (Data Mining)

เหมืองข้อมูล ตามศัพท์ที่ราชบัณฑิตยสถานกำหนดไว้หมายถึง การสกัดหรือวิเคราะห์ ค้นหาข้อมูลที่ต้องการจากข้อมูลจำนวนมากได้ หรือกล่าวอีกนัยหนึ่ง เหมืองข้อมูลคือ ชุดซอฟต์แวร์วิเคราะห์ข้อมูลที่ได้ออกแบบมาเพื่อระบบสนับสนุนความต้องการของผู้ใช้ในการค้นหาข้อมูลที่ต้องการจากข้อมูลจำนวนมากได้

การทำเหมืองข้อมูล คือกระบวนการที่กระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น (Han, J., et al., 2012) เพื่อนำสารสนเทศที่ได้ไปใช้ประโยชน์ในการตัดสินใจ สารสนเทศที่ได้สามารถนำมาใช้ในการพยากรณ์หรือสร้างแบบจำลองสำหรับการจำแนกกลุ่ม หรือแสดงความสัมพันธ์ระหว่างกลุ่มต่างๆ ซึ่งการทำเหมืองข้อมูลจะต้องอาศัยองค์ความรู้จากหลายแขนงร่วมกันเพื่อสร้างตัวแบบ กฎเกณฑ์ รูปแบบการพยากรณ์และข้อความรู้ ในปัจจุบันการทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท ทั้งในด้านธุรกิจที่ช่วยในการตัดสินใจของผู้บริหาร ในด้านวิทยาศาสตร์และการแพทย์รวมทั้งในด้านเศรษฐกิจและสังคม การทำเหมืองข้อมูลเปรียบเสมือนวิวัฒนาการหนึ่งในการจัดเก็บและตีความหมายข้อมูล จากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย ๆ มาสู่การจัดเก็บในรูปแบบข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการทำเหมืองข้อมูลที่สามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูล

2.1 วิวัฒนาการของการทำเหมืองข้อมูล

- ปี ค.ศ 1960 Data Collection มีการนำข้อมูลมาจัดเก็บอย่างเหมาะสมในอุปกรณ์ที่นำเชื่อถือ เพื่อป้องกันการสูญหายได้เป็นอย่างดี
- ปี 1980 Data Access มีการนำข้อมูลที่จัดเก็บมาสร้างความสัมพันธ์ระหว่างกัน เพื่อนำไปวิเคราะห์ และตัดสินใจอย่างมีประสิทธิภาพ

- ปี 1990 Data Warehouse and Decision Support มีการนำข้อมูลมาเก็บลงในฐานข้อมูลขนาดใหญ่ ครอบคลุมการใช้งานทั้งหมดขององค์กร เพื่อช่วยสนับสนุนการตัดสินใจ
- ปี 2000 Data Mining นำข้อมูลจากฐานข้อมูลมาวิเคราะห์และประมวลผล โดยสร้างแบบจำลองและความสัมพันธ์ทางสถิติ

2.2 วัตถุประสงค์ในการใช้การทำเหมืองข้อมูล

- เพื่อการค้นพบองค์ความรู้ใหม่ในฐานข้อมูล (Knowledge Discovery in Databases)
- เพื่อการสกัดองค์ความรู้ที่ซ่อนเร้นอยู่ (Knowledge Extraction)
- เพื่อจัดการกับข้อมูลในอดีต (Data Archeology)
- เพื่อสำรวจข้อมูล (Data Exploration)
- เพื่อค้นหา Pattern ของข้อมูลที่ซ่อนอยู่ (Data Pattern Processing)
- เพื่อใช้ขุดเจาะข้อมูล (Data Dredging)
- เพื่อให้ได้มาซึ่งสารสนเทศที่มีประโยชน์

2.3 ปัจจัยที่ทำให้การทำเหมืองข้อมูลเป็นที่นิยม

- จำนวนและขนาดข้อมูลขนาดใหญ่ถูกผลิตและขยายตัวอย่างรวดเร็ว การสืบค้นความรู้จะมีความหมายก็ต่อเมื่อฐานข้อมูลที่ใช้มีขนาดใหญ่มาก ปัจจุบันมีจำนวนและขนาดข้อมูลขนาดใหญ่ที่ขยายตัวอย่างรวดเร็ว โดยผ่านทาง Internet ดาวเทียม และแหล่งผลิตข้อมูลอื่นๆ เช่น เครื่องอ่านบาร์โค้ด, เครดิตการ์ด, อีคอมเมิร์ซ
- ข้อมูลถูกจัดเก็บเพื่อนำไปสร้างระบบการสนับสนุนการตัดสินใจ (Decision Support System) เพื่อเป็นการง่ายต่อการนำข้อมูลมาใช้ในการวิเคราะห์เพื่อการตัดสินใจ ส่วนมากข้อมูลจะถูกจัดเก็บแยกมาจากระบบปฏิบัติการ (Operational System) โดยจัดอยู่ในรูปของคลังหรือเหมืองข้อมูล (Data Warehouse) ซึ่งเป็นการง่ายต่อการนำเอาไปใช้ในการสืบค้นความรู้
- ระบบคอมพิวเตอร์สมรรถนะสูงมีราคาต่ำลง เทคนิคเหมืองข้อมูลประกอบไปด้วยอัลกอริทึมที่มีความซับซ้อนและความต้องการการคำนวณสูง จึงจำเป็นต้องใช้งานกับระบบคอมพิวเตอร์สมรรถนะสูง ปัจจุบันระบบคอมพิวเตอร์สมรรถนะสูงมีราคาต่ำลง พร้อมด้วยเริ่มมีเทคโนโลยีที่นำเครื่องไมโครคอมพิวเตอร์จำนวนมากมาเชื่อมต่อกันโดยเครือข่ายความเร็วสูง (PC Cluster) ทำให้ได้ระบบคอมพิวเตอร์สมรรถนะสูงในราคาต่ำ
- การแข่งขันอย่างสูงในด้านอุตสาหกรรมและการค้า เนื่องจากปัจจุบันมีการแข่งขันอย่างสูงในด้านอุตสาหกรรมและการค้า มีการผลิตข้อมูลไว้อย่างมากมายแต่ไม่ได้นำมาใช้ให้เกิดประโยชน์ จึงเป็นการจำเป็นอย่างยิ่งที่ต้องควบคุมและสืบค้นความรู้ที่ถูกซ่อนอยู่ในฐานข้อมูล ความรู้ที่ได้รับสามารถนำไปวิเคราะห์เพื่อการตัดสินใจในการจัดการในระบบต่างๆ ซึ่งจะเห็นได้ว่าความรู้เหล่านี้ถือว่าเป็นผลิตภัณฑ์อีกชิ้นหนึ่งเลยทีเดียว

2.4 กระบวนการทำเหมืองข้อมูล (Knowledge Discovery from Database: KDD) มีขั้นตอนหลักอยู่ 3 ขั้นตอน (อุกฤษ ปัจฉิม, 2546) คือ Pre-processing, Data Mining และ Post-processing

2.4.1 Pre-processing คือขั้นตอนการจัดเตรียมข้อมูล ซึ่งการจัดเตรียมข้อมูลนั้นจะต้องทำการคัดข้อมูลที่ไมเกี่ยวข้องหรือข้อมูลเสีย (Noise Data) ออกจากแหล่งข้อมูลดิบเพื่อให้ได้ข้อมูลที่สัมพันธ์กัน ในขั้นตอนนี้แบ่งเป็นขั้นตอนย่อยดังนี้

2.4.1.1 Data Cleaning เป็นขั้นตอนสำหรับการคัดข้อมูลที่ไมเกี่ยวข้องออกไป โดยทั่วไปข้อมูลที่ถูกรวบรวมเก็บอาจมีความผิดปกติต่างๆ ได้ เช่น ข้อมูลบางแอทริบิวต์ขาดหายไป (missing value) ขาดแอทริบิวต์ที่น่าสนใจ หรือขาดรายละเอียดของข้อมูลเป็นข้อมูลรบกวน (noisy data) เช่น ข้อมูลมีค่าผิดพลาด (error) หรือมีค่าผิดปกติ (Outliers) ดังนั้นจึงต้องมีการจัดการข้อมูล การเตรียมข้อมูลเบื้องต้น มีวิธีการดังนี้

ตารางที่ 1 ตัวอย่างข้อมูลประวัติส่วนตัวนิสิต

ID	Sex	Name	Address	Phone	...	Major	GPA
1	นาย	สมภาพ ใจกล้า	86/9 หมู่ 2	0-222-23455	...	ไฟฟ้า	2.3
2	น.ส.	ดวงพร เอี่ยมสุข	54/2 หมู่ 7	0-224-7888	...	โยธา	3.2
...

ตารางที่ 2 ตัวอย่างข้อมูลการลงทะเบียนเรียนของนิสิต

ID	Subject	Section	Term	Year	Grade
1	001	1	1	2537	C+
1	002	1	1	2537	
1	005	1	1	2537	B+
...

- เลือกเฉพาะคอลัมน์สำคัญที่คาดว่าจะสามารถนำมาใช้ประโยชน์ได้ และเป็นคอลัมน์ที่มีข้อมูลค่อนข้างครบถ้วน เมื่อเทียบกับจำนวนนิสิต เช่น จากในตารางที่ 1 คอลัมน์สำคัญที่มีข้อมูลค่อนข้างมาก ได้แก่ ข้อมูลรหัสนิสิต ที่อยู่ อายุ เพศ ประวัติครอบครัวโรงเรียน เกรดเฉลี่ยที่จบการศึกษาในมหาวิทยาลัย เป็นต้น ส่วนในบางคอลัมน์ที่มีความสำคัญ แต่มีข้อมูลน้อยมากนั้นจะไม่นำพิจารณา เช่น ข้อมูลคะแนนสอบเอ็นทรานซ์ในแต่ละวิชา เหตุผลในการสอบเข้า เป็นต้น

- สำหรับคอลัมน์ที่มีค่าสำหรับทุกแถวเป็นค่าเดียวกัน เช่น “สัญชาติไทย” จะเป็นข้อมูลที่ไมสามารถแยกความแตกต่างของแต่ละแถวได้เลย ดังนั้นในการทำ mining จะไม่สามารถใช้ประโยชน์จากคอลัมน์นี้ ดังนั้น จึงไม่นำคอลัมน์นี้มาพิจารณา

- คอลัมน์ที่มีค่าที่ไม่ซ้ำกันเลย จากตารางที่ 1 ได้แก่ “Address”, “Phone” เป็นต้น ข้อมูลเหล่านี้ไม่สามารถหาแถวที่มีข้อมูลสัมพันธ์กันได้เลย การทำ mining จึงไม่สามารถนำข้อมูล เหล่านี้มาใช้ประโยชน์ได้ ดังนั้นในการทำ mining ควรกำจัดคอลัมน์ที่มีข้อมูลไม่ซ้ำกันเลยออก

- ปรับเปลี่ยนข้อมูลให้มีค่าเหมาะสมในการตัดสินใจ เช่น จากตารางที่ 1 ข้อมูลที่เป็นอยู่นั้นไม่สามารถที่จะนำไปใช้โดยตรงได้ เพราะจะเป็นปัญหาดังข้อที่แล้ว คือ ข้อมูลที่อยู่ของนิสิตแต่ละคนไม่ซ้ำกันเลย ดังนั้นจึงต้องปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่จะสามารถนำไปใช้ได้ ในกรณีนี้จะปรับข้อมูลในคอลัมน์ที่อยู่ของนิสิตให้เป็น Bangkok และ Non-Bangkok อย่างใดอย่างหนึ่ง เป็นต้น

- แก้ไขข้อมูลให้ถูกต้องสมบูรณ์ ได้แก่ การแก้ไขค่าว่างของข้อมูล ซึ่งสามารถแก้ไขได้หลายวิธี เช่น แก้ไขโดยจำกัดข้อมูลที่ในแถวเป็นค่าว่าง (NULL) ยกตัวอย่างเช่น จากในตารางที่ 2 ข้อมูลบางแถวค่าในคอลัมน์ Grade หายไป ซึ่งจะเห็นได้ว่าถ้ามีแต่รหัสนิสิตและวิชาที่ลงทะเบียน (Subject) โดยที่ไม่มีข้อมูลเกรดแล้ว เราก็ไม่สามารถจะนำแถวนั้นพิจารณาเพื่อหาความสัมพันธ์ที่น่าสนใจได้

- การจัดกลุ่มข้อมูลเพื่อลดการกระจาย (Binning Data) ทั้งนี้เนื่องมาจากข้อมูลของนิสิตมีจำนวนไม่มาก แต่เกรดในแต่ละวิชา ที่สามารถมีได้นั้นมีจำนวนมากถึง 10 ตัวด้วยกัน คือ {A,B+,B,C+,C,D+,D,F,W,I} ดังนั้นเพื่อลดการกระจายของข้อมูลเกรดของนิสิตที่มีมากเมื่อเทียบกับจำนวนนิสิต จึงได้จัดกลุ่มเกรดของนิสิตเป็น 3 กลุ่ม ดังนี้ คือ เกรด {A,B+,B} เป็น High , เกรด {C+,C} เป็น Medium และ เกรด {D+,D,F,W,I} เป็น Low

2.4.1.2 Data Integration เป็นขั้นตอนการรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน เช่น มีข้อมูลในคลังข้อมูล (Data Warehouse) ในรูปแบบของดาต้าคิวบ์ (Data Cube) และมีข้อมูลในรูปแบบฐานข้อมูลเชิงสัมพันธ์ (Relational Database) จำเป็นต้องทำการรวมข้อมูลให้เป็นข้อมูลชุดเดียวกัน

2.4.1.3 Data Selection เป็นขั้นตอนระบุถึงแหล่งข้อมูลที่จะนำมาทำ mining รวมถึงการนำข้อมูลที่ต้องการออกจากฐานข้อมูลเพื่อสร้างกลุ่มข้อมูลสำหรับพิจารณาในเบื้องต้น

2.4.1.4 Data transformation เป็นการจัดรูปแบบข้อมูลที่ได้จากขั้นตอนการคัดเลือกข้อมูลให้มีความเหมาะสมต่อการทำนาย เช่น การจัดระเบียบข้อมูลที่สัมพันธ์กันมาไว้ในระเบียบชุดเดียวกันหรือการแปลงค่าตัวเลขให้อยู่ในช่วงที่กำหนด เป็นต้น

จากขั้นตอนทั้งหมดจะเห็นว่าในขั้นตอนเบื้องต้นนี้เป็นเพียงแค่ขั้นตอนในการจัดเตรียมข้อมูลเพื่อจะนำไปใช้เท่านั้นยังไม่ใช่การใช้ข้อมูลเพื่อวิเคราะห์แบบจำลอง ซึ่งในขั้นตอนนี้จะใช้เวลานานในการดำเนินงานมากที่สุด หากจัดรูปแบบข้อมูลได้ไม่เหมาะสมแล้วในการทำนายก็จะเป็นความแม่นยำ ผู้ที่ทำการเตรียมข้อมูลจึงต้องมีความรู้เกี่ยวกับข้อมูลเหล่านี้ เช่น การวิเคราะห์โรคอัตโนมัติ จำเป็นจะต้องทราบเกี่ยวกับตัวแปรที่มีอิทธิพลต่อการจำแนกประเภท ของโรค เป็นต้น การทราบถึงปัจจัยที่ส่งผลต่อการทำนายจะทำให้วิเคราะห์ข้อมูลได้ถูกต้องและการทำงานจะมีความแม่นยำ

2.4.2 Data Mining เป็นขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่ โดยใช้เทคนิคและอัลกอริทึมที่เหมาะสมกับข้อมูลที่มีอยู่ เป็นขั้นตอนที่ใช้ในการนำข้อมูลที่พร้อมแล้วมาสร้างแบบจำลอง โดยขั้นแรกจะต้องทำการเลือกเทคนิคที่เหมาะสมกับภาพแบบชุดข้อมูล พิจารณาปัญหา เช่น ต้องการทำนายประเภทของโรคหรือต้องการแบ่งประเภทข้อมูลของโรค หรือต้องการหาปัจจัยที่เกี่ยวข้อง เป็นต้น หลังจากได้เทคนิคที่เหมาะสมแล้วจะทำการสอน (Train) ให้แบบจำลองเรียนรู้ลักษณะของข้อมูลว่า ชุดข้อมูลทั้งหมดมีความสัมพันธ์กันอย่างไร และทิศทางในการวิเคราะห์เป็นอย่างไรโดยในการสอนให้แบบจำลองเรียนรู้นั้นจำเป็นต้องมีการกำหนดพารามิเตอร์ (Parameter) หรือค่าตัวแปรต่างๆ ให้เหมาะสม ซึ่งในการพิจารณาพารามิเตอร์นั้นขึ้นอยู่กับเทคนิคที่เลือกใช้ ประสบการณ์ในการวิเคราะห์และการลองผิดลองถูก จากนั้นจึงนำแบบจำลองที่ได้ไปทดสอบหาความผิดพลาดของแบบจำลอง โดยการนำข้อมูลจริงที่เตรียมไว้สำหรับการทดสอบมาป้อนลงในแบบจำลองแล้วดูผลของการทำนายที่ได้

2.4.3 Post-processing คือขั้นตอนสุดท้ายของการทำเหมืองข้อมูล เป็นขั้นตอนการประเมินผลและนำสิ่งที่ได้มานำเสนอในภาพแบบของการใช้งาน การนำเสนอองค์ความรู้ในขั้นตอนนี้จะเป็นการวิเคราะห์ผลลัพธ์ที่ได้ แปลความหมาย และประเมินผลว่าผลลัพธ์นั้นเหมาะสมหรือตรงวัตถุประสงค์หรือไม่และนำเสนอ ซึ่งในขั้นตอนนี้หากผลการทดสอบ (Test) ไม่เป็นที่น่าพอใจแล้วจะต้องทำการจัดภาพแบบข้อมูลใหม่ (หรือเตรียมข้อมูลใหม่) เพื่อให้ได้ ค่าความถูกต้อง (Accuracy) มากที่สุด หรือค่าความคลาดเคลื่อนในการทำนายน้อยที่สุด (Error) หากค่าความถูกต้องยังน้อยอยู่หรือความผิดพลาดยังคงมีอยู่หลังจากเตรียมข้อมูลใหม่ อาจจำเป็นต้องเลือกเทคนิคในการทำเหมืองข้อมูลใหม่ ซึ่งการเพิ่มค่าความถูกต้องและลดค่าความคลาดเคลื่อนจะใช้วิธีใดนั้นขึ้นอยู่กับปัญหาที่เกิดขึ้น และในส่วนของขั้นตอนนี้แบ่งเป็นขั้นตอนย่อยๆ ดังนี้

- Pattern evaluation การประเมินผลของแบบจำลองที่ได้ว่า เกิดความผิดพลาดมากน้อยเพียงใด ภาพแบบการทำนายที่ได้เป็นไปตามความต้องการหรือไม่
- Knowledge presentation คือการแสดงผลของการทำนายที่ได้จากแบบจำลอง โดยการนำไปทำเป็นภาพแบบที่สามารถนำไปใช้ได้ง่าย สะดวกต่อการใช้งาน เช่น การป้อนข้อมูลเข้า และการแสดงผลลัพธ์ เป็นต้น

เราสามารถสรุปขั้นตอนของการค้นหาความรู้ใหม่ จากกระบวนการทำเหมืองข้อมูลได้ดังนี้

1. เรียนรู้และศึกษาเกี่ยวกับฐานข้อมูลและโปรแกรมที่จะใช้ในการทำเหมืองข้อมูล
2. การกรองข้อมูลและประมวลผล (Data cleaning and preprocessing) ข้อมูลที่เก็บรวบรวมมามีจำนวนมากจะต้องนำมากรอง เพื่อเลือกข้อมูลที่ตรงประเด็น เพราะบางข้อมูลอาจจะไม่เป็นประโยชน์กับเรา ในขั้นตอนนี้เป็นขั้นตอนที่เราจะได้มาซึ่งคุณภาพของข้อมูลที่จะนำไปวิเคราะห์
3. คัดเลือกข้อมูล (Data selection) เป็นการระบุถึงแหล่งข้อมูลที่จะนำมาทำ mining รวมถึงการนำข้อมูลที่ต้องการออกจากฐานข้อมูล เพื่อสร้างกลุ่มข้อมูลสำหรับพิจารณาในเบื้องต้นและทำการแปลงภาพแบบข้อมูล (Data transformation) ลดภาพและจัดข้อมูลให้อยู่ในภาพแบบเดียวกัน มีภาพแบบ (Format) ที่เป็นมาตรฐานและเหมาะสมที่จะนำไปใช้กับอัลกอริทึมและแบบจำลองที่ใช้ทำเหมืองข้อมูล

4. เลือกภาพแบบของการทำเหมืองข้อมูลเช่น Summarization, Classification, Regression, Association และ Clustering เป็นต้น และเลือกอัลกอริทึมที่เหมาะสมกับลักษณะของงาน

5. ทำการค้นหา Patterns ที่เราสนใจ และประเมินผล Pattern และนำเสนอองค์ความรู้ในขั้นตอนนี้จะเป็นการวิเคราะห์ผลลัพธ์ที่ได้ และแปลความหมาย และประเมินผลว่าผลลัพธ์นั้นเหมาะสมหรือตรงวัตถุประสงค์หรือไม่

6. ใช้องค์ความรู้ที่ค้นพบ

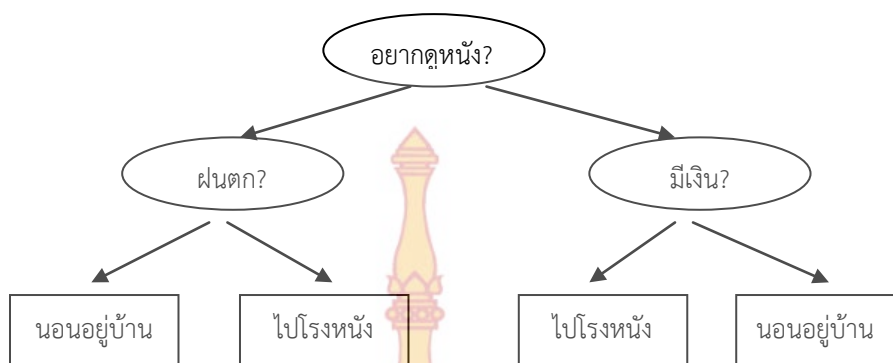
2.5 เทคนิคการทำเหมืองข้อมูล (Data Mining Techniques)

การทำเหมืองข้อมูล จะใช้เทคนิคหรือวิธีการใดนั้นขึ้นอยู่กับความเหมาะสมของการนำไปใช้งาน โดยเทคนิคของการทำเหมืองข้อมูลนั้นมีมากมาย ดังนี้

1. การบรรยายและการแสดงภาพของข้อมูล (Description / Visualization) คือ การวิเคราะห์ข้อมูลที่ไม่ซับซ้อนมากนัก ส่วนใหญ่เป็นการอธิบายรูปแบบและแนวโน้มที่อยู่ในข้อมูลหรือเป็นการสรุปข้อมูล การบรรยาย (Description) เป็นการหาค่าอธิบายถึงสิ่งที่จะเกิดขึ้นโดยอาศัยข้อมูลจากฐานข้อมูล เช่น การวิเคราะห์สหสัมพันธ์ (Correlation Analysis) กลุ่มคนที่มีการศึกษาหรือรายได้น้อย จะเลือกนักรบการเมืองที่มีนโยบายทุนนิยมมากกว่ากลุ่มคนชั้นกลาง ส่วนการแสดงภาพของข้อมูล (Visualization) เป็นการนำเสนอข้อมูลในรูปแบบกราฟิกหรืออาจนำเสนอในแบบ 2 มิติ สร้างรายละเอียดในการนำเสนอให้เข้าใจมากยิ่งขึ้น เช่น องค์กรต้องการหาสถานที่ในขยายสาขาใหม่ที่อยู่ในเขตพื้นที่ภาคเหนือของประเทศ ดังนั้นองค์กรจึงใช้แผนที่ Plot ที่ตั้งขององค์กรคู่แข่งที่มีสาขาอยู่ในเขตพื้นที่เพื่อพิจารณาสถานที่ตั้งที่เหมาะสมที่สุด

2. การจำแนกกลุ่ม (Classification) เป็นการจำแนกข้อมูลที่ต้องการ โดยอาศัยลักษณะที่คล้ายคลึงกันหรือแตกต่างกัน ให้อยู่ในกลุ่มใดกลุ่มหนึ่งที่กำหนดไว้ โดยการสร้างฟังก์ชันเชื่อมโยงระหว่างคุณสมบัติของวัตถุ กับประเภทของวัตถุ เช่น ในทางธุรกิจจะใช้เพื่อดูคุณสมบัติของผู้ที่จะก่อหนี้ดีหรือหนี้เสีย เพื่อประกอบการพิจารณาการอนุมัติเงินกู้ โดยอัลกอริทึมที่นิยมนำมาใช้ ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) โครงข่ายประสาทเทียม (Artificial Neural Network)

วิธีต้นไม้การตัดสินใจ (Decision Tree) เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปของต้นไม้ ซึ่ง Decision Tree นั้นมีการทำงานแบบ Supervised Learning คือ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างของข้อมูลที่ได้กำหนดได้ก่อนล่วงหน้าซึ่งเรียกว่า Training Set ได้อัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วยรูปแบบของ Tree จะประกอบด้วย Node แรกสุดที่เรียกว่า Root Node จาก Root Node ก็จะแตกออกเป็น Node ลูก และที่ Node ลูกก็จะมีลูกของตัวเองซึ่ง Node ที่ระดับสุดท้ายจะเรียกว่า Leaf Node โดยที่ Decision Tree จะมีการสร้างกฎต่างๆ ขึ้นเพื่อใช้ในการตัดสินใจ เนื่องจากความไม่ซับซ้อนของขั้นตอนวิธี ข้อดี คือ สามารถตีความและเข้าใจลักษณะของรูปแบบข้อมูลได้ง่ายเพราะมีการแยกออกเป็นกฎหรือข้อกำหนดต่างๆ แต่มีปัญหาเรื่องการให้น้ำหนักความน่าเชื่อถือหรือการให้ค่าน้ำหนักในแต่ละ Node ซึ่งถ้าให้น้ำหนักผิดไปอาจจะทำให้การตีความผิดไปได้



ภาพที่ 4 ตัวอย่างต้นไม้ตัดสินใจ

จากภาพที่ 4 จะเห็นว่า จาก Root Node จนถึง Leaf Node จะมีเพียงเส้นทางเดียวเท่านั้น ซึ่งเส้นทางนี้จะอธิบายถึงกฎที่ใช้สำหรับการจัดหมวดหมู่ของแต่ละกลุ่ม ซึ่งในแต่ละ Leaf Node นั้น อาจเป็นกลุ่มเดียวกัน ซึ่งเกิดจากเหตุผลที่แตกต่างกันได้

วิธีการที่ใช้สร้างต้นไม้การตัดสินใจ มีขั้นตอนดังนี้

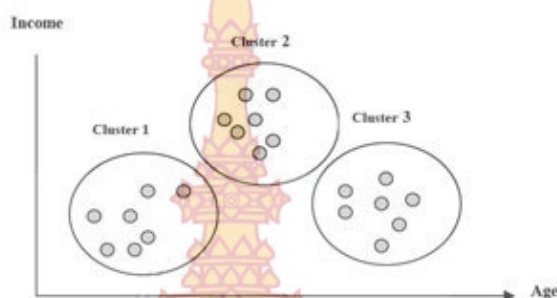
1. หาแอทริบิวต์ที่สำคัญที่สุดมาแบ่งข้อมูลโดยแอทริบิวต์นี้จะถูกนำมาสร้างเป็น Root Node โดยจะมีแอทริบิวต์เป้าหมายเป็นผลลัพธ์ซึ่งเป็น Leaf Node ถูกกำหนดไว้ก่อน
2. นำค่าที่เป็นไปได้ในแอทริบิวต์ที่ถูกเลือกมาแตกออกเป็นกลุ่มของตัวเอง
3. แบ่งข้อมูลทั้งหมดตามกลุ่มที่แตกออกจาก Root Node
4. วนกลับไปทำที่ขั้นตอนแรก คือ หาแอทริบิวต์ที่สำคัญที่สุดจากข้อมูลที่เข้ามาเพื่อหาตัวแบ่งต่อไป

ข้อจำกัดของต้นไม้การตัดสินใจ

- การแบ่งกลุ่มแบบต้นไม้การตัดสินใจกรณีเป็นข้อมูลที่มีค่าต่อเนื่อง เช่น ข้อมูลรายได้ ข้อมูลราคา ต้องทำการแปลงให้อยู่ในช่วงหรือตัดเป็นกลุ่มก่อน
- เมื่ออัลกอริทึมเลือกว่าจะใช้ค่าไหนเป็นตัวแบ่งกลุ่มแล้วก็จะไม่สนใจค่าอื่นที่อาจมีความสำคัญเช่นเดียวกัน
- การจัดการกับข้อมูลที่ไม่ทราบค่าอาจมีผลกระทบกับผลลัพธ์ของต้นไม้การตัดสินใจ
- ต้นไม้ที่มีระดับชั้นมากเกินไปจะทำให้ข้อมูลที่ผ่าน Node แตกออกเป็นชั้นเล็กชั้นน้อย ซึ่งข้อมูลเหล่านั้นจะไม่มีประโยชน์ในการนำมาใช้ทำการวิเคราะห์
- ปัญหาเรื่อง Overfitting / Overtaining เกิดจากการที่แบบจำลองได้เรียนรู้เข้าไปถึงรายละเอียดของข้อมูลมากเกินไปจะทำให้เกิด Node ที่เป็นส่วนเฉพาะเจาะจงกับกลุ่มข้อมูลที่ใช้ในการเรียนรู้ ซึ่งจะต้องหาวิธีการในการตัดกิ่งนี้ออกไป

3. การจัดกลุ่มข้อมูล (Clustering or Segmentation) เป็นการลดขนาดข้อมูลลงด้วยการรวมกลุ่มตัวแปรที่มีลักษณะคล้ายคลึงกันไว้ในกลุ่มเดียวกัน โดยอาศัยความเหมือน (Similarity) หรือ ความใกล้ชิด (Proximity) ดังภาพที่ 5 กล่าวง่ายๆ คือ การรวมกลุ่มข้อมูลที่มีลักษณะเหมือนกัน มี

รูปแบบหรือแนวโน้มที่จะเหมือนกัน โดยจะแบ่งชุดข้อมูลออกเป็นกลุ่ม (Cluster) นำข้อมูลที่มีคุณลักษณะเหมือนกันหรือคล้ายกันจัดไว้ในกลุ่มเดียวกันตัวอย่างเช่น องค์กรต้องการทราบความเหมือนที่มีในกลุ่มลูกค้าของตนเองเพื่อให้สามารถเข้าใจลักษณะเฉพาะของกลุ่มลูกค้าเป้าหมายขององค์กร องค์กรจึงต้องแยกกลุ่มของลูกค้าออกเป็นกลุ่มๆ อัลกอริทึมที่นิยมนำมาใช้ ได้แก่ K-Means Self-Organizing Map (SOM) Kohonen Hierarchical Two-Step clustering EM และ COBWOB



ภาพที่ 5 การจัดกลุ่มข้อมูล (Cluster)

4. การประมาณค่า (Estimation) เป็นการป้อนข้อมูลที่มีอยู่ เพื่อใช้ในการประเมินสิ่งต่าง ๆ ที่จะก่อให้เกิดประโยชน์ จัดว่าเป็นการประเมินที่ไม่สามารถกำหนดค่าหรือคุณสมบัติที่ชัดเจนได้ ใช้จัดการกับค่าที่มีผลแบบต่อเนื่อง เช่น ใช้ประเมินรายได้ของครอบครัว ประเมินความสูงของบุคคลในครอบครัว ประเมินจำนวนเด็ก ๆ ในครอบครัว

5. การทำนาย/พยากรณ์ (Prediction) จะคล้ายคลึงกับกับ Classification และ Estimation ต่างกันตรงที่ Record ถูกแยกจัดลำดับในการทำนายค่าในอนาคต และนำข้อมูลในอดีตมาสร้างเป็นแบบจำลอง ใช้ทำนายสิ่งที่จะเกิดขึ้นในอนาคต เช่น การทำนายการเปลี่ยนแปลงพฤติกรรมของตลาด การทำนายยอดขายของบริษัทในไตรมาสหน้า การทำนายว่าลูกค้ากลุ่มใดที่องค์กรจะสูญเสียไปในอีก 6 เดือนข้างหน้า หรือการทำนายยอดซื้อของลูกค้าจะเป็นเท่าใด ถ้าบริษัทลดราคาสินค้าลง 10% เป็นการทำนายค่าที่ต้องการจากข้อมูลที่มีอยู่ โดยสรุปคือ ทำนายนี้จะมุ่งเน้นศึกษาเกี่ยวกับพฤติกรรมในอนาคตมากกว่าในปัจจุบัน

6. การจัดกลุ่มโดยอาศัยความใกล้ชิดหรือกฎความสัมพันธ์ (Affinity Group or Association Rule) เป็นการหาความสัมพันธ์หรือความเกี่ยวเนื่องของข้อมูล เพื่อแสดงความสัมพันธ์ของเหตุการณ์หรือวัตถุที่เกิดขึ้นพร้อมกัน โดยอาศัยหลักของกฎ “ถ้า สิ่งที่เกิดขึ้น แล้ว ผลที่จะตามมา (If Antecedent then Consequent)” นำไปใช้วิเคราะห์ หรือทำนายปรากฏการณ์ เช่น การวิเคราะห์พฤติกรรมการซื้อสินค้าของลูกค้า เพื่อช่วยในการแนะนำสินค้าเพิ่มเติมได้ โดยพิจารณาสินค้าที่ผู้ซื้อ มักจะซื้อพร้อมกัน ร้านค้าก็อาจจะจัดวางสินค้าสองอย่างให้อยู่ใกล้กัน หรือจับคู่จัดแผนส่งเสริมการขาย เพื่อเพิ่มยอดขาย หรืออาจจะพบว่าหลังจากคนซื้อสินค้า ก แล้ว มักจะซื้อสินค้า ข ด้วย ก็สามารถหาความรู้นี้ไปแนะนำผู้ที่กำลังจะซื้อสินค้า ก ได้ อัลกอริทึมที่นิยมนำมาใช้ ได้แก่ Apriori GRI (Generalized Rule Induction) Carma Sequence Eclat FP-Growth (Frequent Pattern-Growth) OneR (One-Attribute-Rule) ZeroR (Zero-Attribute-Rule)

3. แนวคิดเกี่ยวกับความเครียด

3.1 ความหมายของความเครียด

ปิยะวรรณ เลิศพานิช (2542: 13) กล่าวว่า ความเครียด หมายถึง ภาวะที่ร่างกายและจิตใจตอบสนองต่อสิ่งที่มาคุกคามทั้งภายในและภายนอก เป็นความรู้สึกไม่พอใจ ไม่สบายใจเสมือนจิตใจถูกบีบบังคับให้ต้องเผชิญกับสิ่งเร้า ความรู้สึกดังกล่าวนี้ ทำให้เกิดความแปรปรวนทั้งทางร่างกายและจิตใจ

กรมสุขภาพจิต กระทรวงสาธารณสุข (2548: 67) ได้ให้ความหมายของความเครียดว่าเป็นสภาวะจิตใจและร่างกายที่เปลี่ยนแปลงไป เป็นผลจากการที่บุคคลต้องปรับตัวต่อสิ่ง ที่กระตุ้นหรือสิ่งเร้าต่างๆ ในสิ่งแวดล้อมที่กดดันหรือคุกคามให้เกิดความทุกข์ ความไม่สบายใจ

สมิต อาชนิจกุล (2542: 2) กล่าวว่า ความเครียด คือ สภาวะที่ซับซ้อนของร่างกายและจิตใจ อันเกิดจากภาวะแวดล้อม หรือสถานการณ์ที่บีบบังคับให้ร่างกาย และจิตใจเปลี่ยนแปลงไป และความเครียดจะยังคงอยู่เรื่อยไปจนกว่าสถานการณ์นั้นจะถูกกำจัดให้หมดไป ความเครียดเป็นสิ่งที่เกิดขึ้นได้กับทุกคนในทุกช่วงเวลา และเมื่อเกิดความเครียดในตัวบุคคลแล้วจะทำให้เกิดอาการที่ไม่พึงประสงค์ต่างๆ ทั้งทางร่างกายและจิตใจ เช่น หงุดหงิด กลั้ววิตกกังวล ปวดหัวเวียนหัว เกิดอาการทางกระเพาะ ทำให้อาหารไม่ย่อย นอนไม่หลับ อ่อนเพลีย หหมดเรี่ยวแรง ย้ำคิดย้ำทำ เกิดอาการผิดปกติทางประสาทหรือทางใจถ้าเกิดขึ้นบ่อยครั้ง หรือเกิดระยะเวลาานจะทำให้จิตใจปั่นป่วนและเกิดอาการทุกขใจที่แก้ไขได้ยากยิ่ง

อรพรรณ ลือบุญวัชชัย (2545: 253) กล่าวว่า ความเครียดคือสภาวะที่กดดันในบุคคลที่ เกิดขึ้นเมื่อมีสิ่งหนึ่งสิ่งใดมาคุกคาม ก่อให้เกิดความไม่สมดุลทั้งร่างกาย จิตใจ อารมณ์ สังคม และจิตวิญญาณของบุคคล รวมถึงพัฒนาการทางร่างกายและจิตใจของบุคคลด้วย หรืออาจกล่าวได้ว่าสภาวะเครียดเป็นกลไกการป้องกันตัวที่เกิดขึ้นเมื่อมีภัยอันตรายจากภายนอกมารบกวน

Sarafino, E.P. (Sarafino, E.P., 1990) ได้ให้ความหมายของความเครียดโดยสรุปว่าเป็นการตอบสนองต่อสิ่งที่คุกคามหรือกดดัน ซึ่งแบ่งเป็น 2 องค์ประกอบ คือ 1) องค์ประกอบด้านร่างกาย (Physiological stress) เช่น เหงื่อแตก หายใจถี่ตื่น กล้ามเนื้อเกร็ง ปวดศีรษะ เป็นต้น 2) องค์ประกอบด้านจิตใจ (Psychological stress) แบ่งออกเป็น ด้านพฤติกรรม, ด้านความคิด และด้านอารมณ์

กล่าวโดยสรุปได้ว่า ความเครียด คือ ความรู้สึกในตัวบุคคลที่ทำให้เกิดความไม่สบายทั้งร่างกาย และจิตใจ เป็นทุกข์ เป็นสภาวะกดดันที่มาจากสิ่งกระตุ้นต่างๆ ทั้งภายในและภายนอกโดยมีสาเหตุมาจากสภาพแวดล้อมที่บีบบังคับและมีผลกระทบต่อความรู้สึกนึกคิดของบุคคล โดยบุคคลจะเป็นผู้รับรู้ถึงภาวะดังกล่าว เพื่อเตรียมพร้อมที่จะเผชิญกับเหตุการณ์ ทำให้ร่างกายมีปฏิกิริยาตอบสนองต่อสิ่งที่มาคุกคามทั้งในด้านที่พึงประสงค์และไม่พึงประสงค์ ทั้งจากภายในและภายนอกบุคคล ในขณะที่เดียวกันถ้าบุคคลไม่สามารถปรับตัวกับสภาวะดังกล่าวได้ จะให้เกิดผลเสียต่อบุคคลทั้งทางร่างกายและจิตใจ ซึ่งจะมีผลต่อพฤติกรรมปรับตัว การทำงานและบุคลิกภาพของบุคคลที่ผิดปกติไปจากเดิม

3.2 ชนิดของความเครียด

กรมสุขภาพจิต กระทรวงสาธารณสุข (2546: 22) ได้แบ่งชนิดของความเครียดไว้มี 2 ประเภท คือ

1. ความเครียดที่เกิดขึ้นจากร่างกาย ซึ่งสามารถแบ่งออกตามระยะเวลาของการเกิดได้อีก 2 ชนิด คือ

- ความเครียดชนิดเฉียบพลัน (Emergency Stress) เป็นสิ่งคุกคามชีวิตที่เกิดขึ้นทันทีทันใด เช่น อุบัติเหตุต่างๆ
- ความเครียดชนิดต่อเนื่อง (Continuing Stress) เป็นสิ่งคุกคามที่เกิดขึ้นแล้วดำเนินต่อไปอย่างต่อเนื่อง ได้แก่ การเปลี่ยนแปลงของร่างกายในวัยต่างๆ ที่คุกคามความรู้สึก เช่น การเข้าสู่วัยรุ่น การตั้งครรภ์ ภัยหมดประจำเดือน เป็นต้น และสิ่งแวดล้อมที่ก่อให้เกิดความคับข้องใจอย่างต่อเนื่อง เช่น การจราจรติดขัด ฝุ่นละอองจากสิ่งก่อสร้าง เสียงดังรบกวนในโรงงาน เป็นต้น

2. ความเครียดที่เกิดจากจิตใจ เป็นสิ่งที่คุกคามที่สืบเนื่องมาจากความคิด บางครั้งก็เกิดขึ้นอย่างเฉียบพลัน เพราะเป็นการตอบสนองอย่างทันทีทันใด เช่น เมื่อถูกดูด่าก็จะเกิดความรู้สึกโกรธและเกิดความเครียดขึ้นได้ หรือเป็นสิ่งคุกคามที่ได้รับจากการอ่านหนังสือ จากการชมภาพยนตร์ จากคำบอกเล่าของผู้อื่นทำให้คิดว่าตนจะมีอันตรายจนก่อให้เกิดความเครียดขึ้นได้

3.3 สาเหตุของความเครียด

ความเครียดเกิดขึ้นได้จากสาเหตุหรือปัจจัยหลายๆ อย่าง ความเครียดที่เกิดขึ้นมีความแตกต่างกันในแต่ละบุคคล แม้จะมีสาเหตุเดียวกันแต่อาจมีความรุนแรงไม่เท่ากัน ทั้งนี้มีองค์ประกอบที่เกี่ยวข้องของบุคคลหมายถึงประการคือ พื้นฐานด้านจิตใจ อารมณ์ สังคม วัฒนธรรมและสิ่งแวดล้อมในขณะนั้น และยิ่งขึ้นอยู่กับระดับความมากน้อยของสิ่งที่มากระตุ้น

กรมสุขภาพจิต กระทรวงสาธารณสุข (2541:35-39) ได้ให้ความถึงปัจจัยที่ส่งผลต่อความเครียด คือปัจจัยส่วนบุคคลและปัจจัยด้านสิ่งแวดล้อมดังนี้

3.3.1 ปัจจัยส่วนบุคคล แบ่งได้ทั้งทางร่างกายและจิตใจ

3.3.1.1 ทางด้านร่างกาย สภาวะทางกายบางประการ ทำให้เกิดความเครียดได้ในลักษณะที่เรียกว่า ร่างกายเครียดซึ่งปัจจัยทางร่างกายที่ก่อให้เกิดความเครียด ได้แก่

- คุณลักษณะทางพันธุกรรม ที่ทำให้บุคคลมีลักษณะพื้นฐานที่ต่างกันไป ไม่ว่าจะเป็นเรื่อง เพศ สีผิว ความเข้มแข็งหรือความอ่อนแอ ของระบบการทำงานของร่างกายตลอดจนความพิการทางด้านร่างกายที่เป็นมาแต่กำเนิด
- ความเหนื่อยล้าทางร่างกาย อันเกิดจากการได้ผ่านการทำงานอย่างหนักและติดต่อกันเป็นเวลานาน ซึ่งจะมีความเกี่ยวเนื่องมาจากสภาพ

ความสมบูรณ์แข็งแรงของร่างกายที่จะทำให้แต่ละคนมีความพร้อมในการทำกิจกรรมในชีวิตประจำวันแตกต่างกัน

- การพักผ่อนไม่เพียงพอ
- การเจ็บป่วยทางร่างกาย
- ภาวะโภชนาการได้แก่ ลักษณะนิสัยการรับประทานอาหารที่จะมีผลต่อความสมบูรณ์แข็งแรงของร่างกาย เช่น การรับประทานอาหารไม่ถูกส่วน ไม่เพียงพอ หรือการบริโภคสารอาหารบางประเภท เช่น สุรา บุหรี่ กาแฟ สิ่งเสพติดต่างๆ ล้วนทำให้เกิดความเครียดได้
- ลักษณะท่าทางที่ปรากฏเกี่ยวกับโครงสร้างของกล้ามเนื้อผิวหนัง และการทรงตัว เช่น การเดิน วิ่ง นั่ง นอน หากอยู่ในลักษณะที่ไม่เหมาะสม ย่อมก่อให้เกิดความเครียดได้

3.3.1.2 ทางด้านจิตใจที่สำคัญและเกี่ยวข้องกับการเกิดความเครียดมากที่สุดได้แก่

- บุคลิกภาพแบบเอาจริงเอาจังกับชีวิต เป็นคนที่ต้องการความสมบูรณ์แบบให้กับทุกสิ่งทุกอย่าง ขยัน เจ้าระเบียบ ยึดมั่นกฎเกณฑ์ บุคลิกภาพนี้ทำให้ต้องทำงานหนักและทำทุกอย่างด้วยตนเอง
- บุคลิกภาพแบบพึ่งพิงผู้อื่น เป็นคนที่มีลักษณะขาดความเชื่อมั่นในตนเองไม่กล้าตัดสินใจ จะเป็นคนที่วิตกกังวลง่าย ทำอะไรต้องอาศัยผู้อื่น หวาดหวั่นกับอนาคตอยู่เสมอ
- บุคลิกภาพแบบหุนหันพลันแล่น เป็นคนใจร้อน คบคุมอารมณ์ไม่ได้ อารมณ์เปลี่ยนแปลงง่าย ซึ่งจะมีผลกระทบไปถึงภาวะสมดุลของร่างกาย

3.3.1.3 การเผชิญเหตุการณ์ต่างๆ ในชีวิต (Live Event) เป็นเหตุการณ์ความเปลี่ยนแปลงที่มีผลกระทบต่อการทำ กิจกรรมปกติของบุคคลโดยเหตุการณ์ที่เกิดขึ้น อาจก่อให้เกิดผลทางบวกหรือลบก็ได้ ทำให้บุคคลต้องปรับตัว

3.3.1.4 ความขัดแย้งในใจ (Conflict) เกิดจากการที่บุคคลต้องเผชิญทางเลือกสองทางขึ้นไป แต่ทางเลือกอย่างใดอย่างหนึ่ง อาจจะเป็นทางเลือกในสิ่งที่ไม่ต้องการไม่ยอมได้

3.3.1.5 ความคับข้องใจ (Frustration) เกิดจากการที่บุคคลไม่สามารถบรรลุเป้าหมายที่ต้องการได้ สาเหตุที่ก่อให้เกิดความคับข้องใจ 4 ประการ คือ

- ความรู้สึกบางประเภท ได้แก่ ความรู้สึกสูญเสีย อาจจะเป็นการสูญเสียคนที่รักเกียรติยศ ชื่อเสียง หรือความภาคภูมิใจ ความรู้สึกอันลุ่มเหลวในชีวิต ความรู้สึกเหล่านี้ทำให้หมดหวัง ไม่สามารถควบคุมสิ่งต่างๆ ได้
- อารมณ์ที่ไม่ดีทุกชนิด ได้แก่ ความกลัว ความวิตกกังวล ความโกรธ ความเศร้าความทุกข์ใจ เป็นต้น

- การขาดคุณสมบัติบางประการของบุคคล เช่น การขาดความสามารถ การขาดทักษะในการทำงาน เป็นต้น
- ความล่าช้าของเหตุการณ์หรือการไปสู่วิธีการที่ต้องการอันเกิดจากธรรมชาติหรือลักษณะของเหตุการณ์นั้น หรือเกิดจากปัจจัยแวดล้อมของเหตุการณ์หรือเกิดจากปัจจัยภายในตัวของบุคคล

3.3.2 ปัจจัยด้านสิ่งแวดล้อม

3.3.2.1 ทางกายภาพ เช่น อุณหภูมิ ภูมิอากาศ ความหนาแน่นของประชากร ภัยธรรมชาติ ฯลฯ ลักษณะทางกายภาพเหล่านี้มีผลกระทบต่อทั้งทางตรงและทางอ้อมต่อวิถีการดำเนินชีวิตของบุคคลและทำให้ต้องปรับตัวอยู่ตลอดเวลาซึ่งก่อให้เกิดความเครียดขึ้นได้

3.3.2.2 ทางชีวภาพ สิ่งแวดล้อมที่เป็นสิ่งมีชีวิตและทา ร้ายตัวบุคคลได้ เช่น โรคต่างๆ

3.3.2.3 ทางสังคมวัฒนธรรม ปรัชญาการณต่างๆ ที่เกิดขึ้นในสังคม และสิ่งทีบุคคลกระทำจะมีผลกระทบต่อกันตลอดเวลา ปัจจัยที่ก่อให้เกิดความเครียดได้แก่

- สถานภาพและบทบาทในสังคมของบุคคล ซึ่งอาจจะมีหลายสถานภาพและหลายบทบาท หากบทบาทหลายอย่างนี้มีความขัดแย้งหรือเกิดความแตกต่างกันมากก็อาจทำให้เกิดความยากลำบากในการปรับตัวและเครียดได้
- บรรทัดฐานทางสังคม บางครั้งกฎเกณฑ์ที่เป็นธรรมเนียมปฏิบัติของสังคมไม่สอดคล้องกับความต้องการของบุคคลหรือเป็นกฎเกณฑ์ที่เคร่งครัดมากย่อมก่อให้เกิดความเครียดแก่บุคคลได้
- ระบบการเมืองการปกครอง ซึ่งมีผลต่อวิถีการดำเนินชีวิตของบุคคลหากเป็นระบบการเมืองการปกครองที่เคร่งครัด ทำให้ประชาชนรู้สึกถูกคุกคามหรือถูกควบคุมโดยเฉพาะเรื่องสิทธิ เสรีภาพต่างๆ หรือเป็นระบบการเมืองการปกครองที่ทำให้ประเทศขาดความมั่นคงย่อมก่อให้เกิดกับความเครียดกับประชาชนได้

กนกรัตน์ สุขตุงคะ (2545: 2) ได้อธิบายสาเหตุที่ทำให้บุคคลมีโอกาสดเกิดความเครียดได้ง่าย

ดังนี้

- รู้สึกว่าตัวเองมีปมด้อย
- มีความรู้สึกเชิงแข่งขันสูงต้องการเอาชนะ
- ขาดความยืดหยุ่น ทางความคิด เอาตัวเองเป็นหลัก
- คนที่ชอบเก็บกดไม่กล้าแสดงความรู้สึกจริงๆ ของตนแม้แต่เรื่องเล็กๆ น้อยๆ
- คนที่ยอมรับความแตกต่างของตนเองไม่ได้
- คนที่มีบุคลิกภาพอ่อนแอมาก อ่อนไหวกับเรื่องต่างๆ

- คนซีระแวง ซ้ำสงสัยไม่ไว้วางใจใคร
- เป็นโรคเรื้อรังหรือโรคร้ายแรง
- อยู่ในสถานะที่ไม่สบายใจ โดยหลีกเลี่ยงไม่ได้ เช่น ต้องอยู่กับงานที่ไม่ชอบ หรือเพื่อนร่วมงานที่สร้างปัญหา
- อยู่ในสิ่งแวดล้อมที่เป็นพิษ

วารสารณ์ ตระกูลสฤษดี (2547:117-118) ได้กล่าวไว้ว่าสาเหตุของความเครียดเป็นสิ่งที่มียอิทธิพลต่ออารมณ์สภาวะร่างกายของบุคคลเป็นสิ่งที่คาดว่าจะมีผลก่อให้เกิดความเครียดขึ้นได้ และการเกิดความเครียดนั้นมาจากสาเหตุ 2 ประการคือ

1. จากตัวบุคคลเอง ซึ่งเกี่ยวข้องกับสติปัญญา ประสบการณ์ชีวิต บุคลิกภาพ เฉพาะบุคคล เป็นแบบก้าวร้าว เก็บตัวต่อต้านสังคม รักและหลงตนเองมาก ความคับข้องใจ ความรู้สึกสูญเสีย และรู้สึกผิด ความพิการ ความเจ็บปวดทางด้านร่างกายและอื่นๆ ที่เกี่ยวข้องกับลักษณะของปัจเจกบุคคล
2. จากภายนอกตัวบุคคล เป็นเรื่องของปัญหาที่เกิดจากความสัมพันธ์ระหว่างบุคคล เช่น ปัญหาจากอาชีพการงาน เศรษฐกิจและสังคม ความยากจนและสภาพของสังคมเมืองการเปลี่ยนแปลงชีวิต ความรับผิดชอบ และตำแหน่งหน้าที่การงานและอื่นๆ ที่เกี่ยวข้องและแวดล้อมบุคคลนั้นๆ อยู่

ไพบูลย์ โล่สุนทร (2547: 18) กล่าวถึงสิ่งที่ทำให้เกิดโรคทางด้านจิตใจ และสังคม (Psychosocial agent) สิ่งที่ทำให้เกิดโรคเนื่องมาจากปัจจัยด้านจิตใจและสังคม สภาพปัญหาทางเศรษฐกิจและสังคมปัจจุบัน เงินเดือนหรือรายได้จากการประกอบอาชีพที่ได้รับไม่ได้สัดส่วนที่ดัชนีค่าครองชีพที่สูงขึ้น รายได้ไม่สมดุลกับรายจ่าย ทำให้เกิดความวิตกกังวลเกี่ยวกับการเลี้ยงชีพเกิดความตึงเครียดทางอารมณ์ เป็นสาเหตุของโรคต่างๆ เช่น โรคความดันโลหิตสูง โรคแผลในกระเพาะอาหาร เป็นต้น ความสัมพันธ์ที่ไม่ดีกับเพื่อนร่วมงานกับผู้บังคับบัญชาหรือผู้ใต้บังคับบัญชาอาจก่อให้เกิดความเครียดขึ้นได้ ผลของความเครียดในระยะยาวนานก่อให้เกิดโรคทางกายและทางใจ

จากสาเหตุต่างๆ ดังที่กล่าวมา สามารถสรุปได้ว่าสาเหตุที่ทำให้เกิดความเครียด ได้แก่ สาเหตุภายในและภายนอกตัวบุคคล สำหรับสาเหตุภายใน แบ่งได้เป็นสาเหตุทางร่างกายและสาเหตุทางจิตใจ สาเหตุทางร่างกาย ได้แก่ ความเหนื่อยล้าทางร่างกาย ไม่ได้รับการพักผ่อนเพียงพอ รับประทานอาหารไม่เพียงพอ ความเจ็บป่วยทางร่างกาย และการมีความรู้สึกสูญเสียและรู้สึกผิด ความคับข้องใจ ตลอดจนการเปลี่ยนแปลงในชีวิตของแต่ละบุคคล

3.4 การประเมินความเครียด

เมื่อเกิดความเครียดขึ้นจะมีการเปลี่ยนแปลงทางร่างกายและจิตใจที่จะทราบว่าบุคคลมีความเครียดหรือไม่ และมากน้อยแค่ไหนนั้นการวิจัยครั้งนี้ผู้วิจัยได้ใช้แบบวัดความเครียดจากกรมสุขภาพจิต กระทรวงสาธารณสุข (2541: 90) เพื่อวัดความเครียดของกลุ่มตัวอย่างแบ่งได้ดังนี้

1. ความผิดปกติทางร่างกาย ได้แก่ ปวดศีรษะ ไมเกรน ท้องเสียหรือท้องผูก นอนไม่หลับหรือง่วงเหงาหาวนอนตลอดเวลา ปวดเมื่อยกล้ามเนื้อ เบื่ออาหารหรือกินมากกว่าปกติ ท้องอืดเฟ้อ อาหารไม่ย่อย ประจำเดือนมาไม่ปกติ เสื่อมสมรรถภาพทางเพศ มือเย็นเท้าเย็น เหงื่อออกตามมือตามเท้า ใจสั่น ถอนหายใจบ่อยๆ ผิวหนังเป็นผื่นคัน เป็นหวัดบ่อยๆ แพ้อากาศง่าย ฯลฯ

2. ความผิดปกติทางจิตใจ ได้แก่ ความวิตกกังวล คิดมาก คิดฟุ้งซ่าน หลงลืมง่าย ไม่มีสมาธิ หงุดหงิด โกรธง่าย ใจน้อย เปื่อหน่าย ซึมเศร้า เหงา ว่าเหว่ สิ้นหวัง หมดความรู้สึกสนุกสนาน เป็นต้น

3. ความผิดปกติทางพฤติกรรม ได้แก่ สูบบุหรี่ ดื่มสุรามากขึ้น ใช้สารเสพติด ใช้นานอนหลับ จู้จี้ขี้บ่น ขวนทะเลาะ มีเรื่องขัดแย้งกับผู้อื่นบ่อยๆ ดึงผม กัดเล็บ กัดฟัน ผุดลุกผุดนั่ง เจียบขริ้ม เก็บตัว เป็นต้น

จากแบบประเมินข้างต้นสามารถนำมาวิเคราะห์ความเครียดได้ 5 ระดับดังนี้

1. ระดับคะแนน 0-5 หมายถึง มีความเครียดอยู่ในระดับต่ำกว่าเกณฑ์ปกติ ความเครียดในระดับต่ำมากเช่นนี้ อาจหมายความว่าผู้ตอบแบบประเมินตอบไม่ตรงตามความเป็นจริง อาจเข้าใจคำถามคลาดเคลื่อนไป ผู้ตอบอาจเป็นคนที่ขาดแรงจูงใจ มีความเฉื่อยชา หรือชีวิตประจำวันซ้ำซาก จำเจน่าเบื่อ ปราศจากความตื่นเต้น

2. ระดับคะแนน 6-17 หมายถึง มีความเครียดอยู่ในเกณฑ์ปกติ สามารถจัดการกับความเครียดที่เกิดขึ้นในชีวิตประจำวันและสามารถปรับตัวกับสถานการณ์ต่างๆ ได้อย่างเหมาะสมรู้สึกพึงพอใจเกี่ยวกับตนเองและสิ่งแวดล้อมเป็นอย่างมาก ความเครียดในระดับนี้ถือว่ามีประโยชน์ในการดำเนินชีวิตประจำวัน เป็นแรงจูงใจที่นำไปสู่ความสำเร็จในชีวิตได้

3. ระดับคะแนน 18-25 หมายถึง มีความเครียดอยู่ในระดับสูงกว่าปกติเล็กน้อย ซึ่งถือว่าเป็นความเครียดที่พบได้ในชีวิตประจำวันอาจไม่รู้ว่ามีความเครียดหรืออาจรู้สึกได้จากการเปลี่ยนแปลงของร่างกาย อารมณ์ ความรู้สึกและพฤติกรรมบ้างเล็กน้อย แต่ไม่ชัดเจนและยังพอทนได้ อาจต้องใช้เวลาในการปรับตัวแต่ในที่สุดก็สามารถจัดการกับความเครียดได้และความเครียดระดับนี้ไม่เป็นผลเสียต่อการดำเนินชีวิต ในกรณีนี้ สามารถผ่อนคลายความเครียดด้วยการหากิจกรรมที่เพิ่มพลัง เช่น การออกกำลังกาย เล่นกีฬา ทำสิ่งที่สนุกสนานเพลิดเพลิน เช่น ดูหนัง ฟังเพลง อ่านหนังสือ หรือทำงานอดิเรกต่างๆ หากต้องการป้องกันไม่ให้ความเครียดมากขึ้น ในอนาคตอาจฝึกความเครียดตามวิธีในหัวข้อต่อไปนี้อย่าลืมนพูดคุยกับผู้ที่ไว้วางใจ พิจารณาและลงมือแก้ไขปัญหาลำดับความสำคัญอย่างรอบคอบและมีสติ

4. ระดับคะแนน 26-29 หมายถึง มีความเครียดอยู่ในระดับสูงกว่าปกติปานกลาง ขณะนี้เริ่มมีความตึงเครียดในระดับค่อนข้างสูงและได้รับความเดือนร้อนเป็นอย่างมากจากปัญหาทางอารมณ์ที่เกิดจากปัญหาความขัดแย้ง และวิกฤตการณ์ ในชีวิตเป็นสัญญาณเตือนขั้นต้นว่ากำลังเผชิญกับภาวะวิกฤตและความขัดแย้ง ซึ่งจัดการแก้ไขด้วยความยากลำบากลักษณะดังกล่าวจะเพิ่มความรุนแรง ซึ่งมีผลกระทบต่อการทำงานจำเป็นต้องหาวิธีการแก้ไขข้อขัดแย้งต่างๆ ให้ลดน้อยลงหรือหมดไปด้วยวิธีการอย่างใด อย่างหนึ่ง สิ่งแรกที่ต้องการบริหารจัดการคือ ต้องมีวิธีคลายเครียดที่ดีและสม่ำเสมอทุกวัน วันละ 1-2 ครั้ง ครั้ง ละ 10 นาที โดยนั่งในท่าที่สบายหายใจลึกๆ ให้หน้าท้องขยาย หายใจออกช้าๆ นับ 1-10 ไปด้วย หรือจะใช้วิธีการนั่งสมาธิหรือสวดมนต์ก็ได้ ควรแก้ไขปัญหาให้ดีขึ้น โดยค้นหาสาเหตุของปัญหาที่ทำให้เกิดความขัดแย้งหาวิธีแก้ไขปัญหาหลายๆวิธี พร้อมทั้งพิจารณาผลดี ผลเสียของแต่ละวิธี เลือกวิธีการที่เหมาะสมกับสภาวะของตนเองมากที่สุด ทั้งนี้ต้องไม่สร้างปัญหาให้เพิ่มขึ้น หรือทำให้ผู้อื่นเดือนร้อนวางแผนแก้ไขปัญหาลำดับขั้นตอน และลงมือแก้ไขปัญหา หากไม่สามารถจัดการคลี่คลายหรือแก้ปัญหาด้วยตนเองได้ควรปรึกษากับผู้ให้การรักษาปัญหาสุขภาพจิตตามหน่วยงานต่างๆ

5. ระดับคะแนน 30-60 คือมีความเครียดอยู่ในระดับสูงกว่าปกติมาก กำลังตกอยู่ในภาวะตึงเครียดหรือกำลังเผชิญกับวิกฤตการณ์ในชีวิตอย่างรุนแรง เช่น การเจ็บป่วยที่รุนแรงเรื้อรัง ความพิการ การสูญเสีย ปัญหาความรุนแรงในครอบครัว ปัญหาเศรษฐกิจซึ่งมีผลต่อสุขภาพกาย และสุขภาพจิตอย่างชัดเจน ทำให้ชีวิตไม่มีความสุข ความคิดฟุ้งซ่าน ตัดสินใจผิดพลาด ขาดความยับยั้งชั่งใจ อาจเกิดอุบัติเหตุได้ง่าย บางครั้งอาจมีพฤติกรรมก้าวร้าวรุนแรง เช่น เอะอะโวยวาย ขว้างปาข้าวของ ความเครียดในระดับนี้ถือว่ามีความรุนแรงมาก หากปล่อยไว้โดยไม่ดำเนินการแก้ไขอย่างเหมาะสมและถูกวิธี อาจนำไปสู่ความเจ็บป่วยทางจิตที่รุนแรง ซึ่งส่งผลเสียต่อตนเองและบุคคลใกล้ชิดต่อไปได้ในระดับนี้ควรต้องไปปรึกษาหรือใช้บริการปรึกษาปัญหาสุขภาพจิตทางโทรศัพท์ ซึ่งจะช่วยให้มองเห็นปัญหาและแนวทางแก้ไขที่ชัดเจนและเหมาะสมต่อไป

ความเครียดที่เกิดขึ้นในแต่ละบุคคลนั้นจะมีระดับมากหรือน้อยนั้น ขึ้นอยู่กับความสามารถในการรับรู้และการปรับตัวของบุคคลต่อสิ่งที่มาคุกคาม สาเหตุของความเครียดเดียวกันสามารถก่อให้เกิดการรับรู้และการตอบสนองแตกต่างกันไปในแต่ละบุคคล นอกจากกรมสุขภาพจิตแล้ว จันท์แรม กิจเหมาะ (2539: 23) ยังแบ่งความเครียดไว้ 4 ระดับ ดังนี้

1. ความเครียดระดับปกติ (Normal Stress) หรือความเครียดที่เกิดขึ้นในชีวิตประจำวันเป็นความเครียดที่คนส่วนใหญ่เผชิญอยู่บ่อยๆ จะมีผลกระทบต่อชีวิตเพียงชีวิตเพียงเล็กน้อย ทำให้สามารถทำได้นานและขยับยั้งขึ้น ความเครียดในระดับนี้เกิดในช่วงเวลาสั้นๆ และสามารถกลับคืนสู่ภาวะปกติได้โดยการพักผ่อนวันสุดสัปดาห์ในยามค่ำคืนและวันหยุดที่เป็นครั้งคราว

2. ความเครียดในระดับรุนแรง (Sever Stress) เป็นความเครียดที่เกิดขึ้นเป็นเวลานานต้องได้รับ การบำบัดจากผู้เชี่ยวชาญ ความเครียดในระดับนี้อาจทำให้เป้าหมายในการดำเนินชีวิตลดลงหรือมากขึ้นก็ได้ ถ้าได้รับการจัดการความเครียดในระดับนี้ก็กลับสู่ความเครียดระดับปกติได้

3. ความเครียดในระดับปกติ (Abnormal Stress) ความเครียดในระดับนี้เกี่ยวข้องกับอาการทางโรคประสาทและโรคจิต อาการที่ปรากฏอาจก่อให้เกิดอันตรายต่อตนเองได้และผู้อื่น เช่น มีอาการของโรคซึมเศร้า (Melancholia) มีอาการหลงผิด (Delusion) และมีพฤติกรรมไปจากเดิม

4. ความเครียดระดับสูง (High stress) ความเครียดระดับนี้จะอยู่นานนับสัปดาห์หรืออาจเป็นเดือน หรือเป็นปีก็ได้ เช่น การตายยาก การเจ็บป่วยที่รุนแรง หรือเครียดมากร่างกายและจิตใจแพ้ต่อความเครียด มีการเปลี่ยนแปลงทางร่างกายและจิตใจอย่างเด่นชัดทำให้การดำเนินชีวิตเสียไป การตัดสินใจผิดพลาด ช่วงระยะนี้ต้องการดูแลช่วยเหลือและบำบัดรักษา แม้สาเหตุที่ทำให้เกิดความเครียดจะหมดไปหรือรู้จักผ่อนคลายความเครียดด้วยตนเอง แต่ก็ยังไม่สามารถทำให้ความเครียดหมดไปได้โดยง่าย การช่วยเหลือตนเองอย่างเดียวไม่เพียงพอ ต้องมีผู้อื่นเข้ามาช่วยเหลืออีกด้วย

กล่าวโดยสรุป ความเครียดระดับต่างๆ มีได้ตั้งแต่ระดับความเครียดในระดับที่ต่ำหรือปริมาณน้อย ความเครียดระดับนี้ เป็นภาวะที่บุคคลสามารถปรับตัวกับสถานการณ์ที่กดดันจนเกิดความเคยชิน ไม่มีแรงกดดันในชีวิต ฟังพอใจในสิ่งที่มีอยู่ และมีแรงจูงใจต่ำในการดำเนินชีวิตในด้านต่างๆ ในระดับความเครียดที่เป็นปกติ เป็นภาวะสมดุลที่บุคคลสามารถจัดการกับปัญหาหรือสถานการณ์ต่างๆ ได้เป็นอย่างดีและมีแรงจูงใจในการดำเนินชีวิต สามารถปรับตัวในการดำเนินชีวิตได้อย่างอัตโนมัติ หากระดับความเครียดมีมากขึ้นเล็กน้อย เป็นความเครียดที่ระดับปานกลางก็จะช่วยให้บุคคลนั้นๆ มีความกระตือรือร้น มีความตื่นตัวสูง สำหรับความเครียดระดับนี้เป็นต้นเหตุของความเครียดที่เริ่มรุนแรงใน

ระดับแรกและต้องให้ความสนใจอย่างมาก เนื่องจากบุคคลยังไม่สามารถปรับตัวได้ในระยะเวลาอันสั้น ความเครียดระดับสูงมาก จะมีปฏิกิริยาที่แสดงออกจาจะเป็นลักษณะการปฏิเสธ ก้าวร้าว พุดน้อย กัดฟัน ชีวิตไม่มีความสุข อันเป็นความเครียดที่เรื้อรังแรงและมีผลกระทบต่อตัวบุคคลและผู้อื่น ระดับความเครียดที่สูงมากจะทำให้บุคคลหมดกำลังท้อแท้และเบื่อหน่ายในชีวิต ไม่สามารถจัดการหรือปรับตัวได้ อาจส่งผลให้เกิดการเจ็บป่วยทั้งทางร่างกายและจิตใจ รวมถึงส่งผลในระดับองค์กรและประเทศอย่างมาก

4. โปรแกรม WEKA

โปรแกรม WEKA (Waikato Environment for Knowledge Analysis) : เริ่มพัฒนามาตั้งแต่ปี 1997 โดยมหาวิทยาลัย Waikato ประเทศนิวซีแลนด์ เป็นซอฟต์แวร์สำเร็จ ภาพประกอบประเภทฟรีแวร์ อยู่ภายใต้การควบคุมของ GPL License ซึ่งโปรแกรม WEKA ได้ถูกพัฒนามาจากภาษาจาวาทั้งหมด ซึ่งเขียนมาโดยเน้นกับงานทางด้านการเรียนรู้ด้วยเครื่อง (Machine Learning) และการทำเหมืองข้อมูล (Data Mining) โปรแกรมจะประกอบไปด้วยโมดูลย่อยๆ สำหรับใช้ในการจัดการข้อมูล และเป็นโปรแกรมที่สามารถใช้ Graphic User Interface (GUI) และใช้คำสั่งในการให้ซอฟต์แวร์ประมวลผล และสามารถทำงานได้บนหลายระบบปฏิบัติการและสามารถพัฒนาต่อยอดโปรแกรมได้ โปรแกรม WEKA เป็นเครื่องมือที่ใช้ทำงานในด้านการทำเหมืองข้อมูลที่รวบรวมแนวคิด อัลกอริทึมมากมาย ซึ่งอัลกอริทึมสามารถเลือกใช้งานโดยตรงได้จาก 2 ทางคือจากชุดเครื่องมือที่มีอัลกอริทึมมาให้ หรือเลือกใช้จากอัลกอริทึมที่ได้เขียนเป็นโปรแกรมลงไปเป็นชุดเครื่องมือเพิ่มเติมและชุดเครื่องมือมีฟังก์ชันสำหรับการทำงานร่วมกับข้อมูล ได้แก่ Pre-Processing, Classification, Regression, Clustering, Association rules, Selection และ Visualization โลโก้ของซอฟต์แวร์ WEKA เป็นรูปนกเฉพาะถิ่นของประเทศนิวซีแลนด์

ข้อดีของซอฟต์แวร์ WEKA

1. เป็นซอฟต์แวร์เสรีที่สามารถดาวน์โหลดได้ฟรี
2. สามารถทำงานได้ทุกระบบปฏิบัติการ
3. เชื่อมต่อ SQL Database โดยใช้ Java Database Connectivity
4. มีการเตรียมข้อมูลและเทคนิคในการสร้างแบบจำลองที่ครอบคลุม
5. มีลักษณะที่ง่ายต่อการใช้งาน

ความสามารถของซอฟต์แวร์ WEKA

1. สนับสนุนเกี่ยวกับการทำเหมืองข้อมูล (Data Mining)
2. การเตรียมข้อมูล (Data Preprocessing)
3. การทำเหมืองข้อมูลด้วยเทคนิคการจำแนกข้อมูล (Classification)
4. การทำเหมืองข้อมูลด้วยเทคนิคการจัดกลุ่ม (Clustering)
5. การทำเหมืองข้อมูลด้วยเทคนิคการวิเคราะห์ความสัมพันธ์ (Associating)
6. เทคนิคการคัดเลือกข้อมูล (Selecting Attributes)
7. เทคนิคการนำเสนอข้อมูลด้วยรูปภาพ (Visualization)

สรุปข้อดีของโปรแกรม WEKA คือ มีอัลกอริทึมที่รู้จักกันดีของการทำเหมืองข้อมูลให้เลือกใช้อย่างครบถ้วนและสามารถเขียนฟังก์ชันเพิ่มเข้าไปในโปรแกรมเองได้

5. งานวิจัยที่เกี่ยวข้อง

ศุภกรใจ วุฒิกิจโกศล (2551,น.77) ใช้เทคนิคการทำเหมืองข้อมูลในผู้ป่วยข้อไหล่นัดโรงพยาบาลพระนั่งเกล้าเพื่อหาปัจจัยที่มีผลต่อการรักษาและสร้างกฎความสัมพันธ์ที่น่าสนใจเพื่อเป็นแนวทางช่วยสนับสนุนการตัดสินใจเลือกเทคนิคการรักษาสำหรับนักกายภาพบำบัดโรงพยาบาลพระนั่งเกล้า โดยใช้ข้อมูลผู้ป่วยโรคข้อไหล่นัดที่ไม่มีโรคประจำตัวและไม่ได้เกิดจากอุบัติเหตุมารับการรักษาทางกายภาพบำบัดโรงพยาบาลพระนั่งเกล้าที่รักษาหายในช่วงปี พ.ศ. 2548-2550 จำนวน 550 ราย มีการเตรียมข้อมูลเพื่อให้ได้ข้อมูลที่ถูกต้อง และหาปัจจัยที่มีผลต่อการรักษา จากนั้นวิเคราะห์โดยการทำเหมืองข้อมูลด้วยเทคนิค Clustering และ Association Rule เลือกกฎความสัมพันธ์ที่น่าสนใจด้วยค่าสนับสนุนของกฎไม่ต่ำกว่าร้อยละ 20 และค่าความเชื่อมั่นไม่ต่ำกว่าร้อยละ 90 ผลการศึกษาพบว่า มีปัจจัย อายุ ระยะเวลาที่เป็น องศาการยกแขน การไขว้หลัง ระดับความเจ็บปวด และเทคนิคการรักษาที่มีผลต่อการรักษา ตัวอย่างเช่น เมื่อใช้เทคนิค Association Rule หากกฎความสัมพันธ์ช่วงองศาการยกแขน 90-120 เลือกเทคนิคการรักษาด้วยแผ่นประคบความร้อนบริเวณคอ ความร้อนคลื่นเหนือเสียงบริเวณไหล่ด้านหน้า ความร้อนคลื่นเหนือเสียงบริเวณกล้ามเนื้อ Infraspinus และกระตุ้นไฟฟ้าแบบ Surge บริเวณไหล่ด้านหน้า-หลังจะใช้จำนวนครั้งในการรักษาช่วง 19-24 ครั้งด้วยค่าสนับสนุนของกฎ 43.48 เปอร์เซนต์และค่าความเชื่อมั่นที่ 100%

รักถิ่น เหลลหา (2553,น.45) ศึกษาการพยากรณ์ความเสี่ยงการเกิดโรคมะเร็งปอด โดยใช้ทฤษฎีของการทำเหมืองข้อมูล สร้างระบบที่จัดกลุ่มของผู้ป่วยและพยากรณ์ความเสี่ยงการเกิดโรคมะเร็งปอด ทั้งนี้การจัดกลุ่มจะให้ทราบถึงค่าความเสี่ยงของปัจจัยเสี่ยงที่ทำให้เกิดโรคมะเร็งปอด จากนั้นนำค่าปัจจัยเสี่ยงมาวิเคราะห์และพยากรณ์ผู้ป่วยด้วยต้นไม้ตัดสินใจแบบ C4.5 โดยใช้กลุ่มข้อมูลผู้ป่วยโรงพยาบาลมหาสารคาม ในเดือนสิงหาคม – ธันวาคม 2552 จำนวน 2,215 คน ผู้ป่วยมีสองกลุ่ม คือ ผู้ป่วยที่เป็นโรคมะเร็งปอดจำนวน 118 คน ไม่เป็นมะเร็งปอด 2,097 คนและได้วัดค่าประสิทธิภาพการพยากรณ์จากค่าความแม่นยำและค่าความระลึก จากผลการวิจัยพบว่าปัจจัยที่ส่งผลต่อค่าความเสี่ยงที่ทำให้เป็นมะเร็งปอดมากที่สุดคือ ปัจจัยประวัติด้านพันธุกรรม โดยมีค่าความเสี่ยงเป็น 34.59 เท่าของคนที่ไม่ใช่ประวัติด้านพันธุกรรม รองลงมาคือ ปัจจัยเสี่ยง ด้านประวัติการสูบบุหรี่ ประวัติการดื่มสุรา และอายุ ตามลำดับ ผลการพยากรณ์ให้ค่าความแม่นยำ (precision) ที่ 0.796 ค่าความระลึก (recall) 0.5 และค่าความถูกต้อง (accuracy) ของการจำแนกประเภทเฉลี่ยมีค่าร้อยละ 96.8387

มาลิวรรณ บุญพลอย (2551,น.73) ได้ศึกษาเรื่องเหมืองข้อมูลและการค้นหาคำค้นหาคำความรู้สำหรับโรคกล้ามเนื้ออ่อนแรง นำเสนอวิธีการใช้ขั้นตอนวิธีต้นไม้ช่วยตัดสินใจสองชั้นเพื่อจำแนกประเภทข้อมูลโรคกล้ามเนื้ออ่อนแรงเรื้อรัง ซึ่งกระบวนการทำงานของขั้นตอนวิธีนี้สามารถแบ่งออกเป็นสองขั้นตอนใหญ่ๆ คือ ขั้นตอนการจำแนกกลุ่มผู้ป่วยปกติ (Non-fatigued -NF) และขั้นตอน

การจำแนกกลุ่มผู้ป่วยที่มีอาการผิดปกติแต่ไม่ใช่เป็นโรคกล้ามเนื้ออ่อนแรงเรื้อรัง (Insufficient number of symptoms or fatigue severity - ISF) และ กลุ่มผู้ป่วยที่เป็นโรคกล้ามเนื้ออ่อนแรงเรื้อรัง (Chronic Fatigue Syndrome - CFS) วัตถุประสงค์ของงานวิจัยนี้เพื่อทำการพัฒนาอัลกอริทึมที่สามารถจำแนกประเภทข้อมูลของโรคกล้ามเนื้ออ่อนแรงเรื้อรังให้มีประสิทธิภาพมากขึ้นและทำการค้นหาองค์ความรู้ซึ่งจะทำการเปรียบเทียบประสิทธิภาพของวิธีต้นไม้ช่วยตัดสินใจสองชั้น (Double Decision Trees) วิธีกลุ่มของต้นไม้ (Decision Tree Forests) และวิธีต้นไม้ช่วยตัดสินใจ (Decision Trees) ข้อมูลที่นำมาใช้ในการทดลอง ประกอบด้วยข้อมูลของผู้ป่วยที่เป็นโรคกล้ามเนื้ออ่อนแรงเรื้อรัง ซึ่งข้อมูลของผู้ป่วยที่เป็นโรคกล้ามเนื้ออ่อนแรงนั้นประกอบไปด้วยข้อมูลผลเลือด และข้อมูลอาการป่วย ผลการวิจัยพบว่าประสิทธิภาพในการจำแนกประเภทข้อมูลโรคกล้ามเนื้ออ่อนแรงที่เรื้อรังโดยใช้ขั้นตอนวิธีต้นไม้ช่วยตัดสินใจสองชั้น นั้นสามารถจำแนกประเภทของผู้ป่วยที่เป็นโรคกล้ามเนื้ออ่อนแรงเรื้อรังได้แม่นยำกว่าวิธีกลุ่มของต้นไม้และวิธีต้นไม้ช่วยตัดสินใจ โดยวิธีต้นไม้ช่วยตัดสินใจสองชั้นสามารถจำแนกข้อมูลกลุ่ม NF, CFS และ ISF มีค่าความแม่นยำเท่ากับ 83.60%, 83.30% และ 83.34% ตามลำดับ วิธีกลุ่มของต้นไม้สามารถจำแนกข้อมูลกลุ่ม NF, CFS และ ISF มีค่าความแม่นยำเท่ากับ 81.94%, 82.20% และ 81.79% ตามลำดับ วิธีต้นไม้ช่วยตัดสินใจสามารถจำแนกข้อมูลกลุ่ม NF, CFS และ ISF มีค่าความแม่นยำเท่ากับ 67.00%, 66.26% และ 67.00% ตามลำดับ เนื่องจากวิธีต้นไม้ช่วยตัดสินใจสองชั้นมีการแบ่งการจำแนกประเภทออกเป็น 2 ขั้นตอนโดยอาศัยหลักการคัดกรองข้อมูลผู้ป่วย ซึ่งทำให้ปัญหาการจำแนกประเภทข้อมูลลดความซับซ้อนลงจากเดิมเนื่องจากปัญหาในการจำแนกข้อมูลเป็นแบบหลายคลาส (Multi-class Problem) ได้ถูกลดรูปเป็นปัญหาแบบสองกลุ่ม (Two-class Problem) ซึ่งเป็นปัญหาที่ง่ายทำให้ขั้นตอนวิธีต้นไม้ช่วยตัดสินใจสองชั้นสามารถจำแนกข้อมูลได้อย่างมีประสิทธิภาพมากกว่า

อังคณา พิจารโชติ (2552, น. 73) ได้พัฒนาระบบสนับสนุนการตัดสินใจสำหรับวิเคราะห์ปัจจัยเสี่ยงการเป็นโรคเบาหวานโดยใช้เทคนิคการทำเหมืองข้อมูล เพื่อช่วยในการวิเคราะห์หาความสัมพันธ์ของปัจจัยเสี่ยงต่างๆ ที่ก่อให้เกิดโรคเบาหวาน เช่น พ่อแม่มีประวัติเป็นโรคเบาหวาน มีประวัติการปัสสาวะบ่อย เป็นต้น ข้อมูลความสัมพันธ์ของปัจจัยเสี่ยงที่ได้จากการวิเคราะห์จะเป็นสารสนเทศสำคัญที่เป็นตัวช่วยให้หน่วยงานด้านการแพทย์นำไปใช้สำหรับวางแผนป้องกันโรคโดยระบบที่พัฒนานี้จะประกอบด้วยสามส่วนคือ ส่วนของคลังข้อมูลใช้สำหรับเก็บข้อมูลปัจจัยเสี่ยงที่ได้จากการคัดกรองกลุ่มเสี่ยง ส่วนของเหมืองข้อมูลเป็นตัวค้นหาความสัมพันธ์ของปัจจัยเสี่ยงต่างๆ และส่วนการแสดงผลข้อมูลรายงานซึ่งจะจัดทำในรูปแบบของเว็บแอปพลิเคชันเพื่อให้สะดวกต่อการเรียกใช้งาน ผลการค้นหาค่าความสัมพันธ์ด้วยค่าสนับสนุนที่ 0.5 และค่าความเชื่อมั่นที่ 0.3 พบว่าประวัติครอบครัวเป็นปัจจัยเสี่ยงที่มีอัตราการเสี่ยงมากที่สุดซึ่งพบคู่กับค่าดัชนีมวลกาย มีค่าความเชื่อมั่นที่ 0.94

ในปี 2538 กรมสุขภาพจิต กระทรวงสาธารณสุข ได้ให้ทุนวิจัยโครงการพัฒนาแบบประเมินและวิเคราะห์ความเครียดด้วยตนเองสำหรับประชาชนไทยด้วยคอมพิวเตอร์ (กรมสุขภาพจิต กระทรวงสาธารณสุข, 2541) รวมทั้งคู่มือการใช้โปรแกรมคอมพิวเตอร์ทำแบบทดสอบดังกล่าวขึ้น (กรมสุขภาพจิต กระทรวงสาธารณสุข, 2541) วัตถุประสงค์ของงานวิจัยนี้คือ เพื่อพัฒนาแบบประเมินในการคัดกรองบุคคลที่มีความเครียด แนวโน้มลักษณะบุคลิกภาพที่สัมพันธ์กับความเครียด รวมทั้งหา

แนวทางในการจัดการกับความเครียดและแนวโน้มลักษณะบุคลิกภาพที่สัมพันธ์กับความเครียด จากผลการวิจัยพบว่า แบบประเมินดังกล่าวมีคุณภาพเพียงพอในการนำมาใช้คัดกรองบุคคลที่มีความเครียดแนวโน้มลักษณะของบุคลิกภาพที่สัมพันธ์กับความเครียดพบว่ามี 3 ลักษณะ คือ แนวโน้มลักษณะบุคลิกภาพแบบวิตกกังวล, แนวโน้มลักษณะบุคลิกภาพแบบซึมเศร้า และแนวโน้มแบบที่จะเกิดอาการเจ็บป่วยทางร่างกายเมื่อมีความเครียด แบบประเมินนี้มีข้อคำถามเพื่อใช้ทดสอบความเครียดจำนวน 20 ข้อ ซึ่งปัจจุบันได้ถูกนำมาใช้กันอย่างแพร่หลาย เช่น คู่มือคลายเครียด (สถาบันกัลยาณ์ราชนครินทร์, 2555) งานวิจัยชิ้นนี้จะช่วยให้ประชาชนตระหนักรู้ถึงความเครียด โดยสามารถประเมินความเครียดที่เกิดขึ้นได้ด้วยตนเอง พร้อมกับมีคำแนะนำเพื่อใช้เป็นแนวทางในการจัดการกับความเครียดอย่างเหมาะสม

อย่างไรก็ตาม งานวิจัยชิ้นนี้ยังมีจุดที่สามารถนำมาพัฒนาต่อยอดได้ใน 3 ประเด็นดังนี้

1. ไม่มีการนำคำตอบจากแบบประเมินและวิเคราะห์ความเครียดมาเก็บรวบรวมไว้อย่างเป็นรูปธรรม
2. ขาดการติดตามผลหลังจากผู้เข้ารับการประเมินทำแบบประเมินเรียบร้อยแล้ว ส่งผลให้ไม่สามารถทราบได้ว่าผู้เข้ารับการประเมินมีภาวะความเครียดเป็นอย่างไรเมื่อเวลาผ่านไป
3. แบบประเมินและวิเคราะห์ความเครียดของกรมสุขภาพจิตที่ใช้อยู่นี้ เป็นเพียงการวิเคราะห์สภาวะของผู้รับการประเมินในปัจจุบัน กล่าวคือ ไม่สามารถทำนายหรือคาดการณ์โอกาสในการเกิดความเครียดในอนาคตอันใกล้ได้



บทที่ 3

ระเบียบวิธีการวิจัย

การศึกษาค้นคว้าและวิจัยในครั้งนี้ ผู้วิจัยที่ได้นำแนวคิด ทฤษฎีที่เกี่ยวข้องและข้อมูลพื้นฐานเกี่ยวกับความเครียดและการทำเหมืองข้อมูล มาใช้กำหนดกรอบแนวคิดในการประเมินความเครียดและแบบจำลองเพื่อพยากรณ์ความเครียดของนักศึกษา โดยมีขั้นตอนดังต่อไปนี้

1. ประชากรและกลุ่มตัวอย่าง

ประชากร

ประชากรที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาหลักสูตรบริหารธุรกิจบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจ ชั้นปีที่ 1-4 คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ ปีการศึกษา 2557 จำนวน 400 คน

กลุ่มตัวอย่าง

กลุ่มตัวอย่างที่ใช้ในการศึกษาค้นคว้าครั้งนี้ เป็นนักศึกษาหลักสูตรบริหารธุรกิจบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจ ชั้นปีที่ 1-4 คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ ปีการศึกษา 2557 จำนวน 300 คน ซึ่งได้มาโดยการกำหนดกลุ่มตัวอย่าง โดยมีขั้นตอนในการเลือกกลุ่มตัวอย่างดังนี้

1. กำหนดขนาดของกลุ่มตัวอย่างว่าจากประชากร จำนวน 400 คน โดยใช้การกำหนดกลุ่มตัวอย่างแบบเจาะจงในการศึกษาค้นคว้าครั้งนี้ ผู้วิจัยจะทำการศึกษาจากนักศึกษาหลักสูตรบริหารธุรกิจบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจ ชั้นปีที่ 1-4 คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ ปีการศึกษา 2557 จำนวน 300 คน
2. นักศึกษาหลักสูตรบริหารธุรกิจบัณฑิต สาขาเทคโนโลยีสารสนเทศทางธุรกิจ ชั้นปีที่ 1-4 คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยีราชมงคลรัตนโกสินทร์ จำแนกตามชั้นปีศึกษา จากนั้นทำการกำหนดกลุ่มตัวอย่างในแต่ละชั้นปี โดยกำหนดจำนวน ตามตาราง 3
3. ทำการเลือกกลุ่มตัวอย่างจากนักศึกษาแต่ละชั้นปี โดยใช้วิธีการสุ่มแบบเจาะจง

ตารางที่ 3 แสดงจำนวนประชากรและกลุ่มตัวอย่างจำแนกตามชั้นปี

ชั้นปี	จำนวนประชากร	จำนวนกลุ่มตัวอย่าง
ชั้นปีที่ 1	116	108
ชั้นปีที่ 2	95	75
ชั้นปีที่ 3	87	72
ชั้นปีที่ 4	62	45
รวม	360	300

2. เครื่องมือที่ใช้ในการวิจัย

ผู้วิจัยได้ทำการค้นคว้าตำรา เอกสารต่างๆ ทฤษฎีและงานวิจัยที่เกี่ยวข้องเพื่อศึกษาว่าการประเมินความเครียดของแต่ละบุคคลมีปัจจัยอะไรบ้างในการตรวจสอบ จากการศึกษาพบว่ามีแบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง ที่จัดทำโดยกรมสุขภาพจิต กระทรวงสาธารณสุข ซึ่งเป็นแบบประเมินที่ได้รับการยอมรับและมีการนำมาใช้ประเมินความเครียดกันอย่างแพร่หลายตั้งแต่ปี 2541 โดยแบบประเมินดังกล่าวมีข้อคำถามจำนวน 20 ข้อ ที่เกี่ยวข้องกับอาการ พฤติกรรม หรือความรู้สึกต่างๆ ของผู้รับการประเมิน ดังนั้น ผู้วิจัยจึงใช้แบบประเมินนี้เป็นตัวหลักในการเก็บข้อมูลจากกลุ่มตัวอย่างในงานวิจัยนี้

แบบประเมินที่ใช้นี้จะแบ่งออกเป็น 2 ส่วน คือ

ส่วนที่ 1 เป็นแบบสอบถามเกี่ยวกับข้อมูลทั่วไป ได้แก่ รหัสนักศึกษา, ชื่อ-สกุล, เพศ และชั้นปี

ส่วนที่ 2 เป็นแบบสอบถามเกี่ยวกับอาการ พฤติกรรมหรือความรู้สึกของผู้รับการประเมิน ในช่วง 2 เดือนที่ผ่านมา ลักษณะข้อคำถามเป็นมาตราส่วนประมาณค่า (Rating scale) มี 4 ระดับ จำนวน 20 ข้อ มีเกณฑ์การให้คะแนนของระดับอาการ ดังนี้

ไม่เคยเลย	0	คะแนน
เป็นครั้งคราว	1	คะแนน
เป็นบ่อย	2	คะแนน
เป็นประจำ	3	คะแนน

การแปลความหมายของคะแนนเพื่อจัดระดับความเครียดเป็น 5 ระดับ โดยใช้คะแนนรวม และใช้การประเมินแบบอิงเกณฑ์ (Criterion Reference) ดังนี้

0-5	คะแนน	หมายถึง	มีความเครียดต่ำกว่าเกณฑ์ปกติ / ตอบไม่ตรงความจริง / เข้าใจคำถามคลาดเคลื่อน
6-17	คะแนน	หมายถึง	มีความเครียดอยู่ในเกณฑ์ปกติ / ไม่เครียด
18-25	คะแนน	หมายถึง	มีความเครียดอยู่ในระดับสูงกว่าปกติเล็กน้อย
26-29	คะแนน	หมายถึง	มีความเครียดอยู่ในระดับสูงกว่าปกติปานกลาง
30-60	คะแนน	หมายถึง	มีความเครียดอยู่ในระดับสูงกว่าปกติมาก

3. การเก็บรวบรวมข้อมูล

งานวิจัยนี้ได้ นำแบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง ที่จัดทำโดยกรมสุขภาพจิต กระทรวงสาธารณสุข มาใช้ในการเก็บข้อมูลจากกลุ่มตัวอย่างที่กำหนดไว้ เนื่องจากแบบประเมินดังกล่าวเป็นคำถามที่ให้ผู้รับการประเมินทำการตรวจสอบตนเองว่ามีอาการทางร่างกาย พฤติกรรม หรือความรู้สึกอย่างไรภายในระยะเวลา 2 เดือนที่ผ่านมา ดังนั้น การเก็บข้อมูลในแต่ละครั้งจะมี

ระยะเวลาห่างกัน 2 เดือน โดยมีเงื่อนไขว่า กลุ่มตัวอย่างทั้ง 4 รอบ จะต้องเป็นนักศึกษาในกลุ่มเดิมทั้งหมด

จากที่กล่าวไว้ว่า ผู้รับการประเมิน 1 คนต้องตอบแบบสอบถามจำนวน 4 รอบ เมื่อผู้รับการประเมินทำแบบสอบถามครบทั้ง 4 รอบแล้ว ข้อมูลที่ได้จากผู้รับการประเมิน 1 คน จะถูกนำมาเข้ารหัสเพื่อเก็บลงในโปรแกรม Excel โดยเก็บข้อมูล 1 บรรทัด (เรคคอร์ด) ต่อ 1 ผู้รับการประเมิน ภายในบรรทัดจะประกอบด้วยแททริบิวต์จำนวน 84 แททริบิวต์ นั่นคือ ข้อมูลพื้นฐานจำนวน 4 แททริบิวต์ ได้แก่ ลำดับที่, รหัสนักศึกษา, ชื่อ-สกุล, เพศ และ ข้อคำถาม 80 แททริบิวต์ (จากการเก็บข้อมูลจำนวน 4 ครั้ง ครั้งละ 20 ข้อ) ภาพที่ 6 แสดงตัวอย่างของข้อมูลเบื้องต้นที่ถูกเก็บอยู่ใน Excel

ลำดับ	รหัส	ชื่อ	เพศ	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	ข้อ6	ข้อ7	ข้อ8	ข้อ9	ข้อ10	ข้อ11	ข้อ12	ข้อ13	ข้อ14	ข้อ15	ข้อ16	ข้อ17	ข้อ18	ข้อ19	ข้อ20	ข้อ21	ข้อ22	ข้อ23	ข้อ24	ข้อ25	ข้อ26	ข้อ27	ข้อ28	ข้อ29	ข้อ30					
1	1####	นายพิชชา	ชาย	1	1	1	1	0	2	0	1	1	1	1	2	1	0	1	1	0	1	1	0	1	1	1	2	2	2	2	2	2						
2	2####	นางสาวรินยา	หญิง	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	2	1	3	1	0	1	1	0	1	1	0	0	1	0	1					
3	3####	นายวิศรุต ธี	ชาย	1	0	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	0	1	3	1	1	0	1	1	0	1	0	1				
4	4####	นางสาวธยาณี	หญิง	0	2	1	2	1	0	1	1	1	1	1	1	1	0	0	1	0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	0			
5	5####	นายขจรศักดิ์	ชาย	1	1	1	2	0	1	1	1	0	1	2	2	2	1	0	1	2	3	2	1	1	1	0	1	0	1	0	0	0	0	0	0			
6	6####	นายเจษฎา ติ	ชาย	2	1	0	1	0	3	0	0	0	1	1	1	0	1	0	2	1	2	0	2	0	0	3	1	1	1	3	2	2	2	2	2	2		
7	7####	นายอภิรักษ์	ชาย	1	1	0	1	0	2	1	0	0	0	1	1	2	0	0	2	2	1	1	0	2	1	0	2	1	0	0	0	2	1	0	0	0		
8	8####	นายสนนาร	ชาย	0	1	0	1	0	1	0	0	0	0	1	1	1	0	1	1	1	1	1	0	0	1	0	1	0	1	1	0	1	1	1	0	0		
9	9####	นางสาวอรุณ	หญิง	1	1	1	1	1	3	1	0	0	1	2	1	3	2	1	2	0	2	3	0	1	1	0	1	0	1	0	2	0	0	0	0	1		
10	10####	นายอริญ คช	ชาย	1	3	0	3	3	2	3	0	0	1	2	2	1	0	3	0	0	0	0	1	2	0	2	1	2	1	0	0	1	0	0	1	0		
11	11####	นายอมเรศ	ชาย	0	1	0	1	0	1	0	0	0	0	2	1	1	1	1	2	2	1	1	1	1	0	1	0	1	0	2	0	0	0	0	1	0		
12	12####	นางสาวสุจินทา	หญิง	1	2	0	0	0	1	0	0	0	0	1	1	1	0	1	1	3	2	1	0	0	3	1	1	1	1	0	0	1	0	0	1	0		
13	13####	นายชนสิทธิ์	ชาย	1	0	0	1	0	1	1	1	0	1	1	1	0	1	3	1	2	1	0	2	3	0	2	2	0	1	1	0	1	1	0	1	0	0	
14	14####	นายศกร	ชาย	2	1	1	3	1	0	0	1	0	3	2	1	1	1	0	2	0	2	1	0	0	1	0	1	0	1	1	1	1	1	0	1	1	0	
15	15####	นายธเรศวร	ชาย	0	3	1	2	1	1	0	0	0	1	3	1	2	1	1	1	3	1	2	0	2	2	0	1	0	2	1	0	0	1	0	0	1	0	
16	16####	นางสาว	หญิง	1	1	1	1	1	3	1	0	0	1	1	1	2	0	0	0	1	1	0	0	1	0	2	2	2	1	2	1	0	0	1	1	0	1	
17	17####	นางสาว	หญิง	1	1	0	1	1	0	1	0	0	1	2	0	1	1	0	0	0	1	1	0	0	1	1	0	2	2	1	1	0	2	1	0	0	1	
18	18####	นายธเรศวร	ชาย	1	1	0	1	1	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	19####	นายยุทธนา	ชาย	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	2	1	0	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2
20	20####	นายศศิกรินทร์	ชาย	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	21####	นายพิระพงษ์	ชาย	2	2	2	1	3	1	1	1	0	1	1	2	2	1	1	1	1	2	2	1	0	1	2	3	2	1	0	1	2	3	2	3	2	3	
22	22####	นายปรกติ	ชาย	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	
23	23####	นางสาว	หญิง	2	2	1	2	1	1	1	1	0	2	3	1	1	1	1	2	1	1	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
24	24####	นายสราวุฒิ	ชาย	1	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	0	0	0	

ภาพที่ 6 ข้อมูลที่นำมาใช้ในการศึกษาซึ่งอยู่ในรูปแบบ Excel

4. การเตรียมข้อมูล (Pre-processing) และวิเคราะห์ข้อมูล

เนื่องจากข้อมูลที่ถูกจัดเก็บไว้ยังไม่อยู่ในรูปแบบที่สามารถนำไปใช้ในการสร้างและทดสอบแบบจำลองด้วยเครื่องมือการทำเหมืองข้อมูล จึงต้องมีการเตรียมข้อมูลในเบื้องต้นก่อนนำไปใช้เพื่อให้ได้ผลการทดลองที่มีประสิทธิภาพสูงสุด ในขั้นตอนนี้จะประกอบด้วยขั้นตอนย่อยตามลำดับคือ 1) การตัดข้อมูล 2) การแก้ไขข้อมูลให้ถูกต้อง 3) การเพิ่มข้อมูล

4.1 การตัดข้อมูล

เนื่องจากข้อมูลมีบางแททริบิวต์ที่ไม่เกี่ยวข้องกับกระบวนการวิเคราะห์และประเมินความเครียดอยู่ จึงจำเป็นต้องมีการตัดแททริบิวต์บางตัวออกไป ได้แก่ ลำดับที่, รหัสนักศึกษา, ชื่อ-สกุล นอกจากนี้ยังมีกลุ่มตัวอย่างบางคนที่ทำแบบประเมินไม่ครบทั้ง 4 ครั้ง ดังนั้นจึงทำการตัดข้อมูลบางเรคคอร์ดออกไป ดังแสดงในภาพที่ 7 เมื่อตัดแททริบิวต์เหล่านี้ออกไป ทำให้คงเหลือแททริบิวต์จำนวน 81 แททริบิวต์

4.2 การแก้ไขข้อมูลให้ถูกต้อง

โดยทั่วไปข้อมูลที่ถูกรวบรวมอาจมีความผิดปกติต่างๆ ได้ เช่น ข้อมูลบางแอทริบิวต์ขาดหายไป (Missing Value) ขาดแอทริบิวต์ที่น่าสนใจหรือขาดรายละเอียดข้อมูล เป็นข้อมูลรบกวน (Noisy Data) เช่น ข้อมูลมีค่าผิดพลาด (error) หรือมีค่าผิดปกติ (Outliers) ซึ่งอาจเกิดจากการตอบแบบสอบถามที่ไม่ถูกต้อง การบันทึกข้อมูลลงคอมพิวเตอร์ผิดพลาดของเจ้าหน้าที่ เป็นต้น

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI				
1	ลำดับ	รหัส	ชื่อ	เพศ	ครั้งที่ 1															ครั้งที่ 2																			
2	1	###	นายพิชิต	ชาย	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	ข้อ6	ข้อ7	ข้อ8	ข้อ9	ข้อ10	ข้อ11	ข้อ12	ข้อ13	ข้อ14	ข้อ15	ข้อ16	ข้อ17	ข้อ18	ข้อ19	ข้อ20	ข้อ1	ข้อ2	ข้อ3	ข้อ4	ข้อ5	ข้อ6	ข้อ7	ข้อ8	ข้อ9	ข้อ10					
3	2	###	นางสาววิภา	หญิง	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	0	1	1	0	1	2	1	0	1	1	1	1	2	2				
4	3	###	นายวิศรุต	ชาย	1	0	1	1	0	1	1	0	0	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	3	1	1	0	1	1	0	1		
5	4	###	นางสาวอรุณี	หญิง	0	2	1	2	1	0	1	1	1	1	1	1	1	0	0	1	0	1	0	0	1	0	1	0	1	1	0	1	1	0	1	1	0	0	0
6	5	###	นายจตุศักดิ์	ชาย	1	1	1	2	0	1	1	1	0	1	2	2	2	1	0	1	2	3	2	1	1	1	1	0	1	0	1	0	1	0	0	0	0		
7	6	###	นายเจษฎา	ชาย	2	1	0	1	0	3	0	0	0	0	1	1	1	0	1	0	2	1	2	0	2	0	2	0	0	3	1	1	1	1	3	2			
8	7	###	นายอภิรักษ์	ชาย	1	1	0	1	0	2	1	0	0	0	1	1	2	0	0	2	2	1	1	0	2	1	1	0	2	1	0	0	0	2	1	0	0	0	
9	8	###	นายมาธ	ชาย	0	1	0	1	0	1	0	0	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	1	1	0	1	1	0
10	9	###	นางสาวอรุณ	หญิง	1	1	1	1	1	3	1	0	0	1	2	1	3	2	1	2	0	2	3	0	1	1	0	1	1	0	1	0	2	0	0	0	1	1	
11	10	###	นายอริย ุคช	ชาย	1	3	0	3	3	2	3	0	0	1	2	2	2	1	0	3	0	0	0	0	1	2	2	0	1	2	0	2	1	2	1	0	0	1	
12	11	###	นายอมเรศ	ชาย	0	1	0	1	0	1	0	0	0	0	2	1	1	1	1	2	2	2	1	1	1	1	1	1	0	1	0	2	0	0	0	1	0	1	
13	12	###	นางสาวศุจินา	หญิง	1	2	0	0	1	0	0	0	0	1	1	1	0	1	1	3	2	1	0	0	3	1	1	1	1	1	1	1	0	1	0	1	0	1	
14	13	###	นายชนสิทธิ์	ชาย	1	0	0	1	0	1	1	1	1	0	1	1	1	0	1	3	1	2	1	0	2	3	0	2	3	0	2	2	0	1	1	0	1	1	
15	14	###	นายศุภกร	ชาย	2	1	1	3	1	0	0	1	0	3	2	1	1	1	0	2	0	2	1	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1	
16	15	###	นายอภิสิทธิ์	ชาย	0	3	1	2	1	1	0	0	0	1	3	1	2	1	1	1	3	1	2	0	2	2	0	1	0	2	1	0	2	1	0	0	1	0	
17	16	###	นางสาว	หญิง	1	1	1	1	3	1	0	0	0	1	1	1	2	0	0	1	1	0	2	2	0	0	1	1	0	2	2	1	2	1	0	0	1	1	
18	17	###	นางสาว	หญิง	1	1	0	1	1	0	1	0	0	1	2	0	1	1	0	0	1	1	0	2	2	1	1	0	2	1	1	0	2	1	0	0	1	0	
19	18	###	นายอริย ุคช	ชาย																																			
20	19	###	นายอนุชา	ชาย	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	2	1	0	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
21	20	###	นายศิริพันธ์	ชาย	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	21	###	นายพิรพัฒน์	ชาย	2	2	2	1	3	1	1	0	1	1	2	2	1	1	1	1	2	2	1	0	1	2	2	1	0	1	2	3	2	1	0	1	2	3	
23	22	###	นายปรกต์	ชาย	0	1	0	1	0	1	1	1	1	1	1	0	0	0	2	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
24	23	###	นางสาว	หญิง	2	2	1	2	1	1	1	1	0	2	3	1	1	1	2	1	1	2	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	
25	24	###	นายสุภาณี	ชาย	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	0	0	0	0	

ภาพที่ 7 การตัดแอทริบิวต์และเรคอร์ดที่ไม่ใช้ในการวิเคราะห์และประเมินความเครียด

ภาพที่ 8 การแก้ไขข้อมูลบางแอทริบิวต์ที่ขาดหายไป

จากมากไปน้อยในแต่ละแอทริบิวต์เพื่อหลีกเลี่ยงข้อมูลที่มีค่าผิดพลาดใดๆ พบว่ามีแอทริบิวต์ที่ขาดหายไปจำนวน 9 ค่า จึงแก้ไขโดยการนำค่าจากคำตอบข้อเดียวกันของครั้งที่แล้วของเรคอร์ดเดียวกันมาเติมให้ ตัวอย่างดังภาพที่ 8 เช่น ในการเก็บแบบประเมินครั้งที่ 2 ข้อมูลในข้อ 3 ของเรคอร์ดที่ 16 ได้หายไป จึงได้นำคำตอบข้อ 3 ของเรคอร์ดเดียวกันในแบบประเมินครั้งที่ 1 มาเติมให้

4.3 การเพิ่มข้อมูล

ในขั้นตอนการเพิ่มข้อมูลนี้ การตอบแบบประเมินแต่ละครั้งใน Excel จะมีการเพิ่มแอทริบิวต์จำนวน 2 แอทริบิวต์ ดังภาพที่ 9 คือ คะแนนรวม และระดับความเครียด ดังนั้น สำหรับผู้รับการประเมิน 1 คน (1 เรคอร์ด) จะมีแอทริบิวต์เพิ่มเท่ากับ 2 แอทริบิวต์ X 4 ครั้ง เมื่อเสร็จขั้นตอนนี้ทำให้มีแอทริบิวต์จำนวน 89 แอทริบิวต์ สำหรับแอทริบิวต์ระดับความเครียดจะมีการแทนค่าที่เป็นไปได้ตั้งแต่ 0-4 ดังต่อไปนี้

- คะแนนรวม 0-5 คะแนน แทนค่าด้วย 0
- คะแนนรวม 6-17 คะแนน แทนค่าด้วย 1
- คะแนนรวม 18-25 คะแนน แทนค่าด้วย 2
- คะแนนรวม 26-29 คะแนน แทนค่าด้วย 3
- คะแนนรวม 30-60 คะแนน แทนค่าด้วย 4

ภาพที่ 9 การเพิ่มข้อมูลคะแนนรวมและระดับความเครียด

ก่อนจะทำการสร้างและทดสอบแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูล ผู้วิจัยนำข้อมูลที่เก็บได้มาทำการวิเคราะห์โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป โดยดำเนินการวิเคราะห์ข้อมูลและนำเสนอผลการวิเคราะห์ตามลำดับดังต่อไปนี้

- 1) วิเคราะห์ข้อมูล เพศ ชั้นปี โดยการแจกแจงความถี่และค่าร้อยละ
- 2) วิเคราะห์ข้อมูลอาการ พฤติกรรมหรือความรู้สึกจากข้อคำตอบจำนวน 20 ข้อ จำนวน 4 รอบ โดยการแจกแจงความถี่และค่าร้อยละ
- 3) วิเคราะห์คะแนนรวมของข้อคำตอบจำนวน 20 ข้อ จำนวน 4 รอบ โดยการแจกแจงความถี่และค่าร้อยละ
- 4) วิเคราะห์ระดับความเครียด 5 ระดับ จำนวน 4 รอบ โดยการแจกแจงความถี่และค่าร้อยละ

ข้อมูลที่ใช้ในการทดลองทั้งหมดสามารถวิเคราะห์แยกตามแต่ละข้อได้ดังรายละเอียดในตารางที่ 4 - 8 และภาพที่ 10 - 14

ตารางที่ 4 ผลการวิเคราะห์ค่าของข้อมูลเพศและชั้นปี

ชั้นปี	เพศ	จำนวน (เรคอร์ด)	ร้อยละ
ชั้นปีที่ 1	ชาย	77	25.67
	หญิง	31	10.33
ชั้นปีที่ 2	ชาย	51	17.00

	หญิง	24	8.00
ชั้นปีที่ 3	ชาย	46	15.33
	หญิง	26	8.67
ชั้นปีที่ 4	ชาย	38	12.67
	หญิง	7	2.33
รวม	ชาย	212	70.67
	หญิง	88	29.33











ภาพที่ 10 การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 1



ภาพที่ 11 การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 2

ภาพที่ 12 การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 3



ภาพที่ 13 การวิเคราะห์คะแนนรวมของคำตอบจำนวน 20 ข้อ รอบที่ 4

ภาพที่ 14 การวิเคราะห์ระดับความเครียด 5 ระดับ จำนวน 4 รอบ

จากภาพที่ 14 แสดงให้เห็นว่าผลการแจกแจงความถี่ทั้ง 4 รอบสอดคล้องไปในทิศทางเดียวกันนั่นคือ นักศึกษาส่วนใหญ่ไม่เครียด รองลงมาคือเครียดสูงกว่าเกณฑ์ปกติเล็กน้อย

5. การสร้างแบบจำลองเพื่อพยากรณ์ความเครียด

ในการศึกษาและสร้างแบบจำลองเพื่อพยากรณ์นี้ใช้เทคนิคการจำแนกกลุ่มซึ่งเป็นเทคนิคหนึ่งของการทำเหมืองข้อมูล โดยเทคนิคนี้จะให้ผลลัพธ์เป็นกฎเพื่อให้สามารถทำความเข้าใจได้ง่ายและสามารถนำไปใช้พยากรณ์ความเครียดได้ ในที่นี้ ผู้วิจัยได้เลือกเทคนิคที่ได้รับความนิยมสำหรับการจำแนกกลุ่ม 3 เทคนิคเช่นเดียวกับงานวิจัยของซัชชฎา และคณะ (2556) นั่นคือ เทคนิคการเรียนรู้แบบเบย์ (Bayesian Learning) เทคนิคต้นไม้ตัดสินใจ (Decision Tree: 4.5) และเทคนิคโครงข่ายประสาทเทียม (Neural Network) สำหรับเทคนิคการเรียนรู้แบบเบย์จะแบ่งออกเป็น 2 อัลกอริทึมคือ Bayesian Network และ Naïve Bayesian พร้อมกันนี้ยังมีการนำอัลกอริทึม Decision Table และ อัลกอริทึม Partial Rules (PART) เข้ามาทดลองเปรียบเทียบกับ Decision Tree:4.5 เช่นเดียวกับงานวิจัยของเสถียร (2553) ด้วย ส่วนเทคนิคโครงข่ายประสาทเทียมนั้นเลือกใช้อัลกอริทึม Multilayer Perceptron (MLP) เช่นเดียวกับงานวิจัยของซัชชฎา และคณะ (2556) โดยสรุปแล้วอัลกอริทึมในการจำแนกกลุ่มที่ถูกลำเอียงมาใช้สร้างแบบจำลองพยากรณ์ความเครียดในงานวิจัยนี้มีจำนวน 6 อัลกอริทึม คือ

1. Bayesian Network (Kantardzic M., 2003) หรือเรียกย่อว่า Bayesnet เป็นวิธีการลดข้อจำกัดของ Naïve Bayesian ในสมมุติฐานของความไม่ขึ้นต่อกันระหว่างแอทริบิวต์ในวิธีของ Naïve Bayesian แต่ในความเป็นจริงแล้วแอทริบิวต์บางตัวจะขึ้นต่อกัน และควรที่จะนำความขึ้นต่อกันนี้เข้ามาใส่ไว้ในแบบจำลองด้วย จึงใช้ Bayesnet ในการอธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไขระหว่างตัวแปรเพื่อให้กระบวนการเรียนรู้มีประสิทธิภาพ
2. Naïve Bayesian (Kantardzic M., 2003) ทำนายผลโดยหลักการจำแนกประเภทโดยทฤษฎีของเบย์ (Bayes Theorem) การเรียนรู้เบย์อย่างง่าย(Naïve Bayesian) เป็นวิธีจำแนกประเภทข้อมูลที่มีประสิทธิภาพวิธีหนึ่งเหมาะกับกรณีของเซตตัวอย่างมีจำนวน

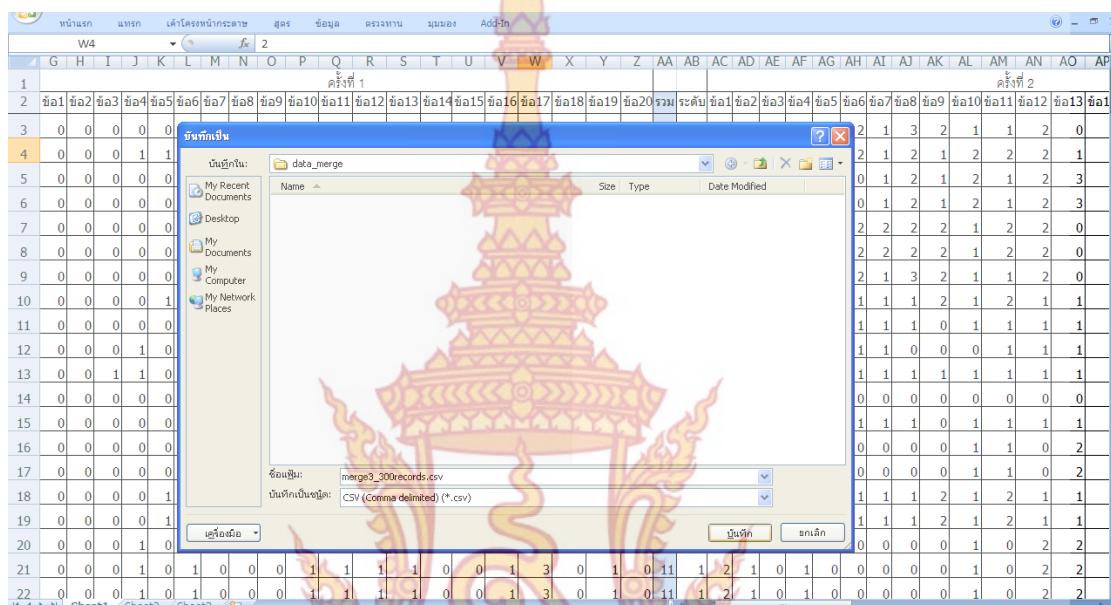
มากและแอทริบิวต์ของตัวอย่างไม่ขึ้นต่อกัน มีการจำแนกประเภทเบี่ยงอย่างง่ายไปประยุกต์ใช้ในงานด้านการวินิจฉัย (Diagnosis) และพบว่าให้ประสิทธิภาพที่ดี

3. Decision Tree: C4.5 (Kantardzic M., 2003) หรือในโปรแกรม Weka ใช้ชื่อว่า Classifiers trees J48 เป็นการจัดกลุ่มโดยผลลัพธ์ที่ได้อยู่ในรูปของต้นไม้การตัดสินใจที่มีกิ่งต้นไม้จำนวนมาก โดยค่าของข้อมูลของแอทริบิวต์ที่มีค่าต่อเนื่องจะใช้เครื่องหมาย "<", "<=", ">=", ">" ในการพิจารณา ส่วนข้อมูลที่เป็นค่าไม่ต่อเนื่องจะใช้เครื่องหมาย "=" และใช้เงื่อนไข "AND" ในการเชื่อมโยงระหว่างแอทริบิวต์ จะเชื่อมโยงไปจนถึงแอทริบิวต์สุดท้ายที่จะแบ่งกลุ่มได้
4. Decision Table (Kantardzic M., 2003) หรือในโปรแกรม Weka ใช้ชื่อว่า Classifiers Decision Table เป็นการจัดกลุ่มในลักษณะตารางการตัดสินใจโดยการนำผลที่ได้จากการเรียนรู้จากข้อมูลที่น่าเข้ามาที่มีลักษณะของข้อมูลที่ใกล้เคียงกันให้มาอยู่ในกลุ่มเดียวกัน โดยแอทริบิวต์สุดท้ายของตารางจะใช้กำหนดกลุ่มที่ต้องการแบ่ง
5. Decision Rules: Partial Rules (Kantardzic M., 2003) หรือในโปรแกรม Weka ใช้ชื่อว่า Classifiers PART เป็นการจัดกลุ่มโดยผลลัพธ์ที่ได้จะอยู่ในรูปของกฎที่น่าเอาข้อมูลแต่ละแอทริบิวต์มาเป็นเงื่อนไขในการตัดสินใจลักษณะของข้อมูลที่ใกล้เคียงกันให้มาอยู่ในกลุ่มเดียวกันโดยค่าของข้อมูลของแอทริบิวต์ที่มีค่าต่อเนื่องจะใช้เครื่องหมาย "<", "<=", ">=", ">" ในการพิจารณา ส่วนข้อมูลที่เป็นค่าไม่ต่อเนื่องจะใช้เครื่องหมาย "=" และใช้เงื่อนไข "AND" ในการเชื่อมโยงระหว่างแอทริบิวต์ จะเชื่อมโยงไปจนถึงแอทริบิวต์สุดท้ายที่จะแบ่งกลุ่มได้
6. Multilayer Perceptron เป็นอัลกอริทึมรูปแบบหนึ่งของกรรมวิธี Neuron ที่ต้องการคาดการณ์ข้อมูลที่เป็นลักษณะเชิงปริมาณ โดยการทำงานเป็นการประมวลผลโดยใช้การคำนวณการให้น้ำหนักของการเชื่อมต่อในแต่ละโหนด ลักษณะการเชื่อมต่อเป็นแบบเครือข่ายและมีการปรับค่าถ่วงน้ำหนักของการส่งผ่านข้อมูลไปในแต่ละโหนด ซึ่งจะมีการส่งค่าย้อนกลับเพื่อปรับค่าน้ำหนักจนได้ค่าผลลัพธ์ที่ยอมรับได้ ในการสร้างเครือข่ายจะมีระดับโหนดข้อมูลเข้า (input node) และระดับโหนดซ่อน (hidden node) จำนวนของชั้นโหนดซ่อนอาจมีหลายชั้นก็ได้

ข้อมูลที่ใช้ในการสร้างและทดสอบแบบจำลองมีจำนวนทั้งสิ้น 300 เรคอร์ด สำหรับการทดสอบได้ใช้วิธี Cross-validation test เนื่องจากเป็นวิธีที่นิยมใช้กันมากที่สุด (เอกสิทธิ์, 2557) เมื่อเปรียบเทียบกับวิธี Split test หรือวิธีสร้าง Testing set ขึ้นมาเอง นอกจากนี้วิธี Cross-validation test ยังเหมาะสมกับข้อมูลจำนวนไม่มากนัก (เช่น 300 เรคอร์ด) ทำให้สามารถทดสอบซ้ำหลายรอบได้และส่งผลให้การทดสอบโดยวิธี Cross-validation test มีความน่าเชื่อถือกว่าวิธีการทดสอบแบบอื่น (เอกสิทธิ์, 2557) ดังนั้น งานวิจัยนี้จึงใช้ 10-folds cross-validation แบ่งกลุ่มข้อมูลออกเป็น 10 กลุ่ม ในแต่ละรอบจะนำกลุ่มข้อมูลจำนวน 9 กลุ่มเป็นกลุ่มศึกษา (Training Set) และกลุ่มข้อมูลที่เหลือเป็นกลุ่มทดสอบ (Testing Set) โดยจะทำซ้ำจำนวน 10 รอบ เพื่อเปลี่ยนกลุ่มทดสอบให้ครบทุกกลุ่ม และนำผลลัพธ์ที่ได้มาหาค่าเฉลี่ย โดยเกณฑ์ที่ใช้ในการพิจารณาคุณภาพของแบบจำลองที่ได้จาก

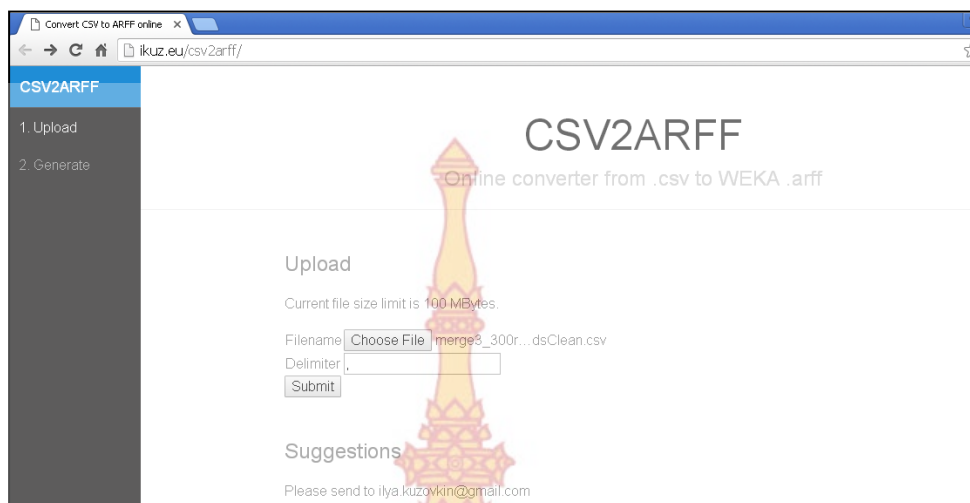
แต่ละอัลกอริทึม จะพิจารณาจากค่าความถูกต้อง (Accuracy), ค่าความแม่นยำ (Precision), ค่าความระลึก (Recall) และค่าความเหวี่ยง (F-measure) ที่มากที่สุด

การนำเข้าข้อมูลของโปรแกรม Weka 3.7.3 ต้องทำการแปลงไฟล์ข้อมูลจาก .xls ให้เป็นนามสกุล csv ก่อนดังภาพที่ 15 จึงจะสามารถนำข้อมูลเข้าในโปรแกรม Weka ได้

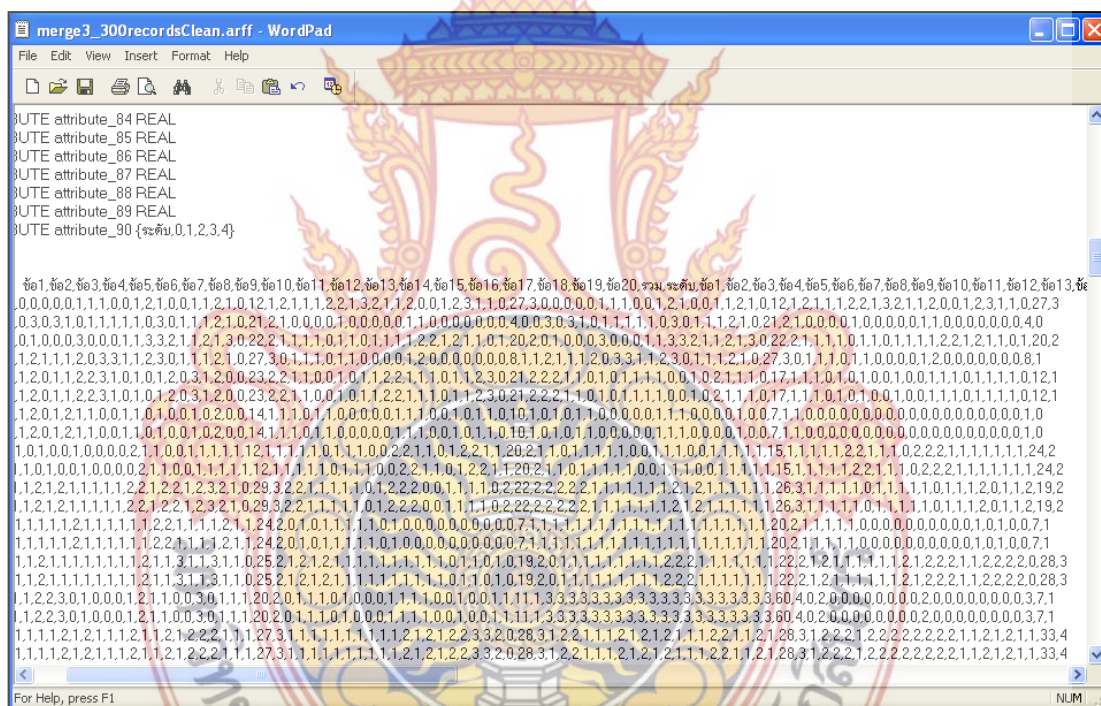


ภาพที่ 15 การบันทึกข้อมูลในรูปแบบนามสกุล *.csv

ขั้นตอนต่อมาคือการแปลงไฟล์ .csv ให้เป็นไฟล์ .arff เพื่อให้เป็นรูปแบบมาตรฐานสำหรับใช้ในโปรแกรม Weka โดยเข้าไปแปลงไฟล์ได้ที่ <http://ikuz.eu/csv2arff/> ดังภาพที่ 16
 แฟ้มข้อมูล .arff จะสามารถแจกแจงข้อมูลในแต่ละแอทริบิวต์ให้เป็นค่าไม่ต่อเนื่องได้ ซึ่งโดยปกติเมื่อทำการบันทึก เป็น .arff แล้ว ค่าปริยายของแต่ละแอทริบิวต์ที่เป็นตัวเลขจะถูกกำหนดให้เป็น numeric สำหรับการแก้ไขข้อมูลในไฟล์ .arff สามารถกระทำได้โดยเปิดแฟ้มข้อมูลด้วยโปรแกรม WordPad ดังภาพที่ 17



ภาพที่ 16 เว็บไซต์ในการแปลงไฟล์ *.csv เป็น *.arff

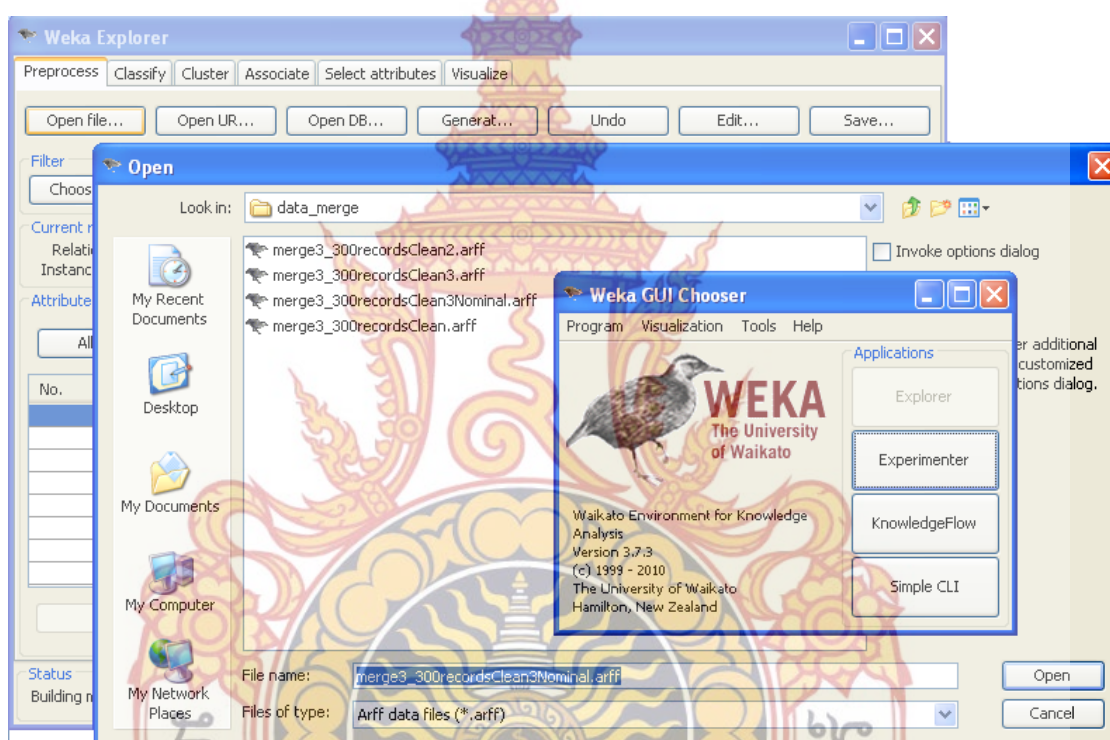


ภาพที่ 17 การใช้โปรแกรม WordPad เปิดไฟล์ *.arff ขึ้นมาแก้ไข

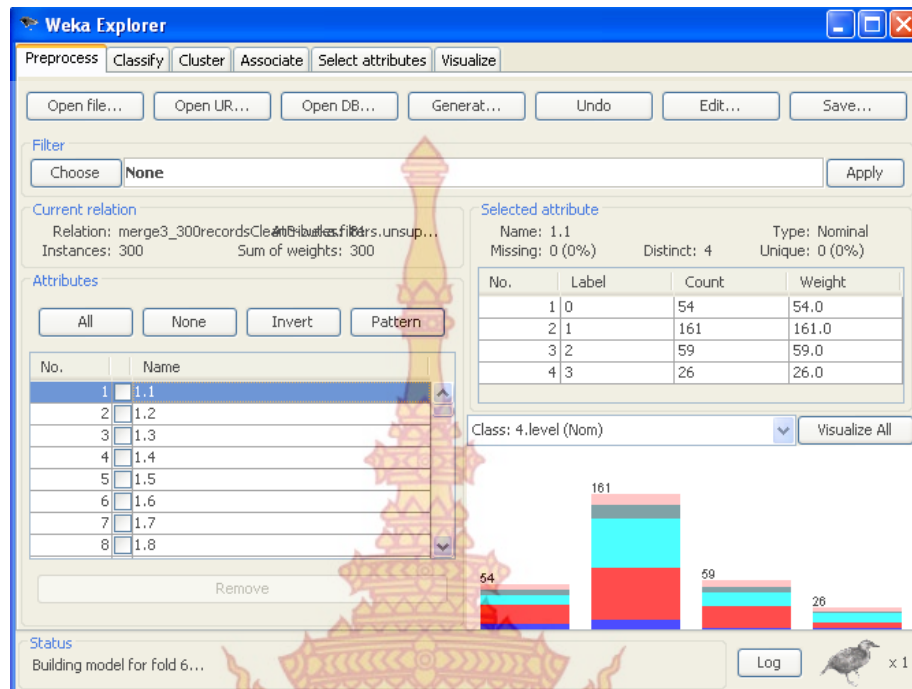
โปรแกรม Weka เลือก Explorer >> Preprocess>> Open file... เลือกไฟล์ข้อมูลที่ต้องการ ดังภาพที่ 18 สำหรับตัวอย่างการอ่านค่าของหน้าจอแรกนี้แสดงไว้ดังภาพที่ 19 นั่นคือ ไฟล์ชื่อ merge3_300recordsClean3 มีทั้งหมด 300 เรคคอร์ด (Instances) แต่ละเรคคอร์ดมี 89 แอททริบิวต์ แอททริบิวต์ที่ถูกเลือกในขณะนี้คือ 1.1 (คำถามข้อที่ 1 ของการเก็บข้อมูลรอบที่ 1) มีค่าที่เป็นไปได้ทั้งหมด 4 ค่าคือ 0, 1, 2, 3 โดยที่ จำนวนเรคคอร์ดที่มีค่า 0 มีจำนวน 54 เรคคอร์ด (18.00%) จำนวนเรคคอร์ดที่มีค่า 1 มีจำนวน 161 เรคคอร์ด (53.67%) จำนวนเรคคอร์ดที่มีค่า 2 มีจำนวน 59 เรคคอร์ด

(19.67%) และจำนวนเรคคอร์ดที่มีค่า 3 มีจำนวน 26 เรคคอร์ด (8.67%) ซึ่งสอดคล้องกับข้อมูลในตารางที่ 5

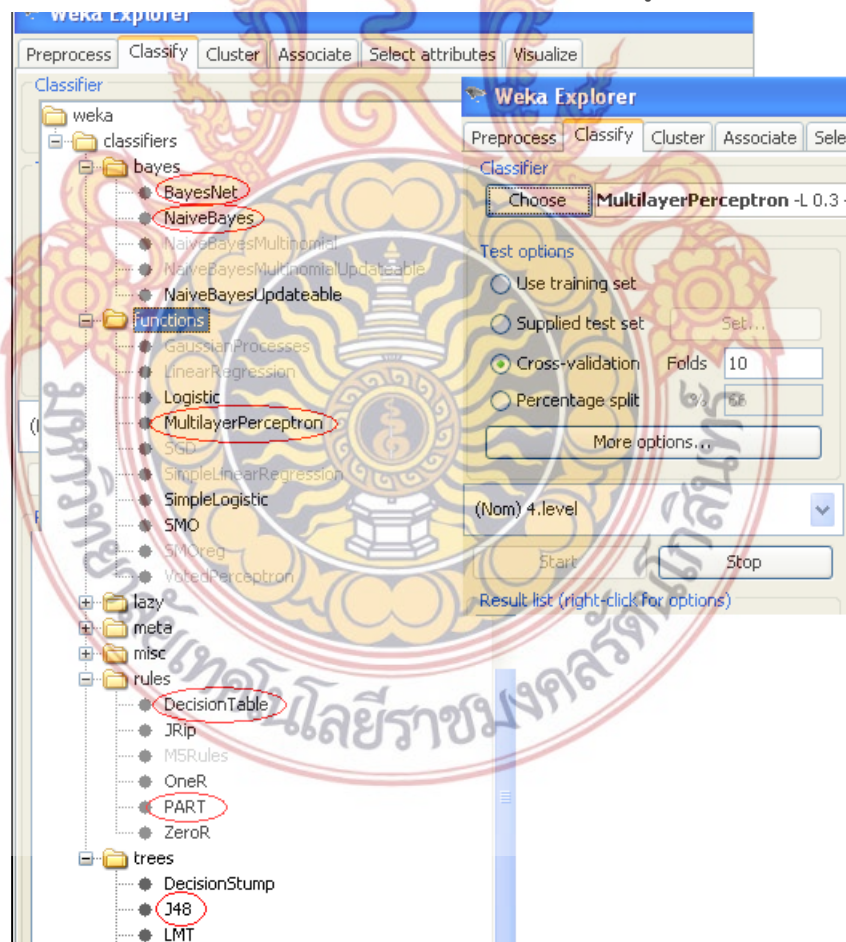
สำหรับการสร้างแบบจำลองเพื่อการพยากรณ์ความเครียดนั้น จะทำการสร้างแบบจำลอง 6 แบบจาก 6 อัลกอริทึมเพื่อใช้ในการเปรียบเทียบว่าอัลกอริทึมใดให้ผลดีที่สุด โดยการทดลองจะทำตามลำดับอัลกอริทึม ดังนี้ Bayesian Network, Naïve Bayesian, Multilayer Perceptron, Decision Table, Decision Rules: Partial Rules, Decision Tree: C4.5 จากหน้าจอโปรแกรม Weka เลือก Explorer >> Classify แล้วเลือก tree พร้อมทั้งกำหนด Cross-validation fold 10 ทั้ง 6 อัลกอริทึม ดังภาพที่ 20



ภาพที่ 18 การใช้โปรแกรม Weka เพื่อเปิดไฟล์ *.arff ขึ้นมาทำงาน



ภาพที่ 19 การใช้โปรแกรม Weka เพื่อวิเคราะห์ข้อมูลในเบื้องต้น



ภาพที่ 20 การใช้โปรแกรม Weka เพื่อสร้างแบบจำลอง 6 อัลกอริทึม

บทที่ 4 ผลการวิจัย

สำหรับผลการวิจัย ผู้วิจัยได้นำเสนอตามลำดับคือ การประเมินประสิทธิภาพ การออกแบบ การทดลอง และผลการทดลองเพื่อประเมินว่า 1) อัลกอริทึมใดที่เหมาะสมในการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด 2) แอทริบิวใดที่เหมาะสมในการนำมาใช้สร้างแบบจำลองเพื่อพยากรณ์ความเครียด

1. การประเมินประสิทธิภาพ

การประเมินประสิทธิภาพจะดูจากผลการทำนายของแบบจำลองแต่ละอัลกอริทึมซึ่งหาได้จาก Confusion matrix แสดงไว้ดังตารางที่ 9 โดย matrix นี้จะเป็นการประเมินผลลัพธ์การทำนายกับผลลัพธ์จริงๆ ที่หาได้ และค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าความเหวี่ยง (F-Measure) (Yang G., Mukherjee S., and I. V. Ramakrishnan, 2003) สามารถหาได้จากการนำเอาข้อมูลในตารางที่ 9 มาคำนวณด้วยสมการที่ 1 ถึงสมการที่ 4 ตามลำดับ จากหน้าจอโปรแกรม Weka จะพิจารณาจากค่าเฉลี่ยถ่วงน้ำหนัก (Weighted Avg.) จากทุกๆ คลาส ดังแสดงในภาพที่ 21

ตารางที่ 9 Confusion matrix

Actual Class	Predicted Class	
	Class = Yes	Class = No
Class = Yes	a (TP)	b (FN)
Class = No	c (FP)	d (TN)

$$Accuracy = \frac{a + d}{a + b + c + d} \quad (1)$$

$$Precision = \frac{a}{a + c} \quad (2)$$

$$Recall = \frac{a}{a + b} \quad (3)$$

$$F - Measure = \frac{2a}{2a + b + c} \quad (4)$$

The screenshot shows the Weka software interface with the Classifier tab selected. The chosen classifier is MultilayerPerceptron. The test options are set to Cross-validation with 10 folds. The classifier output shows the following metrics:

- Correctly Classified Instances: 247 (82.3333 %)
- Incorrectly Classified Instances: 53 (17.6667 %)
- Kappa statistic: 0.7513
- Mean absolute error: 0.082
- Root mean squared error: 0.2406
- Relative absolute error: 28.7284 %
- Root relative squared error: 63.7564 %
- Coverage of cases (0.95 level): 93.3333 %
- Mean rel. region size (0.95 level): 27.6 %
- Total Number of Instances: 300

The Detailed Accuracy By Class table is as follows:

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.85	0.011	0.85	0.85	0.85	0.991	0
	0.889	0.027	0.954	0.889	0.92	0.976	1
	0.899	0.109	0.802	0.899	0.848	0.954	2
	0.531	0.063	0.5	0.531	0.515	0.915	3
	0.625	0.022	0.769	0.625	0.69	0.97	4
Weighted Avg.	0.823	0.057	0.829	0.823	0.824	0.963	

The Confusion Matrix is as follows:

	a	b	c	d	e	<-- classified as
a	17	3	0	0	0	a = 0
b	3	104	10	0	0	b = 1
c	0	2	89	7	1	c = 2
d	0	0	10	17	5	d = 3
e	0	0	2	10	20	e = 4

ภาพที่ 21 วิธีการวัดประสิทธิภาพของแบบจำลอง

2. การออกแบบการทดลอง

การเก็บข้อมูลของงานวิจัยนี้ใช้วิธีการให้นักศึกษาตอบแบบประเมินความเครียดจากการสังเกตอาการ, พฤติกรรมหรือความรู้สึกของตนเองในระยะเวลา 2 เดือน โดยกำหนดให้ช่วงเวลาของการเก็บข้อมูลในแต่ละรอบเป็นดังนี้ รอบที่ 1 เก็บสิ้นเดือนที่ 2 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 1 - 2 รอบที่ 2 เก็บสิ้นเดือนที่ 4 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 3 - 4 รอบที่ 3 เก็บสิ้นเดือนที่ 6 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 5 - 6 และรอบที่ 4 เก็บสิ้นเดือนที่ 8 เพื่อให้นักศึกษาประเมินพฤติกรรมของตนเองในเดือนที่ 7 - 8 ความหมายของการทดลองแต่ละแบบมีดังนี้

1. การทดลองที่ 1 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 - 8 รวมระยะเวลาย้อนหลัง 8 เดือนต่อเนื่อง
2. การทดลองที่ 2 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 3 - 8 รวมระยะเวลาย้อนหลัง 6 เดือนต่อเนื่อง
3. การทดลองที่ 3 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 5 - 8 รวมระยะเวลาย้อนหลัง 4 เดือนต่อเนื่อง
4. การทดลองที่ 4 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 7 - 8 รวมระยะเวลาย้อนหลัง 2 เดือนต่อเนื่อง
5. การทดลองที่ 5 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียดตั้งแต่เดือนที่ 1 - 6 รวมระยะเวลา 6 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้ในรอบที่ 4

(เดือนที่ 7 – 8) มาใช้ จึงเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 6 เพื่อพยากรณ์ความเครียด ตอนปลายเดือนที่ 8

6. การทดลองที่ 6 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียด ตั้งแต่เดือนที่ 1 – 4 รวมระยะเวลา 4 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้ในรอบที่ 3 – 4 (เดือนที่ 5 – 8) มาใช้ จึงเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 4 เพื่อพยากรณ์ความเครียด ตอนปลายเดือนที่ 8
7. การทดลองที่ 7 คือ การพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 โดยใช้ข้อมูลความเครียด ตั้งแต่เดือนที่ 1 – 2 รวมระยะเวลา 2 เดือน โดยไม่ได้นำข้อมูลที่เก็บได้ในรอบที่ 2, 3, 4 (เดือนที่ 3 – 8) มาใช้ นับเป็นการใช้ข้อมูลตั้งแต่เดือนที่ 1 – 2 เพื่อพยากรณ์ความเครียด ตอนปลายเดือนที่ 8

โดยสรุปคือ การทดลองที่ 1, 2, 3, 4 เป็นการทดลองที่ช่วงเวลาของการเก็บข้อมูลกับช่วงเวลาของการพยากรณ์ต่อเนื่องกัน ส่วนการทดลองที่ 5, 6 และ 7 นั้น เป็นการทดลองที่ช่วงเวลาของการเก็บข้อมูลกับช่วงเวลาในการพยากรณ์ไม่ต่อเนื่องกัน ดังนั้น ในแต่ละอัลกอริทึมจะทำการทดลอง จำนวน 7 ครั้ง โดยแต่ละครั้งจะมีความแตกต่างกันตรงจำนวนแอทริบิวต์นำมาใช้ในการประมวลผล ดังสรุปไว้ใน ตารางที่ 10

ตารางที่ 10 แอทริบิวต์ที่ใช้ในแต่ละการทดลอง

แอทริบิวต์ที่ใช้ในการทดลอง	การทดลองที่						
	1	2	3	4	5	6	7
ข้อมูลรอบที่ 1 จำนวน 20 แอทริบิวต์	✓				✓	✓	✓
ข้อมูลรอบที่ 2 จำนวน 20 แอทริบิวต์	✓	✓			✓	✓	
ข้อมูลรอบที่ 3 จำนวน 20 แอทริบิวต์	✓	✓	✓		✓		
ข้อมูลรอบที่ 4 จำนวน 20 แอทริบิวต์	✓	✓	✓	✓			
ระดับความเครียดของการเก็บข้อมูลรอบที่ 4	✓	✓	✓	✓	✓	✓	✓

3. ผลการทดลอง

ผลของการทดลองจะทำให้ทราบว่า 1) อัลกอริทึมใดที่เหมาะสมในการสร้างแบบจำลองเพื่อพยากรณ์ความเครียด 2) แอทริบิวต์ใดที่เหมาะสมในการนำมาใช้สร้างแบบจำลองเพื่อพยากรณ์ความเครียด ประสิทธิภาพของการทดลองโดยใช้อัลกอริทึม Bayesian Network, Naïve Bayesian, Multilayer Perceptron (MLP), Decision Table, Partial Rules Decision (PART), Decision Tree (C4.5) ถูกแสดงไว้ในตารางที่ 11 ถึงตารางที่ 16 ตามลำดับ

ตารางที่ 11 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Bayesian Network

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	75%	0.762	0.75	0.753
2	76.67%	0.785	0.767	0.772
3	82.33%	0.829	0.823	0.824
4	82.667%	0.84	0.827	0.828
5	41.67%	0.435	0.417	0.422
6	44.33%	0.454	0.443	0.446
7	35.33%	0.354	0.353	0.348

ตารางที่ 12 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Naïve Bayesian

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	74.67%	0.754	0.747	0.748
2	77.67%	0.796	0.777	0.783
3	81.67%	0.824	0.817	0.817
4	82%	0.832	0.82	0.82
5	40.67%	0.418	0.407	0.409
6	44.67%	0.456	0.447	0.449
7	36%	0.354	0.36	0.35

ตารางที่ 13 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Multilayer Perceptron (MLP)

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	79.67%	0.795	0.797	0.794
2	81%	0.81	0.81	0.809
3	79%	0.791	0.79	0.79
4	87.67%	0.875	0.877	0.874
5	55%	0.55	0.55	0.545
6	56.33%	0.562	0.563	0.562
7	52.67%	0.522	0.527	0.524

ตารางที่ 14 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Decision Table

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	59%	0.598	0.59	0.578
2	58.33%	0.584	0.583	0.57
3	58.33%	0.59	0.583	0.571
4	60%	0.624	0.607	0.595
5	39%	0.353	0.39	0.345
6	37.67%	0.349	0.377	0.337
7	36%	0.345	0.36	0.325

ตารางที่ 15 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Partial Rules Decision (PART)

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	66.67%	0.672	0.667	0.668
2	69.33%	0.692	0.693	0.692
3	68.67%	0.695	0.687	0.686
4	76.33%	0.766	0.763	0.761
5	49%	0.484	0.49	0.485
6	48.67%	0.473	0.487	0.477
7	46%	0.457	0.46	0.456

ตารางที่ 16 ประสิทธิภาพของแต่ละการทดลองโดยใช้ Decision Tree (C4.5)

การทดลองที่	Accuracy	Precision	Recall	F-Measure
1	72.33%	0.717	0.723	0.716
2	77%	0.771	0.77	0.765
3	76.33%	0.765	0.763	0.762
4	73%	0.726	0.73	0.727
5	51.33%	0.5	0.513	0.503
6	51%	0.49	0.51	0.497
7	47.67%	0.462	0.477	0.466


```

Classifier output
=== Summary ===

Correctly Classified Instances      243          81  %
Incorrectly Classified Instances    57           19  %
Kappa statistic                     0.7315
Mean absolute error                  0.0805
Root mean squared error              0.254
Relative absolute error              28.1939 %
Root relative squared error          67.3183 %
Coverage of cases (0.95 level)      88  %
Mean rel. region size (0.95 level)  25.1333 %
Total Number of Instances           300

=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
                0.7      0.011    0.824      0.7      0.757      0.989     0
                0.889    0.071    0.889      0.889    0.889      0.964     1
                0.859    0.095    0.817      0.859    0.837      0.929     2
                0.563    0.056    0.545      0.563    0.554      0.882     3
                0.688    0.026    0.759      0.688    0.721      0.95      4
Weighted Avg.   0.81      0.068    0.81       0.81     0.809      0.944

=== Confusion Matrix ===

  a  b  c  d  e  <-- classified as
14  6  0  0  0  |  a = 0
 3 104 8  2  0  |  b = 1
 0  7 85  6  1  |  c = 2
 0  0  8 18  6  |  d = 3
 0  0  3  7 22  |  e = 4

```

ภาพที่ 22 ผลลัพธ์บางส่วนของการทดลองที่ 2 อัลกอริทึม Multilayer Perceptron (MLP)

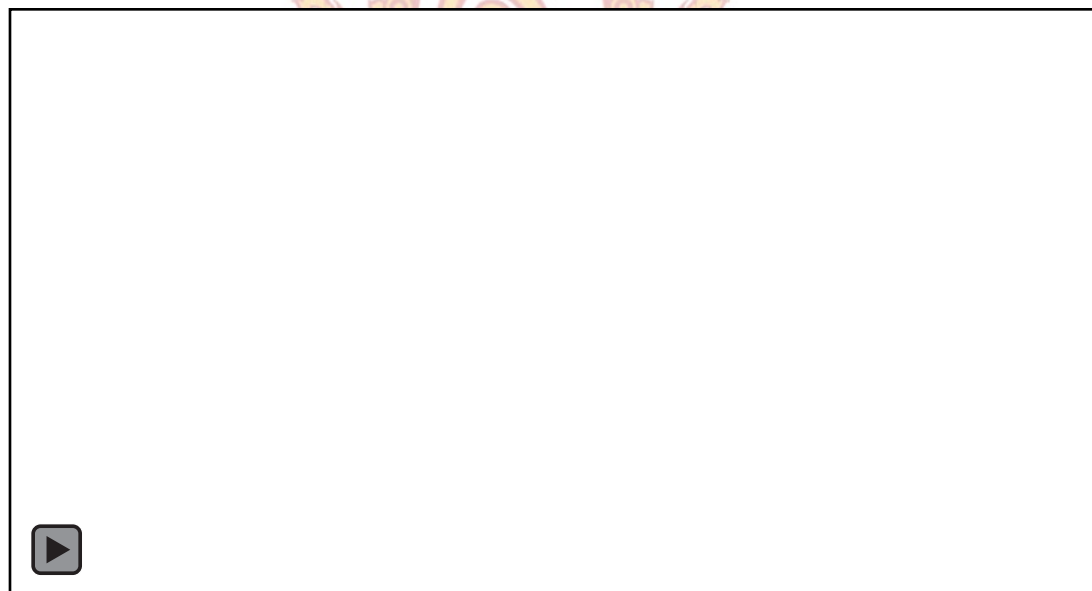
3.1 การเลือกอัลกอริทึมที่เหมาะสม

ภาพที่ 23 – ภาพที่ 29 เป็นการนำข้อมูลจากตารางที่ 11 ถึงตารางที่ 16 มาแสดงในรูปแบบของแผนภูมิเพื่อเปรียบเทียบค่า Accuracy, Precision, Recall, F-Measure จำแนกตามอัลกอริทึมตั้งแต่การทดลองที่ 1 ถึงการทดลองที่ 7 ตามลำดับ สามารถสรุปได้ดังนี้

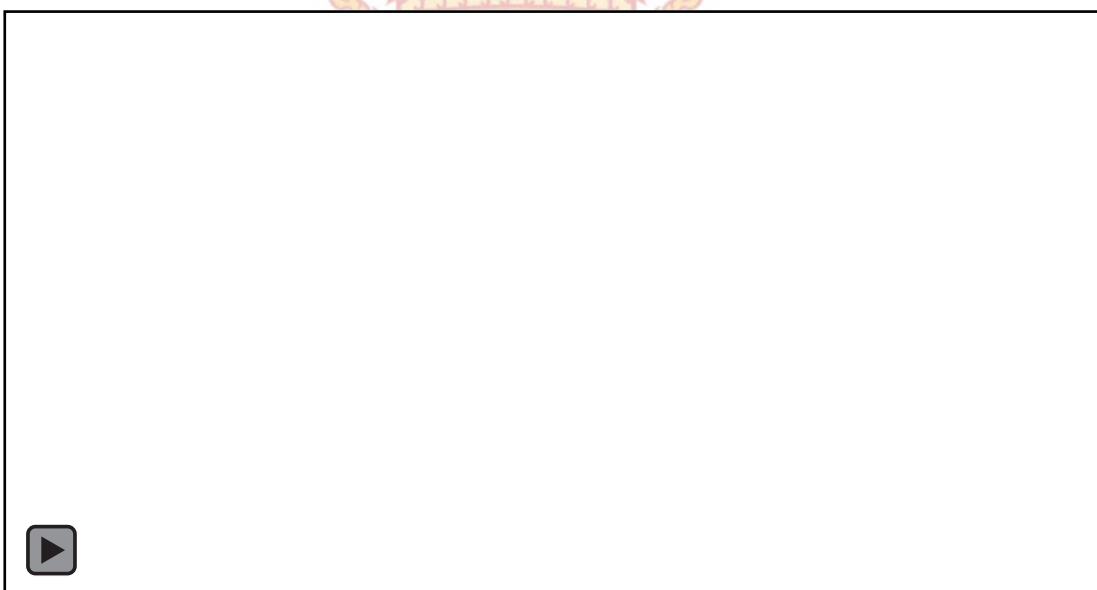
- จากภาพที่ 23 การทดลองที่ 1 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
 - จากภาพที่ 24 การทดลองที่ 2 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
 - จากภาพที่ 25 การทดลองที่ 3 อัลกอริทึม Bayesian Network ให้ผลดีที่สุด
 - จากภาพที่ 26 การทดลองที่ 4 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
 - จากภาพที่ 27 การทดลองที่ 5 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
 - จากภาพที่ 28 การทดลองที่ 6 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
 - จากภาพที่ 29 การทดลองที่ 7 อัลกอริทึม Multilayer Perceptron (MLP) ให้ผลดีที่สุด
- ดังนั้น อัลกอริทึม Multilayer Perceptron (MLP) คืออัลกอริทึมที่เหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียด



ภาพที่ 23 ผลการทดลองที่ 1 จำแนกตามอัลกอริทึม







ภาพที่ 29 ผลการทดลองที่ 7 จำแนกตามอัลกอริทึม

3.2 การเลือกแอทริบิวต์ที่เหมาะสม

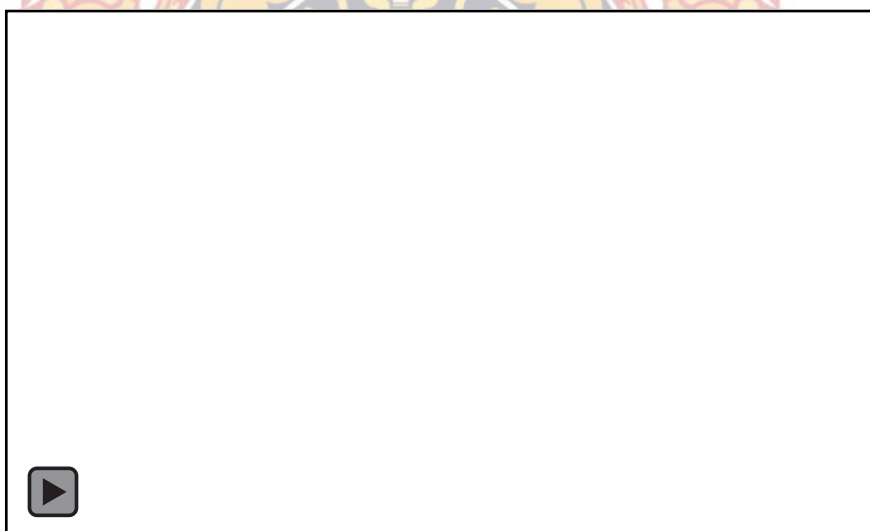
เพื่อให้เห็นภาพที่ชัดเจนมากขึ้น ผู้วิจัยจึงนำผลการทดลองจาก ตารางที่ 11 ถึงตารางที่ 16 มาสร้างแผนภูมิดังภาพที่ 30 ถึงภาพที่ 35 เพื่อทำการเปรียบเทียบว่าแอทริบิวต์ตามการทดลองที่เท่าไรเหมาะสมที่จะนำมาใช้ในการพยากรณ์ความเครียด จากแผนภูมิจะเห็นว่า การทดลองที่ 4 ของทุกอัลกอริทึมยกเว้น Decision Tree (C4.5) ให้ค่า Accuracy, Precision, Recall, F-Measure สูงที่สุด เหตุผลเนื่องจากตัวแปรตามคือระดับความเครียดของการเก็บข้อมูลรอบที่ 4 นั้นเอง หากไม่นับ

การทดลองที่ 4 แล้ว การทดลองที่ให้ผลดีรองลงมามีความแตกต่างกันในแต่ละอัลกอริทึม สามารถสรุปผลได้ดังนี้

- จากภาพที่ 30 สำหรับอัลกอริทึม Bayesian Network, การทดลองที่ 3 ให้ผลดีที่สุด
- จากภาพที่ 31 สำหรับอัลกอริทึม Naïve Bayesian, การทดลองที่ 3 ให้ผลดีที่สุด
- จากภาพที่ 32 สำหรับอัลกอริทึม Multilayer Perceptron, การทดลองที่ 2 ให้ผลดีที่สุด
- จากภาพที่ 33 สำหรับอัลกอริทึม Decision Table, การทดลองที่ 1 ให้ผลดีที่สุด
- จากภาพที่ 34 สำหรับอัลกอริทึม Partial Rules Decision, การทดลองที่ 2 ให้ผลดีที่สุด
- จากภาพที่ 35 สำหรับอัลกอริทึม Decision Tree (C4.5), การทดลองที่ 2 ให้ผลดีที่สุด

จะเห็นว่า โดยส่วนใหญ่แล้วการทดลองที่ 2 ให้ผลดีรองลงมาจากการทดลองที่ 4 จึงกล่าวได้ว่าการทดลองที่ 2 ซึ่งใช้ข้อมูลต่อเนื่อง 6 เดือนย้อนหลัง เช่น ใช้ข้อมูลเดือนมีนาคม – สิงหาคมเพื่อพยากรณ์ความเครียดตอนสิ้นเดือนสิงหาคมเหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียดการใช้ข้อมูลย้อนหลังที่เพิ่มขึ้นจาก 6 เดือนเป็น 8 เดือน เช่น ใช้ข้อมูลเดือนมกราคม – สิงหาคม (แบบการทดลองที่ 1)กลับทำให้ประสิทธิภาพในการพยากรณ์ลดลง

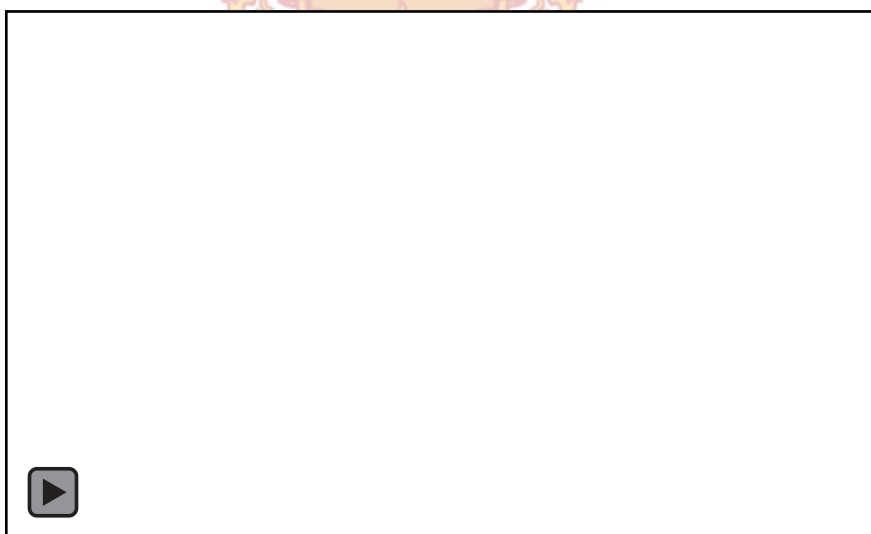
สำหรับการทดลองที่ 5, 6 และ 7 นั้น เป็นการทดลองที่ช่วงเวลาของการเก็บข้อมูลกับช่วงเวลาในการพยากรณ์ไม่ต่อเนื่องกัน จึงทำให้ผลการทดลองของทุกอัลกอริทึมต่ำกว่าการทดลองที่ 1, 2, 3 และ 4 (ที่ช่วงเวลาของการเก็บข้อมูลต่อเนื่องกับช่วงเวลาในการพยากรณ์) จากการเปรียบเทียบระหว่างการทดลองที่ 5, 6 และ 7 พบว่าการทดลองที่ 7 ให้ค่า Accuracy, Precision, Recall, F-Measure ต่ำที่สุด เนื่องจากมีระยะห่างถึง 6 เดือนระหว่างช่วงการเก็บข้อมูลและช่วงการพยากรณ์ เช่น เก็บข้อมูลความเครียดช่วงสิ้นเดือนกุมภาพันธ์เพียงครั้งเดียวเพื่อนำไปใช้พยากรณ์ความเครียดตอนสิ้นเดือนสิงหาคม จึงทำให้ความถูกต้องในการพยากรณ์ต่ำนั่นเอง



ภาพที่ 30 ประสิทธิภาพของอัลกอริทึม Bayesian Network โดยใช้ข้อมูลในตารางที่ 11



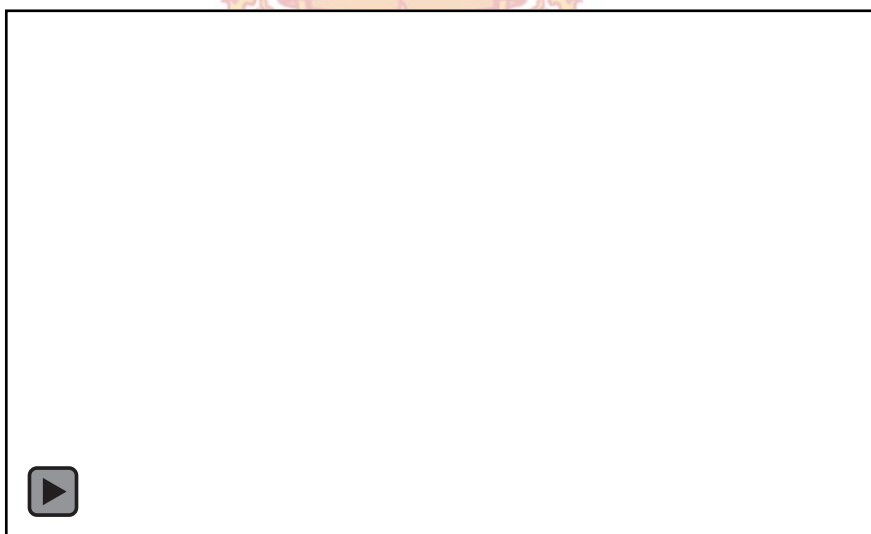
ภาพที่ 31 ประสิทธิภาพของอัลกอริทึม Naïve Bayesian โดยใช้ข้อมูลในตารางที่ 12



ภาพที่ 32 ประสิทธิภาพของอัลกอริทึม Multilayer Perceptron โดยใช้ข้อมูลในตารางที่ 13



ภาพที่ 33 ประสิทธิภาพของอัลกอริทึม Decision Table โดยใช้ข้อมูลในตารางที่ 14



ภาพที่ 34 ประสิทธิภาพของอัลกอริทึม Partial Rules Decision โดยใช้ข้อมูลในตารางที่ 15



ภาพที่ 35 ประสิทธิภาพของอัลกอริทึม Decision Tree (C4.5) โดยใช้ข้อมูลในตารางที่ 16



บทที่ 5

สรุปผล อภิปรายผลและข้อเสนอแนะ

1. สรุปผลการวิจัย

การพัฒนาแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูลนี้ เริ่มต้นจากการเก็บข้อมูลความเครียดของกลุ่มตัวอย่างจำนวน 300 คนตลอดระยะเวลา 8 เดือน เพื่อนำมาสร้างแบบจำลองเพื่อใช้ในการพยากรณ์โอกาสการเกิดความเครียดเมื่อสิ้นเดือนที่ 8 โดยมีการนำอัลกอริทึมทางด้านการทำเหมืองข้อมูลจำนวน 6 อัลกอริทึมมาสร้างแบบจำลอง ได้แก่ Bayesian Network, Naïve Bayesian, Multilayer Perceptron, Decision Table, Decision Rules: Partial Rules และ Decision Tree: C4.5 นอกจากนี้ในแต่ละอัลกอริทึมจะมีการทดลองจำนวน 7 ครั้ง โดยการทดลองแต่ละครั้งจะมีความแตกต่างกันตรงแอทริบิวต์ที่นำมาทดลอง ทั้งในส่วนของจำนวนของแอทริบิวต์และช่วงระยะเวลาในการเก็บข้อมูลแต่ละแอทริบิวต์นั้น งานวิจัยนี้มุ่งเน้นตอบโจทย์วิจัยใน 2 ประเด็น คือ 1) แบบจำลองที่สร้างโดยอัลกอริทึมใดให้ผลดีที่สุดในการพยากรณ์ความเครียด 2) ควรใช้ระยะเวลาเท่าใดในการเก็บข้อมูลจึงจะให้ผลดีที่สุดในการพยากรณ์ความเครียด

2. การอภิปรายผล

จากการทดลองสามารถตอบโจทย์วิจัยได้ดังนี้ 1) อัลกอริทึม Multilayer Perceptron (MLP) คืออัลกอริทึมที่เหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียด ซึ่งโดยส่วนใหญ่แล้วการทดลองที่ 2 ให้ผลดีรองลงมาจากทดลองที่ 3 จึงกล่าวได้ว่า 2) ควรใช้ข้อมูลย้อนหลัง 6 เดือนต่อเนื่องเพื่อพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 เนื่องจากการทดลองที่ 2 ซึ่งใช้ข้อมูลต่อเนื่อง 6 เดือนย้อนหลัง เช่น ใช้ข้อมูลเดือนที่ 3 – 8 เพื่อพยากรณ์ความเครียดตอนสิ้นเดือนที่ 8 เหมาะสมที่สุดในการนำมาใช้เพื่อพยากรณ์ความเครียด ส่วนการใช้ระยะเวลาที่ไม่ต่อเนื่องระหว่างช่วงการเก็บข้อมูลกับช่วงการพยากรณ์ เช่น นำข้อมูลที่เก็บได้เมื่อสิ้นเดือนกุมภาพันธ์มาใช้พยากรณ์ความเครียดตอนสิ้นเดือนสิงหาคม จะทำให้ประสิทธิภาพของการพยากรณ์ไม่เท่าที่ควร

3. ข้อเสนอแนะ

สำหรับการต่อยอดเพื่อให้งานวิจัยมีความละเอียดมากยิ่งขึ้นนั้น มีแนวทางคือ

1. นอกจากแบบประเมินความเครียดของกรมสุขภาพจิตแล้ว แบบประเมินสามารถนำมาจากแหล่งอื่นๆ ที่มีความน่าเชื่อถือ แล้วนำปัจจัยทั้งหมดมารวมกันเพื่อนำมาวิเคราะห์ปัจจัย (Factors Analysis) ในการเลือกทำแบบสอบถามเพื่อให้เกิดความน่าเชื่อถือในแบบสอบถามและในตัวงานวิจัย

2. กลุ่มตัวอย่างควรมีความหลากหลายเพราะความเครียดอาจมีเหตุผลจากปัจจัยที่หลากหลาย เช่น อายุ อาชีพ พื้นที่ การศึกษา และที่สำคัญคือช่วงเวลา เช่น ช่วงการสอบ ช่วงการประท้วงทางการเมือง ช่วงเทศกาลต่างๆ เป็นต้น ซึ่งความเครียดจะแตกต่างกันไป จึงควรนำปัจจัยเหล่านี้มาร่วมวิเคราะห์ด้วย

3. ในอนาคตอาจมีการจำแนกการทดลองโดยแยกข้อมูลออกเป็นภูมิภาคต่างๆ เช่น ภาคเหนือ ภาคกลาง เป็นต้น เพราะลักษณะของคนในแต่ละภูมิภาคมีความแตกต่างกัน จะทำให้งานวิจัยมีความละเอียดมากยิ่งขึ้น



บรรณานุกรม

- กนกรัตน์ สุขตุงคะ. **ความเครียดและวิถีคลายความเครียด**. กรุงเทพฯ : ภาควิชาเวชศาสตร์ คณะแพทยศาสตร์ศิริราชพยาบาลมหาวิทยาลัยมหิดล, 2545.
- กรมสุขภาพจิต กระทรวงสาธารณสุข. **ความเครียดของคนไทย : การศึกษาระดับชาติปี 2546**. กรุงเทพฯ : ปิยอนด์ พับลิชชิ่ง., 2546.
- กรมสุขภาพจิต กระทรวงสาธารณสุข. **การพัฒนาแบบประเมินและการวิเคราะห์ความเครียดด้วยตนเองสำหรับประชาชนชาวไทยด้วยคอมพิวเตอร์**. กรุงเทพฯ: กระทรวงสาธารณสุข, 2541.
- กรมสุขภาพจิต กระทรวงสาธารณสุข. **คู่มือการส่งเสริมสุขภาพวัยทำงาน**. นนทบุรี :กรมสุขภาพจิต, 2548.
- กรมสุขภาพจิต กระทรวงสาธารณสุข. **คู่มือคลายเครียด**. พิมพ์ครั้งที่ 6. กรุงเทพฯ; ดีไซน์คอนดักซ์, 2541.
- กรมสุขภาพจิต กระทรวงสาธารณสุข. **“รายงานวิจัยการพัฒนาแบบประเมินและวิเคราะห์ความเครียดด้วยตนเองสำหรับประชาชนชาวไทยด้วยคอมพิวเตอร์”**, กรุงเทพฯ: สำนักพิมพ์ทีคอม (T-COM), 2541.
- กิตติศักดิ์ สุมาลย์. **“การคัดกรองสุขภาพเบื้องต้นโดยใช้เทคนิคเหมือนข้อมูล”**. สารนิพนธ์ วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยธุรกิจบัณฑิต, 2555.
- ชัชชฎา วันดี, จิรัฏฐา ภูบุญอบ, ฉัตรเกล้า เจริญผล. **“ประสิทธิภาพการจำแนกข้อมูลการเลือกอาชีพโดยอัตโนมัติด้วยเทคนิคเหมือนข้อมูล”**. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม ฉบับพิเศษ การประชุมทางวิชาการมหาวิทยาลัยมหาสารคาม ครั้งที่ 9. (2556): 263-270.
- ชาญยุทธน์ วิหคโต และนิตยา พันธุ์เวทย์. **ร่วมมือ...ร่วมใจ...ป้องกันโรคเรื้อรังในปัจจุบันและพร้อมเข้าสู่ประชาคมเศรษฐกิจอาเซียน (AEC) กัน: Act Now**. (ระบบออนไลน์) แหล่งข้อมูล:<http://thaincd.com/good-stories-view.php?id=8193> (30 สิงหาคม 2557)
- นิภา นีรุตติกุล. **การพยากรณ์การขาย**. กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, 2549.
- ปราณี สุทธิสุนทร, จุฑาธิป ศีลบุตร, ดุชนิ คามิ และเกรียงศักดิ์ ชื้อเลื่อม. **“ระดับความเครียดของบุคลากรสถาบันพัฒนาการสาธารณสุขอาเซียน”**. วารสารสาธารณสุขและการพัฒนา. 7, 1 (มกราคม-เมษายน 2552): 61.
- ปิยะวรรณ เลิศพานิช . **“การศึกษาความเครียด สาเหตุของความเครียด และวิธีการเผชิญความเครียดทางการเรียนของนักเรียนระดับมัธยมศึกษาตอนปลายในโรงเรียนมัธยมศึกษา สังกัดกรมสามัญศึกษา ส่วนกลาง กลุ่มที่4”**. ปริญญาานิพนธ์ กศ.ม.(จิตวิทยาและการแนะแนว). กรุงเทพฯ.บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ, 2542.

- มาลิวรรณ บุญพลอย. “เหมืองข้อมูลและการค้นหาองค์ความรู้สำหรับโรคกล้ามเนื้ออ่อนแรง”, วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2551.
- รักถิ่น เหลาหา. “การพยากรณ์ความเสี่ยงการเกิดโรคมะเร็งปอดโดยใช้ทฤษฎีของการทำเหมืองข้อมูล”, วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น, 2553.
- วีระยุทธ มายุศิริ, จารีย์ ทองคำ, วาทีนี สุขมาก. “การพัฒนาแบบจำลองเพื่อการพยากรณ์การรักษาซ้ำของผู้ป่วยโรคจิตเภทโดยเทคนิคเหมืองข้อมูล.” วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยมหาสารคาม ฉบับพิเศษ. (2557): 144-153.
- ศุภกรใจ วุฒิกิจโกศล. “การใช้เทคนิคการทำเหมืองข้อมูลในผู้ป่วยข้อไหล่ติด โรงพยาบาลพระนั่งเกล้า”. วิทยานิพนธ์สาขาวิชาการจัดการเทคโนโลยีสารสนเทศ. นครศรีธรรมราช. บัณฑิตวิทยาลัย มหาวิทยาลัยวลัยลักษณ์, 2551.
- สถาบันกัลยาณ์ราชนครินทร์. คู่มือคลายเครียด (ฉบับปรับปรุงใหม่). กรุงเทพฯ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด, 2555.
- สมิต อาชวนิจกุล. เครียดเป็นบ้า. พิมพ์ครั้งที่ 3 กรุงเทพฯ : สามัคคีสาร(ดอกหญ้า), 2542.
- เสถียร วิชาเรือง. “การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในกระบวนการอนุมัติสินเชื่อ.” งานค้นคว้าอิสระ วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยธุรกิจบัณฑิต, 2553.
- สุภาพ หวังช่อกลาง. “การศึกษาความสัมพันธ์ระหว่างความเครียดและการเผชิญความเครียดของนักศึกษาสาขาวิชาสาธารณสุขศาสตร์ คณะสาธารณสุขศาสตร์และเทคโนโลยีสุขภาพวิทยาลัยนครราชสีมา.” วิทยานิพนธ์ วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยศรีนครินทรวิโรฒ, 2554.
- อรพรรณ ลีบุญธวัชชัย. การพยาบาลสุขภาพจิตและจิตเวช. กรุงเทพฯ: ด้านสุขภาพการพิมพ์, 2545.
- อังคณา พิจารโชติ. “ระบบสนับสนุนการตัดสินใจสำหรับวิเคราะห์ปัจจัยเสี่ยงการเป็นโรคเบาหวานโดยใช้เทคนิคคาด้าไมน์นิง”, วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น, 2553.
- อัจฉรา จันทร์ฉาย. การพยากรณ์เพื่อการตัดสินใจทางธุรกิจ. กรุงเทพฯ: สำนักพิมพ์แห่งมหาวิทยาลัยจุฬาลงกรณ์, 2544.
- เอกสิทธิ์ พัทธวงศ์ศักดิ์. การวิเคราะห์ข้อมูลด้วยเทคนิคคาด้าไมน์นิงเบื้องต้น. กรุงเทพฯ: บริษัท เอเชีย ดิจิตอลการพิมพ์ จำกัด, 2557.
- อุกฤษ ปัจฉิม. “การประยุกต์ใช้เทคนิคการทำเหมืองข้อมูลในการทำนายระดับน้ำสูงสุด”. วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมทรัพยากรน้ำ. คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2546.
- Guizhen Yang, Saikat Mukherjee, and I. V. Ramakrishnan, "On precision and Recall of MultiAttribute Data Extraction from Semistructured Sources [Electronic version]," 2003.

Han, J., Kamber, M., and Pei, J. **Data Mining Concepts and Techniques**. (3rd ed.).

United States of America: Morgan Kaufman Publishers, 2012.

Kantardzic M. **Data Mining Concepts, Models, Methods, and Algorithms**. United States of America: IEEE Press, 2003.

Sarafino, E.P. **Health Psychology**. (2nd ed.). Singapore: John Wiley & Sons Inc., 1990.





แบบประเมินและวิเคราะห์ความเครียดด้วยตนเอง

ในระยะเวลา 2 เดือนที่ผ่านมา ท่านมีอาการ พฤติกรรม หรือความรู้สึกต่อไปนี้มากน้อยเพียงใด โปรดขีดเครื่องหมาย " X " ลงในช่องแสดงระดับอาการที่เกิดขึ้นกับตัวท่านตามความเป็นจริงมากที่สุด

อาการ พฤติกรรม หรือความรู้สึก	ระดับอาการ			
	0 ไม่เคยเลย	1 เป็นครั้งคราว	2 เป็นบ่อยๆ	3 เป็นประจำ
1. นอนไม่หลับเพราะคิดมากหรือกังวลใจ				
2. รู้สึกหงุดหงิด หรือรำคาญใจ				
3. ทำอะไรไม่ได้เลยเพราะประสาทตึงเครียด				
4. มีความวุ่นวายใจ				
5. ไม่อยากพบปะผู้คน				
6. ปวดหัวข้างเดียวหรือปวดบริเวณขมับ ทั้ง 2 ข้าง				
7. รู้สึกไม่มีความสุขและเศร้าหมอง				
8. รู้สึกหมดหวังในชีวิต				
9. รู้สึกว่าชีวิตตนเองไม่มีคุณค่า				
10. กระวนกระวายอยู่ตลอดเวลา				
11. รู้สึกว่าตนเองไม่มีสมาธิ				
12. รู้สึกเพลียจนไม่มีแรงจะทำอะไร				
13. รู้สึกเหนื่อยหน่ายไม่อยากทำอะไร				
14. มีอาการหัวใจเต้นแรง				
15. เสียงสั่น ปากสั่น หรือมือสั่น เวลาไม่พอใจ				
16. รู้สึกกลัวผิดพลาดในการกระทำสิ่งต่างๆ				
17. ปวดหัวหรือเกร็งกล้ามเนื้อบริเวณท้ายทอย หรือหัวไหล่				
18. ตื่นเต้นง่ายกับเหตุการณ์ที่ไม่คุ้นเคย				
19. มึนงงหรือเวียนศีรษะ				
20. ความสุขทางเพศลดลง				

ผลการประเมินและวิเคราะห์ความเครียด

คะแนนที่ได้	ความหมายของคะแนน
0 - 5	<p>ท่านมีความเครียดในระดับต่ำกว่าเกณฑ์อย่างมาก ทฤษฎีถือว่ามีความเป็นไปได้เพียงเล็กน้อยเท่านั้น ที่จะมีความเครียดในระดับต่ำมากเช่นนี้ ในกรณีของท่านอาจมีความหมายว่า</p> <ul style="list-style-type: none"> * ท่านตอบคำถามไม่ตรงตามความเป็นจริง หรือ * ท่านอาจเข้าใจคำสั่งหรือข้อความคลาดเคลื่อนไป * ท่านอาจจะเป็นคนที่ขาดแรงจูงใจในการดำเนินชีวิต มีความเฉื่อยชา ชีวิตประจำวันซ้ำซาก จำเจ น่าเบื่อหน่าย ปราศจากความตื่นเต้น
6 - 17	<p>ท่านมีความเครียดอยู่ในเกณฑ์ปกติ สามารถจัดการกับความเครียดที่เกิดขึ้นในชีวิตประจำวัน และสามารถปรับตัวกับสถานการณ์ต่างๆ ได้อย่างเหมาะสม รู้สึกพึงพอใจเกี่ยวกับตนเองและสิ่งแวดล้อมเป็นอย่างมาก รู้สึกมีพลัง มีชีวิตชีวา กระตือรือร้น มองสิ่งเร้าหรือเหตุการณ์รอบตัวว่าเป็นสิ่งท้าทายความสามารถ มีความสามารถในการจัดการกับสิ่งต่าง ๆ ได้เหมาะสม ผลผลิตของการปฏิบัติงานอยู่ในระดับสูง ความเครียดในระดับนี้ถือว่ามิมีประโยชน์ในการดำเนินชีวิต ประจำวัน เป็นแรงจูงใจที่จะนำไปสู่ความสำเร็จได้</p>
18 - 25	<p>ท่านมีความเครียดอยู่ในระดับสูงกว่าปกติเล็กน้อย มีความไม่สบายใจอันเกิดจากปัญหาในการดำเนินชีวิตประจำวัน โดยที่ปัญหาหรือข้อขัดแย้งของท่านอาจจะยังไม่ได้รับการคลี่คลายหรือแก้ไข ซึ่งถือว่าเป็นความเครียดที่พบได้ในชีวิตประจำวัน อาจไม่รู้ตัวว่ามีความเครียดหรืออาจรู้สึกได้จากการเปลี่ยนแปลงของร่างกาย อารมณ์ความรู้สึกและพฤติกรรมบ้างเล็กน้อย แต่ไม่ชัดเจนและยังพอทนได้ แม้ว่าท่านจะมีความยุ่งยากในการจัดการกับปัญหาอยู่บ้างและอาจต้องใช้เวลาในการปรับตัวมากขึ้นกว่าเดิม ท่านก็สามารถจัดการกับความเครียดได้ และไม่เป็ผลเสียต่อการดำเนินชีวิตในกรณีนี้ท่านควรผ่อนคลายความเครียดด้วยการหาความเพลิดเพลินใจ เช่น ออกกำลังกาย ดูหนัง ฟังเพลง สัน่สรรรค์กับเพื่อน ฯลฯ</p>
26 - 29	<p>ท่านมีความเครียดอยู่ในระดับสูงกว่าปกติปานกลาง ขณะนี้ท่านเริ่มมีความตึงเครียดในระดับค่อนข้างสูงและได้รับความเดือดร้อนเป็นอย่างมากจากปัญหาทางอารมณ์ที่เกิดจากปัญหาความขัดแย้งและวิกฤติในชีวิต เป็นสัญญาณเตือนขั้นต้นว่า ท่านกำลังเผชิญกับภาวะวิกฤติ และความขัดแย้งซึ่งท่านจัดการแก้ไขด้วยความยากลำบาก ลักษณะอาการดังกล่าวจะเพิ่มพูนความรุนแรงซึ่งมีผลกระทบต่อการทำงาน จำเป็นต้องหาวิธีแก้ไข</p>

คะแนนที่ได้	ความหมายของคะแนน
	ข้อขัดแย้งต่าง ๆ ให้ลดน้อยลง หรือหมดไปด้วยวิธีการอย่างใดอย่างหนึ่งและควรคลี่คลายความเครียดด้วยการพูดคุยปรึกษากับคนที่ไว้วางใจ หรือปรึกษาทางโทรศัพท์ (Hotline)
30 - 60	ท่านมีความเครียดอยู่ในระดับสูงกว่าปกติมาก กำลังตกอยู่ในภาวะตึงเครียดหรือกำลังเผชิญกับวิกฤติการณ์ในชีวิตอย่างรุนแรง เช่น การเจ็บป่วยที่รุนแรง เรือร้าง ความพิการ การสูญเสีย ปัญหาความรุนแรงในครอบครัว ปัญหาเศรษฐกิจ ซึ่งส่งผลต่อสุขภาพกาย และสุขภาพจิตอย่างชัดเจน ทำให้ชีวิตไม่มีความสุข ฟุ้งซ่าน ตัดสินใจผิดพลาด ขาดความยับยั้งชั่งใจ อาจเกิดอุบัติเหตุได้ง่าย ความเครียดในระดับนี้ถือว่ามีความรุนแรงมาก หากปล่อยไว้โดยไม่ดำเนินการอย่างใดอย่างเหมาะสมและถูกวิธีอาจนำไปสู่ความเจ็บป่วยทางจิตที่รุนแรง ซึ่งส่งผลเสียต่อตนเองและบุคคลใกล้ชิดต่อไปได้ ในระดับนี้ควรขอรับบริการHotline และคลินิกคลายเครียด ซึ่งมีอยู่ตามสถานบริการทั่วไป



ประวัติผู้วิจัย



ประวัติผู้วิจัย

1. ชื่อ สกุล ดร.ปวีณา ชัยวนารมย์
2. ตำแหน่งปัจจุบัน อาจารย์ประจำสาขาเทคโนโลยีสารสนเทศทางธุรกิจ
คณะบริหารธุรกิจ (บพิตรพิมุข จักรวรรดิ)
3. หน่วยงานที่สามารถติดต่อได้
คณะบริหารธุรกิจ (บพิตรพิมุข จักรวรรดิ)
264 ถนนจักรวรรดิ แขวงจักรวรรดิ เขตสัมพันธวงศ์ กรุงเทพฯ 10100
หมายเลขโทรศัพท์ 02-222-2814 โทรสาร 02-226-4879
e-mail: paweena.c@rmutr.ac.th, cc.paweena@gmail.com
4. ประวัติการศึกษา
 - Ph.D in Computer Science, Chulalongkorn University, Thailand, 2014.
 - M.Sc. Program in Computer Science, National Institute of Development Administration (NIDA), Thailand, 2004.
 - B.Sc. Program in Computer Science (2nd Class Honors), King Mongkut's University of Technology North Bangkok, Thailand, 2000.
5. สาขาวิชาการที่มีความชำนาญพิเศษ
Data Mining, Information Retrieval, Social Network Analysis
6. ประสบการณ์ที่เกี่ยวข้องกับการบริหารงานวิจัย
 - หัวหน้าโครงการวิจัยงบประมาณเงินแผ่นดิน ประจำปี 2558: การพัฒนาแบบจำลองเพื่อพยากรณ์โอกาสการเกิดความเครียดในหลายระดับด้วยเทคนิคการทำเหมืองข้อมูล
 - หัวหน้าโครงการวิจัยงบประมาณเงินแผ่นดิน ประจำปี 2557: เครื่องมือช่วยบริหารการพัฒนาและทดสอบระบบสำหรับผู้ประกอบการอุตสาหกรรมซอฟต์แวร์.
 - หัวหน้าโครงการวิจัยงบประมาณเงินรายได้ ประจำปี 2556: ตัวแบบวิเคราะห์โอกาสการทำงานวิจัยร่วมกันโดยใช้ความคล้ายคลึงกันของผลงานวิจัยตีพิมพ์
 - หัวหน้าโครงการวิจัยงบประมาณเงินรายได้ ประจำปี 2555: การพัฒนาต้นแบบช่วยประเมินความเข้าใจการเลือกเกมส์เพื่อส่งเสริมทักษะแก่เยาวชน