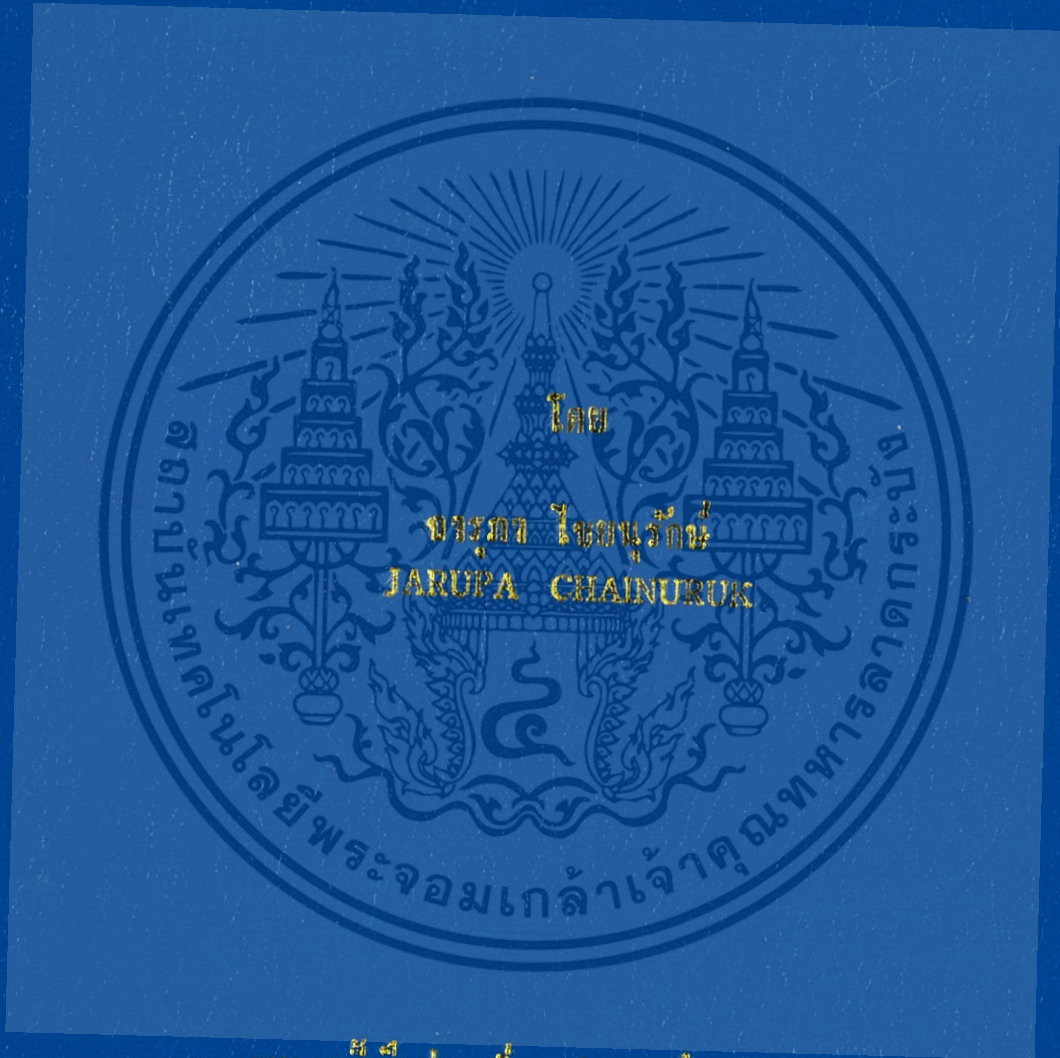


การทำนายพฤติกรรมนักเรียนโดยใช้เทคนิคต้นไม้ตัดสินใจ
STUDENT'S BEHAVIOR PREDICTION USING
DECISION TREE TECHNIQUE



รายงานชิ้นนี้เป็นส่วนหนึ่งของผลงานการศึกษาระดับ 2
ระดับปริญญาตรี สาขาวิศวกรรมไฟฟ้า ภาควิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ภาคเรียนที่ 2 ปีการศึกษา 2557

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การทำนายพฤติกรรมนักเรียนโดยใช้เทคนิคต้นไม้ตัดสินใจ

STUDENT'S BEHAVIOR PREDICTION USING
DECISION TREE TECHNIQUE

โดย



T144575



จารุภา ไชยนุรักษ์

JARUPA CHAINURUK

อาจารย์ที่ปรึกษา

ผศ.ดร.กนต์พงษ์ วรรณปัญญา

เลขหมู่.....
เลขทะเบียน.....144575
วัน,เดือน,ปี.....25 ๗.๑. 2559

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

รายงานนี้เป็นส่วนหนึ่งของวิชาการศึกษาระดับ 2
หลักสูตรวิทยาศาสตรมหาบัณฑิตสาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2557

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**STUDENT 'S BEHAVIOR PREDICTION USING
DECISION TREE TECHNIQUE**



**A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS OF THE COURSE
INDEPENDENT STUDY 2
MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

2 / 2014

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ใบรับรองการศึกษาอิสระ 2 (INDEPENDENT STUDY 2)

เรื่อง

การทำนายพฤติกรรมนักเรียนโดยใช้เทคนิคต้นไม้ตัดสินใจ STUDENT'S BEHAVIOR PREDICTION USING DECISION TREE TECHNIQUE

นางสาวจรรุภา ไชยนุรักษ์
รหัสประจำตัว 56606169

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด
รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของการศึกษาวิชาการศึกษาอิสระ 2 หลักสูตรวิทยาศาสตรมหาบัณฑิต (เทคโนโลยีสารสนเทศ) ภาคเรียนที่ 2 ปีการศึกษา 2557

.....อาจารย์ที่ปรึกษา
(ผศ.ดร.กนต์พงษ์ วรรณปัญญา)

.....กรรมการสอบ
(รศ.ดร.จันทร์บุรณัฐ สถิตวิริยวงศ์)

.....กรรมการสอบ
(ผศ.ดร.ปานวิทย์ ฐานะนาค)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2015

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	การทำนายพฤติกรรมนักเรียนโดยใช้เทคนิคต้นไม้ตัดสินใจ
นักศึกษา	นางสาวจรรุภา ไชยนุรักษ์
รหัสนักศึกษา	56606169
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	เทคโนโลยีสารสนเทศและการจัดการ
ปีการศึกษา	2556
อาจารย์ที่ปรึกษา	ผศ.ดร.กัณฑ์พงษ์ วรรณันปัญญา

บทคัดย่อ

พฤติกรรมกรทะเลาะวิวาทของวัยรุ่นเป็นปัจจัยสำคัญที่นำไปสู่การก่ออาชญากรรมในทั่วประเทศ การคาดการณ์ในช่วงต้นของพฤติกรรมนี้จะช่วยป้องกันความก้าวร้าวและความรุนแรงของวัยรุ่น ในการศึกษานี้มีวัตถุประสงค์เพื่อทำนายพฤติกรรมทะเลาะวิวาทของวัยรุ่นโดยใช้การประเมินพฤติกรรมดังต่อไปนี้: เขกหัว กระชาก ตี ตบแขน ศีรษะ ตบหน้าหรือท้ายทอย จับหัวโขก ตะ ต่อย ปาวัตถุใส่ กระตืบ รุมทำ แกล้ง โยนบกโยนน้ำ แบ่งพรรคแบ่งพวก บีบคอ บังคับให้กินของอันตราย เอาอาหารร้อนสาด นาบ จี้ ทำร้ายด้วยวัตถุ ทำร้ายด้วยอาวุธ รุมทำร้าย สาดสารเคมี กรอกสารเคมี เป็นต้น การแบ่งกลุ่มพฤติกรรมเหล่านี้ได้ใช้เทคนิคต้นไม้ตัดสินใจ โดยใช้การแบ่งระดับพฤติกรรมความรุนแรงออกเป็น 5 ระดับ จากนั้น นำเอาโมเดลไปทำการประเมินกลุ่มเป้าหมายโดยใช้กลุ่มข้อมูลเพื่อสำหรับเป็นข้อมูลการสอนให้กับโมเดล 6,757 กลุ่มตัวอย่าง และ 2,895 กลุ่มตัวอย่างใช้สำหรับการทดสอบกับโมเดล ผลการศึกษาพบว่า การจำแนกกลุ่มตัวอย่างนั้นมีความถูกต้อง 99.93% และแสดงให้เห็นว่าความเป็นไปได้ของโมเดล ที่จะทำการทำนายนั้นถูกต้องดียิ่งขึ้น เมื่อนำเอาโมเดลดังกล่าวนี้ไปใช้ร่วมกับชุดการตรวจสอบกลุ่มข้อมูล

Title	Student's Behavior Prediction Using Decision Tree Technique
Student	Miss Jarupa Chainuruk
Student ID.	56606169
Degree	Master of Science
Program	Information Technology
Major	Information Technology and Management
Academic Year	2014
Advisor	Assistant Professor Dr.Kuntpong Woraratpanya

ABSTRACT

Quarrelsome behavior of teenagers is an important factor which leads to crimes in the nationwide. Early prediction of this behavior helps prevent adolescent aggression and teenage violence. This study aims to predict the quarrelsome behavior of teenagers by evaluating the following behaviors: knocking, jerking, hitting, slap, striking, kicking, punching, throwing, trample, attacking, teasing, fling, discriminate, choke, force, splash, press down, harm, attack, assault and pour. These behaviors are classified by using decision tree technique with behavior rating to define the violence. Then this model is evaluated by the target group criteria. Based on training set with 6,757 samples, the model-generated is tested by 2,895-sample test set. The results showed that correctly classified instance is 99.93%. It is revealed that the possibility of the model will make the better accuracy prediction when applied to validation data set

กิตติกรรมประกาศ

โครงการศึกษาเครื่องมือที่ช่วยในการวิเคราะห์พฤติกรรม จะสำเร็จลุล่วงไปไม่ได้เลย ถ้าไม่ได้รับการช่วยเหลือและแรงสนับสนุนจากบุคคลสำคัญหลายท่าน ดังต่อไปนี้

ข้าพเจ้าขอกราบขอบพระคุณ ผศ.ดร.กัณฑ์พงษ์ วรรัตน์ปัญญา ซึ่งเป็นอาจารย์ที่ปรึกษาโครงการนี้ที่ให้ความกรุณาให้คำแนะนำและปรึกษา ข้าพเจ้ารู้สึกซาบซึ้งและขอกราบขอบพระคุณเป็นอย่างสูง

ขอกราบพระคุณคณาจารย์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง คณะเทคโนโลยีสารสนเทศทุกท่านที่ได้ประสิทธิ์ประสาทวิชาให้กับข้าพเจ้า

ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ทุกคนที่ให้คำแนะนำ และคอยให้กำลังใจเสมอมา

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และให้การสนับสนุนในทุกเรื่องๆ ทำให้ข้าพเจ้าสามารถทำโครงการพัฒนาระบบงานฉบับนี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์จากโครงการพัฒนาระบบงานฉบับนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

จารุภา ไชยนุรักษ์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	2
1.3 ขอบเขตของงาน.....	2
1.4 แนวทางการศึกษาและพัฒนา.....	2
1.5 ข้อตกลงเบื้องต้น.....	2
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.7 เครื่องมือที่ใช้ในการพัฒนาระบบ.....	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	4
2.1 ทฤษฎีที่เกี่ยวข้องกับวัยรุ่น.....	4
2.2 ทฤษฎีและแนวคิดเกี่ยวกับพฤติกรรม.....	5
2.3 พฤติกรรมความรุนแรง.....	8
2.4 ทฤษฎีการทำเหมืองข้อมูล.....	12
2.5 งานวิจัยที่เกี่ยวข้อง.....	23
บทที่ 3 วิธีการดำเนินการศึกษา.....	25
3.1 ข้อมูลที่ใช้กับระบบ.....	25
3.2 ขั้นตอนการดำเนินงาน.....	28
บทที่ 4 ผลการทดลอง.....	35
4.1 การใช้งานระบบ.....	35
4.2 ข้อมูลที่ใช้กับโปรแกรมประยุกต์.....	51
4.3 ผลการทดลองและวิเคราะห์.....	51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
บทที่ 5 สรุปผลการทดลองและข้อเสนอแนะ.....	52
5.1 สรุปผลการศึกษา.....	52
5.2 ประโยชน์ที่ได้รับจากการศึกษาและพัฒนาระบบ	52
5.3 ข้อเสนอแนะ.....	53
บรรณานุกรม	54
ประวัติผู้เขียน	55



สารบัญตาราง

	หน้า
2.1 ตารางตัวอย่างข้อมูลการให้บริการ.....	17
2.2 ตัวอย่างข้อมูลสภาพภูมิอากาศที่มีผลต่อการเล่นกีฬาเทนนิส.....	20
3.1 ตารางเก็บค่านำหนักคะแนน.....	30
3.2 ตารางค่านำหนักของแต่ละคำถาม.....	31
4.1 ตารางเก็บข้อมูลประวัติการทดสอบระบบ.....	32



สารบัญรูป

รูปที่	หน้า
2.1 แสดงโครงสร้างต้นไม้ตัดสินใจ (Decision Tree).....	15
2.2 ตัวอย่างต้นไม้ตัดสินใจ จำแนกกลุ่มลูกค้าใช้งานรถยนต์แต่ละยี่ห้อ.....	16
2.3 แสดงโครงสร้างต้นไม้ตัดสินใจ.....	17
2.4 แสดงการเลือก Root Node.....	22
2.5 แสดงภาพ Tree ที่สมบูรณ์.....	22
3.1 ตัวอย่างข้อมูลต้นแบบที่นำมาใช้.....	25
3.2 ตัวอย่างแบบสอบถามที่ใช้ประเมินพฤติกรรมที่เป็นผู้ถูกระทำ.....	26
3.3 ตัวอย่างแบบสอบถามที่ใช้ประเมินพฤติกรรมด้านความรู้สึก.....	27
3.4 ขั้นตอนการทำงาน.....	28
3.5 ตารางเก็บข้อมูลต้นฉบับ.....	33
3.6 รูปตัวอย่างตารางเก็บข้อมูลคำตอบของกลุ่มนักเรียนตัวอย่าง.....	33
3.7 ตารางคำตอบที่ทำการแบ่งกลุ่มผู้ตอบตามระดับความรุนแรง.....	34
4.1 โปรแกรม WEKA.....	35
4.2 โปรแกรม WEKA ในรูปแบบ Explorer.....	36
4.3 การเลือกแฟ้มเอกสารที่ต้องการนำเข้าสู่ WEKA.....	37
4.4 รูปการแสดงผลข้อมูลหลังจากนำเข้าข้อมูลเข้า WEKA.....	37
4.5 ผลลัพธ์จากการเลือกอัลกอริทึมในการจำแนกข้อมูล.....	38
4.6 ผลลัพธ์จากการนำเข้าข้อมูลเข้า WEKA.....	39
4.7 การเลือกให้ WEKAแสดงผลออกมาเป็นรูปภาพ.....	40
4.8 ผลลัพธ์ที่ได้จากโมเดลต้นแบบแสดงTree.....	41
4.9 เลือกเมนูการ Test.....	41
4.10 ผลลัพธ์ที่ได้หลังจากให้ WAKA ประมวลผล.....	42
4.11 ข้อมูลผลลัพธ์ Test Set.....	43
4.12 ขั้นตอนการเลือกการวัดประสิทธิภาพการทำนาย.....	46
4.13 ผลลัพธ์จากการ Cross Validation.....	47
4.14 ตารางข้อมูลเพื่อการทำนาย.....	48
4.15 ขั้นตอนการเลือกแฟ้มเอกสารที่จะต้องการทำนาย.....	48
4.16 การกำหนด สำหรับการทำนาย.....	49

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.17 ขั้นตอนการเลือกโมเดลต้นแบบในการทำนาย.....	50
4.18 ผลลัพธ์ที่ได้จากการทำนาย.....	50



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

เยาวชนเป็นกำลังของประเทศชาติในอนาคต เป็นวิถีที่ยังคงใช้ได้มาจนถึงปัจจุบัน การที่ประเทศจะพัฒนานั้นย่อมจะต้องมีการส่งเสริมสนับสนุนในด้านต่างๆ โดยเริ่มมาจากหน่วยที่เล็กที่สุดในสังคมก็คือ ครอบครัว ซึ่งหน้าที่หลักของสถาบันครอบครัวคือการให้ความรักและความเอาใจใส่ การให้การศึกษาเล่าเรียน ซึ่งหน้าที่หลักในการจัดการศึกษานั้นคือ โรงเรียน ซึ่งครูจะเป็นบุคคลที่มีหน้าที่ให้ความรู้ อบรมสั่งสอน ทั้งด้านคุณธรรมและจริยธรรม ส่งเสริมให้นักเรียนเป็นคนดีมีจิตสำนึกในความเป็นพลเมืองไทยและเป็นพลโลก ตามหลักสูตรแกนกลางการศึกษาขั้นพื้นฐาน 2551 และเนื่องมาจากสภาพสังคมปัจจุบันที่เปลี่ยนแปลงไปอย่างรวดเร็ว ทั้งทางด้านวัฒนธรรม ค่านิยม สภาพสังคมและสิ่งแวดล้อม สิ่งเหล่านี้อาจเป็นปัจจัยที่มีส่วนทำให้เกิดพฤติกรรมและการดำรงชีวิตของคนในปัจจุบันเปลี่ยนไป ทั้งในด้านพฤติกรรมที่พึงประสงค์และพฤติกรรมที่ไม่พึงประสงค์ ถ้าเป็นพฤติกรรมที่พึงประสงค์นั้นก็ส่งผลดีกับตนเองและคนรอบข้าง แต่ในทางกลับกันหากเป็นพฤติกรรมไม่พึงประสงค์ก็อาจส่งผลในด้านลบต่อตนเองและคนรอบข้างได้เช่นกัน โดยอาจจะมองย้อนลงไปเริ่มตั้งแต่จากตัวบุคคล ปัญหาภายในครอบครัว ปัญหาการคบเพื่อน ปัญหาการเรียน ตลอดจนปัญหาทางด้านสังคม สิ่งเหล่านี้อาจเป็นจุดกำเนิดที่อาจส่งผลในเรื่องของการมีพฤติกรรมรุนแรง ซึ่งเป็นสิ่งที่สังคมไทยกำลังเกิดขึ้นและมักพบเจอตามสื่อทั่วไป ไม่ว่าจะเป็น โทรทัศน์ หนังสือพิมพ์ ฯลฯ

เหตุการณ์ความรุนแรงที่เป็นข่าวกึกโครมในปัจจุบันมีส่วนที่เกิดจากเยาวชนเพิ่มมากขึ้นเรื่อยๆ โดยพบในวัยรุ่นอยู่ในช่วงอายุระหว่าง 13 - 18 ปี โดยการเกิดเหตุการณ์รุนแรงที่มาจากเยาวชนนั้นนับวันยิ่งจะทวีความรุนแรงมากขึ้นและขยายวงกว้างออกไปทุกขณะ หากลองมองให้ลึกลงไปก็จะมีสิ่งที่สะท้อนให้เห็นในปัญหาของความล้มเหลวในการพัฒนาประเทศอยู่มาก โดยอาจรวมถึงปัจจัยต่างๆ ที่ได้กล่าวมาข้างต้นทำให้การพัฒนาเยาวชนในชาติไม่ประสบความสำเร็จเท่าที่ควร

ดังนั้นจึงมีความต้องการที่จะจัดทำเครื่องมือที่ช่วยในการวิเคราะห์พฤติกรรม โดยพิจารณาองค์ประกอบในด้านพฤติกรรมของนักเรียน ครอบครัว เพื่อน และ โรงเรียน เพื่อนำข้อมูลที่ได้มาทำนายพฤติกรรมของนักเรียนว่ามีแนวโน้มที่อาจจะมีลักษณะหรือโอกาสที่มีพฤติกรรมความรุนแรงอยู่ในระดับใด และอาจจะนำมาช่วยเป็นแนวทางในการรับมือและปฏิบัติต่อนักเรียนที่มี

เอกสารแนบไว้จะมีพฤติกรรมความรุนแรง อีกทั้งยังสามารถเป็นเครื่องมือเพื่อช่วยให้ครูในการเฝ้าระวังไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ติดตามนักเรียนและพัฒนาเจตคติให้ดีขึ้น ตลอดจนและเป็นแนวทางในการปรับเปลี่ยนพฤติกรรม ซึ่งส่งผลต่อคุณภาพชีวิตที่ดีขึ้นและเป็นพลเมืองที่ดีของสังคมในอนาคต

1.2 วัตถุประสงค์

1. เพื่อทำนายพฤติกรรมของนักเรียนว่ามีพฤติกรรมความรุนแรงอยู่ในระดับใด
2. เพื่อเป็นเครื่องมือที่ช่วยครูในการสอดส่องดูแลและติดตามนักเรียนพฤติกรรมความรุนแรง

1.3 ขอบเขตของงาน

1. ใช้ในการจำแนกระดับพฤติกรรมความรุนแรงของนักเรียน
2. ใช้ในการวิเคราะห์ระดับพฤติกรรมความรุนแรงของนักเรียน
3. เป็นเครื่องมือต้นแบบในการทำนายพฤติกรรมความรุนแรงของนักเรียน

1.4 แนวทางการศึกษาและพัฒนา

1. ศึกษาจากข้อมูลกลุ่มพฤติกรรมของนักเรียน จากแบบสอบถาม
2. ศึกษาความต้องการจากกลุ่มครูผู้สอน
3. สร้างโมเดลต้นแบบเพื่อการทำนายระดับพฤติกรรม

1.5 ข้อตกลงเบื้องต้น

1. เนื่องจากการศึกษาอิสระที่จัดทำมีการต่อยอดมาจากงานวิจัยอื่นซึ่งมีข้อจำกัดหลายประการ
2. งานวิจัยที่นำมาทำมีวัตถุประสงค์เพื่อต้องการทราบถึงสาเหตุที่เด็กแสดงพฤติกรรมรุนแรงเป็นการแปรผลสถิติแบบพรรณนา
3. การพยากรณ์ครั้งนี้เป็นการวัดด้านพฤติกรรมเท่านั้นไม่ได้วัดด้านเจตคติ
4. เนื่องการสร้างโมเดลในการศึกษาอิสระครั้งนี้ผู้จัดทำได้ใช้กลุ่มตัวอย่างจากนักเรียน โดยกลุ่มนักเรียนดังกล่าวซึ่งได้มาจาก สำนักงานส่งเสริมสวัสดิภาพและพิทักษ์เด็ก เยาวชน ผู้ด้อยโอกาสและผู้สูงอายุ กระทรวงพัฒนาสังคมและความมั่นคงของมนุษย์ เพื่อมาทดสอบกับโมเดล ซึ่งได้ค่าความเชื่อมั่นในระดับหนึ่ง ทั้งนี้หากต้องการให้โมเดลมีความแม่นยำ และมีค่าความเชื่อมั่นมากยิ่งขึ้น ควรนำข้อมูลของผู้ที่เคยกระทำผิด เช่น ผู้ที่ถูกจองจำ อาชญากรมาสร้างโมเดลเพื่อใช้ในการตรวจสอบความถูกต้องแม่นยำของโมเดลมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. ช่วยในการวิเคราะห์ระดับพฤติกรรมความรุนแรงของนักเรียน
2. ช่วยในการจำแนกกลุ่มนักเรียนเพื่อให้สอดคล้องกับการเรียนการสอน
3. ช่วยในการพัฒนาพฤติกรรมนักเรียนได้
4. ช่วยลดโอกาสที่นักเรียนที่มีพฤติกรรมความรุนแรง

1.7 เครื่องมือที่ใช้ในการพัฒนาระบบ

1. โปรแกรมประยุกต์ Weka version 3.7
2. โปรแกรม Microsoft Excel 2010



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในหัวข้อนี้จะกล่าวถึงทฤษฎีพื้นฐานต่างๆ ที่เกี่ยวข้อง ได้แก่ ทฤษฎีพฤติกรรม พฤติกรรมความรุนแรง พื้นฐานการทำเหมืองข้อมูลต้นไม้การตัดสินใจ (Decision Tree) รวมทั้งสถาปัตยกรรมและระบบการทำงานของอัลกอริทึมที่ใช้ในการจำแนกกลุ่มข้อมูลซึ่งเนื้อหาทั้งหมดนี้จำเป็นสำหรับการศึกษา และพัฒนาโปรแกรมประยุกต์ต่อไป

2.1 ทฤษฎีที่เกี่ยวข้องกับวัยรุ่น

2.1.1 ความหมายของวัยรุ่น

วัยรุ่น คือวัยแห่งการเจริญเติบโต และมีการพัฒนาทั้งทางร่างกาย อารมณ์ สติปัญญาและสังคม เพื่อให้พร้อมไปสู่การเป็นผู้ใหญ่ที่มีวุฒิภาวะ ซึ่งในขณะที่เป็นวัยรุ่นที่ไม่ควรได้รับการปฏิบัติอย่างวัยเด็กแต่ก็ยังไม่พร้อมที่จะปฏิบัติอย่างในวัยผู้ใหญ่

2.1.2 ลักษณะของวัยรุ่น

1. เป็นช่วงของการเปลี่ยนวัยคือเด็กที่กำลังจะเป็นผู้ใหญ่ ซึ่งมักจะเกิดความสับสนว่าตนเป็นเด็กอยู่แต่จะทำในสิ่งที่ผู้ใหญ่ทำ ซึ่งเกิดความกังวล ไม่แน่ใจ เป็นช่วงวัยหัวเลี้ยวหัวต่อ
2. เป็นช่วงวัยที่การเปลี่ยนทางร่างกายอย่างเต็มที่มีการสนใจตนเองและร่างกายมากขึ้น มีความพึงพอใจในเรื่องของการคบเพื่อนต่างเพศต่างวัย ซึ่งต้องอาศัยเวลาในการปรับตัว
3. เป็นช่วงวัยที่จะต้องเตรียมความพร้อมสำหรับการประกอบอาชีพในอนาคต ซึ่งการเลือกอาชีพเป็นสิ่งสำคัญต่อความต้องการของเด็ก ซึ่งสัมพันธ์กับ จิตใจ อารมณ์ รวมถึงบุคคลที่เกี่ยวข้อง เด็กอาจจะเกิดความสับสนหากอยู่ภายใต้การบีบบังคับ ข้อจำกัดต่างๆ ซึ่งอาจจะทำให้เด็กเกิดความสับสนในชีวิตได้

2.1.3 การแบ่งช่วงอายุของวัยรุ่น

1. วัยรุ่นตอนต้นอยู่ในช่วงอายุ 13 - 15 ปี ส่วนใหญ่จะเน้นในช่วงการเปลี่ยนแปลงทางด้านร่างกายอย่างชัดเจนทั้งในเด็กชายและเด็กหญิง ซึ่งในเพศหญิงก็จะมีประจำเดือนครั้งแรก ส่วนเด็กผู้ชายก็มีการเปลี่ยนแปลงเช่นกันแต่ไม่บ่งชัด
2. วัยรุ่นตอนกลาง อยู่ในช่วงอายุ 15 - 18 ปี เป็นช่วงที่มีการเปลี่ยนด้านร่างกาย จิตใจ มีการเปลี่ยนแปลงค่อยเป็นค่อยไปและมากขึ้นเรื่อยๆ ส่วนในด้านจิตใจเป็นผลจากการเปลี่ยนแปลงของร่างกาย ซึ่งการเปลี่ยนทางด้านจิตใจนั้นแตกต่างกันไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. วัยรุ่นตอนปลาย อยู่ในช่วงอายุ 18 - 21 ปี เป็นช่วงที่มีการเปลี่ยนแปลงด้านวุฒิภาวะอย่างสมบูรณ์เพื่อเตรียมที่จะเป็นผู้ใหญ่ในอนาคต เริ่มมีความคิดเป็นของตนเอง อิสระ และมักจะขัดแย้งกับผู้ใหญ่มากขึ้น

2.2 ทฤษฎีและแนวคิดเกี่ยวกับพฤติกรรม

2.2.1 ความหมายของพฤติกรรม

ความหมายของคำว่า พฤติกรรม ตามราชบัณฑิตยสถาน หมายถึง การกระทำหรืออาการที่แสดงออกทางกล้ามเนื้อ ความคิด และความรู้สึกที่ตอบสนองต่อสิ่งเร้า

ซิมบาร์โด (Zimbardo) กล่าวว่า พฤติกรรม (Behavior) หมายถึง ผลที่เกิดจากการแสดงปฏิกิริยาตอบสนองต่อสิ่งเร้าในสถานการณ์ต่างๆ พฤติกรรมหรือการกระทำของมนุษย์ เมื่อเกิดขึ้นจะสังเกตได้ชัดเจน เช่น การกระทำทางกาย ได้แก่ การเดิน ยืน นอน และการกิน การกระทำทางวาจา ได้แก่ การพูด การสนทนา

นักจิตวิทยา เชื่อว่าพฤติกรรมที่เกิดขึ้นต่างๆ เป็นผลที่เกิดมาจากการทำปฏิกิริยาของมนุษย์หรืออินทรีย์ (Organise) กับสิ่งแวดล้อม (Environment)

สรุปได้ว่า พฤติกรรมนั้นหมายถึง การกระทำของมนุษย์ อันได้แก่ การกระทำทางด้านกาย วาจา ใจ ที่มีปฏิกิริยาตอบสนองจากสิ่งเร้า

2.2.2 ประเภทของพฤติกรรม

1. พฤติกรรมที่มีมาตั้งแต่กำเนิด จะเกิดขึ้น โดยที่ไม่สามารถทราบล่วงหน้า และเป็นไปตามความรู้สึก เช่น ปฏิกิริยาสะท้อนกลับ (Reflex Action) ยกตัวอย่าง การที่ปลายนิ้วมือไปแตะถูกวัตถุที่มีความร้อน ปฏิกิริยาสะท้อนกลับคือ การดึงมือเข้าหาตัวเพื่อหนีออกจากวัตถุนั้นๆ
2. พฤติกรรมที่ได้มาภายหลังเกิดจากพฤติกรรมที่บุคคลได้มาภายหลัง ซึ่งมาจากการเรียนหรือฝึกฝนด้วยตนเองได้รับอิทธิพลมาจากกลุ่มหรือของสังคมตลอดจนสภาพแวดล้อม

2.2.3 รูปแบบพฤติกรรมของมนุษย์

รูปแบบพฤติกรรมของมนุษย์แบ่งได้เป็น 2 อย่างคือ

1. พฤติกรรมภายนอก (Overt Behavior) เป็นพฤติกรรมที่บุคคลนั้นแสดงออกมา เพื่อให้ผู้อื่นสามารถมองเห็นและสังเกตได้เช่นการหัวเราะการเคลื่อนไหวต่างๆ การพูดและการสื่อสาร เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. พฤติกรรมภายใน (Covert Behavior) เป็นพฤติกรรมที่บุคคลนั้นแสดงออกมาแล้วแต่บุคคลอื่นไม่สามารถสังเกตเห็นได้โดยตรงจนกว่าบุคคลนั้นจะเป็นผู้บอกหรือแสดงบางอย่างเพื่อให้บุคคลอื่นรับรู้ได้เช่นความคิดอารมณ์ความรู้สึก และการรับรู้เป็นต้น ซิกมันด์ فروยด์ ได้วิเคราะห์จิตมนุษย์ออกเป็น 3 ส่วนคือ Id Ego และ Super Ego ส่วนทั้งสามนี้ประกอบเป็นโครงสร้างทางจิต

อิด คือ สัญชาตญาณของดิบของมนุษย์ จะเป็นสิ่งที่คอยขับเคลื่อนให้มนุษย์แสดงพฤติกรรมต่างๆ อิดเป็นจิตใต้สำนึก แรงผลักดันนี้อยู่โดยที่บุคคลนั้นไม่รู้สึกรู้ตัว เป็นแรงผลักดันไร้สำนึก อิดจะผลักดันให้จิตอีกส่วนหนึ่งซึ่งเป็นส่วนใหญ่และส่วนที่รู้ตัว ที่เรียกว่า อีโก้

อีโก้ คือ กระทำในสิ่งต่างๆ ตามที่อิดมีความต้องการ เป็นตัวตอบสนองในความต้องการของอิด ส่วนของจิตที่ทำหน้าที่ควบคุมอีโก้ คือ ซุปเปอร์อีโก้

ซูปเปอร์อีโก้ เป็นความรู้สึกผิดชอบชั่วดี เป็นสิ่งที่เกิดจากการอบรมสั่งสอน การได้รับศึกษา ทำให้อิดและอีโก้มีพฤติกรรมอยู่ในทางที่ถูกที่ชอบแรงผลักดันจากอิดจะทำให้เกิดความตึงเครียด อีโก้จึงจำเป็นที่จะต้องพยายามตอบสนองความต้องการของอิดเพื่อลดความตึงเครียด แต่ความต้องการของอิดบางอย่างอีโก้ก็ไม่สามารถทำได้ทุกอย่าง เพราะจะขัดกับความรู้สึกผิดชอบชั่วดีในซูปเปอร์อีโก้ จึงทำให้เกิดความเครียดและความวิตกกังวลในใจเกิดขึ้น ซึ่งความวิตกกังวลที่เกิดขึ้นจะสามารถเปลี่ยนแปลงเป็นแรงขับเคลื่อนพฤติกรรมอีกทางหนึ่ง เพื่อให้เกิดกลไกการปกป้องตนเองให้รอด อีโก้จึงต้องพัฒนาพฤติกรรมป้องกันที่เรียกว่า “กลไกป้องกัน” ซึ่งเป็นไปโดยไม่รู้สึกรู้ตัว ได้แก่

1. การเก็บกด (Repression) เป็นกลไกในการป้องกันจากความผิดหวัง วิตกกังวลที่เกิดจากความขัดแย้งความคับข้องใจ การเก็บกดเป็นการทำให้สภาวะที่ก่อให้เกิดความรู้สึกผิดหวังหรือบางสิ่งบางอย่างที่ตนมีความไม่พอใจถูกเก็บไว้ในจิตไร้สำนึกเพื่อไม่ให้จดจำความรู้สึกนั้น

2. การแสดงปฏิกิริยาตรงกันข้ามกับความรู้สึกที่แท้จริง (Reaction Formation) คือ การเก็บกดความรู้สึกที่แท้จริงของตนเองซึ่งสังคมไม่ยอมรับแล้วแสดงออกถึงพฤติกรรมที่ตรงข้ามกับความรู้สึกของตนเพื่อไม่ให้ถูกตำหนิรักษาศักดิ์ศรีของตนเอง

3. การหาเหตุผลเข้าข้างตนเอง (Rationalization) คือการหาคำอธิบายหรือเหตุผลมาอ้างอิงการกระทำของตนเองเพื่อให้เป็นที่ยอมรับการหาเหตุผลต่างๆ มาอ้างอิงนี้ก็เพื่อลดความเครียดความวิตกกังวลความไม่สบายใจและเป็นการรักษาหน้าและศักดิ์ศรีของตน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. การแสดงพฤติกรรมถอยหลัง (Regression) คือ การแสดงพฤติกรรมย้อนกลับ ไปมีลักษณะเหมือนพฤติกรรมที่เคยทำมาในอดีต โดยเฉพาะพฤติกรรมในวัยเด็กเกิดขึ้นโดยไม่ได้ตั้งใจและไม่รู้ตัวพฤติกรรมนี้จะเกิดขึ้นเมื่อประสบความไม่มั่นคงทางจิตใจมีความคับข้องใจ

5. การปฏิเสธความจริง (Denial) คือ การปฏิเสธความจริงที่มีผลกระทบกระเทือน ต่อจิตใจอย่างรุนแรงจนไม่สามารถยอมรับได้จึงทำให้บุคคลนั้นไม่ยอมรับรู้สิ่งที่เกิดขึ้นกับตนเพื่อ เป็นการหลอกตัวเองให้คลายทุกข์การใช้กลไกป้องกันตนเองชนิดนี้จะพบได้บ่อยเมื่อบุคคลสูญเสีย คนรักการปฏิเสธ โรคร้ายแรงบางชนิดหรือกระทำความผิดบางอย่าง

6. การระบายไปที่อื่น (Displacement) คือ การระบายความคับข้องใจหรืออารมณ์ โกรธที่มีอยู่ขณะนั้น ไปสู่คนหรือสิ่งของที่ไม่ได้เป็นต้นเหตุของความโกรธหรือความคับข้องใจการใช้กลไกชนิดนี้ทำเพื่อเป็นการระบายและทำให้ความรู้สึกดีขึ้น โดยไม่มีผลกระทบต่อตนเองและสังคม

7. การโทษผู้อื่น (Projection) คือ การโทษผู้อื่นเป็นกลไกป้องกันตนเองที่นิยมใช้กันมากเกิดจากการที่บุคคลเกิดความผิดพลาดจากสิ่งที่ตนกระทำและทำให้เกิดความรู้สึกผิดแต่ เพื่อให้ความรู้สึกผิดนั้นลดน้อยลงจึงซัดทอดความผิดไปให้ผู้อื่น

8. การชดเชย (Compensation) คือ กลไกป้องกันตนเองที่บุคคลนำมาใช้เมื่อพบว่าตนมีข้อบกพร่องบางอย่างอยู่ในตัวเองและทำให้เกิดเป็นปมด้อยจึงพยายามสร้างปมเด่นมาชดเชยปมด้อย

9. การถอยหนี (Withdrawal) คือ กลไกป้องกันตนเองที่ใช้ลดความกลัวความวิตกกังวลและความเครียด โดยการหลีกเลี่ยงออกจากสถานการณ์ตัวบุคคลและสิ่งแวดล้อมที่ก่อให้เกิดความรู้สึกดังกล่าวการใช้กลไกป้องกันตนเองชนิดนี้ก่อให้เกิดผลเสียมากกว่าผลดีเนื่องจากบุคคลจะไม่สามารถแก้ไขปัญหาหรือสถานการณ์ในแบบเผชิญหน้า

2.2.4 พฤติกรรมที่เกิดขึ้นจากแรงขับเคลื่อนภายในตัวของมนุษย์และสิ่งแวดล้อม

สกินเนอร์ (Skinner) ให้แนวคิดว่าพฤติกรรมมนุษย์ถูกควบคุมโดยทฤษฎีเงื่อนไข การเสริมแรง และเงื่อนไขการลงโทษ ด้วยเหตุผลนี้ มนุษย์จึงไม่มีเสรีภาพแต่ประการใด สกินเนอร์ชี้ให้เห็นว่าผลการกระทำของคนเรามีอยู่ 2 ทางคือ ผลของการกระทำที่ทำให้เกิดความพอใจ ซึ่งจะทำหน้าที่เป็นการเสริมแรงให้กับการกระทำนี้ดำเนินต่อไป และผลการกระทำที่ทำให้

เกิดความไม่พอใจ ซึ่งจะเป็นต้นเหตุที่ทำให้คนหยุดการกระทำหรือหยุดอันจะนำมาซึ่งผลการกระทำเช่นนี้ในอนาคต

พฤติกรรมที่ทำให้เกิดความพึงพอใจ ยกตัวอย่างเช่น พฤติกรรมที่ทำแล้วได้รับการชื่นชมยินดีได้รับตำแหน่ง เลื่อนตำแหน่ง ได้รับเงิน ได้รับการยกย่องสรรเสริญ ฯลฯ ซึ่งสามารถจะมีโอกาสเกิดขึ้นในครั้งหน้า และในขณะที่พฤติกรรมที่ทำให้เกิดความรู้สึกไม่พึงพอใจ คือ ทำแล้วถูกตำหนิติเตียน เสียตำแหน่ง เสียเงิน ถูกหมิ่นเหยียดหยาม เป็นต้น ก็จะหยุดหรืออาจจะมีโอกาสเกิดขึ้นได้น้อยในครั้งหน้า เพราะฉะนั้นการเกิดพฤติกรรมของคนต่างๆ ไปมักถูกควบคุมโดยเงื่อนไขของผลการกระทำทั้ง 2 อย่าง

2.3 พฤติกรรมความรุนแรง

2.3.1 ความหมายของพฤติกรรมความรุนแรง

ในปัจจุบันเยาวชนสังคมไทยได้มีพฤติกรรมได้เปลี่ยนแปลงไปมากโดยเฉพาะด้านความรุนแรง และทวีความรุนแรงขึ้นเรื่อยๆ ซึ่งมักจะปรากฏให้เห็น โดยทั่วไปไม่ว่าจะเป็นหนังสือพิมพ์ โทรทัศน์ หรือแม้กระทั่งตามสื่อโซเชียลมีเดียต่างๆ เช่น เด็กนักเรียนยกพวกตีกัน ทำร้ายร่างกาย กลุ่มเด็กวัยรุ่นรุมข่มขืนเด็กผู้หญิง หญิงสาวรุมตบตีกันเพื่อแย่งผู้ชาย ทำร้ายด้วยอาวุธร้ายแรง การฆ่าตัวตายความรุนแรงในเด็กและเยาวชนวัยรุ่นเป็นช่วงวัยของการเปลี่ยนแปลงในด้านต่างๆ ทั้งทางร่างกาย จิตใจ อารมณ์ การปรับตัวในการเข้าสังคม

เด็กวัยรุ่นเรียนรู้ชีวิตจากชีวิตจริงที่เขาดำเนินอยู่ ขณะเดียวกันเขาก็เรียนรู้หรือซึมซับความรุนแรงจากชีวิตจริงที่เขาได้เห็น ได้สัมผัสจากคน ใกล้ชิด ครอบครัว ชุมชน สังคม โรงเรียน สิ่งแวดล้อมตลอดจนจากสื่อต่างๆ ที่มีผลกระทบอย่างมากต่อเด็กเล็กๆ และเด็กวัยรุ่นในสังคมปัจจุบัน

พฤติกรรมการใช้ความรุนแรงที่มีต่อความรุนแรงระหว่างบุคคลได้ถูกให้คำจำกัดความว่าเป็นพฤติกรรมของบุคคลที่ ต่อสู้กับบุคคลอื่น ซึ่งตั้งใจหรือพยายามที่จะคุกคาม หรือทำให้ร่างกายได้รับบาดเจ็บทุกข์ทรมานอย่างแท้จริง พฤติกรรมเหล่านี้ถูกรวมอยู่ในคำจำกัดความที่กว้างกว่าคำจำกัดความของความก้าวร้าว

ความรุนแรง พจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2542 ให้คำจำกัดความไว้ว่า

ความ หมายถึง [ความ] น. เรื่อง เช่น เนื้อความ เกิดความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รุนแรง หมายถึง หนักมาก แรงมาก เกินปกติ เช่น ดูอย่างรุนแรง คัดค้านอย่างรุนแรง
ความคิดเห็นรุนแรง

ความรุนแรง หมายถึง การกระทำใดๆที่ส่งผลกระทบต่อบุคคลอื่น ไม่ว่าจะเป็นการ
กระทำ หรือคำพูด ที่มาจากเจตนาทำให้ผู้อื่น ได้รับความบาดเจ็บ เกิดความรู้สึกละอายใจ กระทบกระเทือน
ตั้งแต่เล็กน้อยถึงขั้นเสียชีวิต โดยพฤติกรรมความรุนแรงดังกล่าว อาจมาจากปัจจัยหลายอย่างเช่น
โอกาส สิ่งแวดล้อม ความเชื่อ เพศ ฮอว์โมน ช่วงอายุ วุฒิภาวะ ฯลฯ โดยความรุนแรงนี้อาจเป็น
ผู้กระทำและอาจเป็นผู้ถูกกระทำได้

2.3.2 ประเภทของพฤติกรรมความรุนแรง

1. ความรุนแรงในวัยรุ่นจากข่าวที่พบในปัจจุบันมีการแสดงออกในด้านความ
รุนแรงที่เกิดจากการกระทำในวัยรุ่นเป็นจำนวนมากขึ้นเรื่อยๆ เนื่องจากเป็นวัยที่มีความคิดคะนอง
มักชอบทำในสิ่งที่ท้าทายต้องการอยากทดลองหรือค้นพบในสิ่งที่แปลกใหม่และค่านึงแค่เพียงว่าผล
ที่จะเกิดขึ้นแต่บุคคลอื่นไม่เกี่ยวข้องกับตนเอง จึงไม่คิดให้รอบคอบว่าจะมีผลเสียอย่างไรซึ่งจะเป็น
พฤติกรรมเสี่ยงที่จะทำให้เกิดแสดงพฤติกรรมที่เป็นปัญหาต่างๆตามมา

2. ความรุนแรงภายในครอบครัว เป็นการแสดงออกถึงความรุนแรงที่เกิดจากคน
ในครอบครัว เป็นการกระทำทารุณ ทางร่างกาย จิตใจ ทางเพศ การปล่อยปละละเลยไม่เลี้ยงดู

3. ความรุนแรงต่อตนเอง เป็นความรุนแรงโดยตรงต่อตนเอง เช่น การทำร้าย
ร่างกายตนเอง การฆ่าตัวตาย

2.3.3 สาเหตุของพฤติกรรมความรุนแรงของวัยรุ่น

ข้อมูลจากองค์การอนามัยโลกได้ให้รายละเอียดดังนี้

1. ปัจจัยส่วนบุคคล เกี่ยวข้องกับทางร่างกายพฤติกรรม การแสดงออก โดย
ธรรมชาติของวัยรุ่นเป็นวัยที่มีความใจร้อน มุทะลุ ก้าวร้าว ขาดการควบคุมของพฤติกรรม
การแสดงออกที่เหมาะสม ขาดการให้ความรัก ดูแลเอาใจใส่และมีประสบการณ์ในการทำร้ายผู้อื่นมี
ความเชื่อแบบผิดๆมีประสบการณ์ถูกทำโทษทางกายหรือเคยพบเห็นความรุนแรง

2. ปัจจัยครอบครัวและคนใกล้ชิด สิ่งแวดล้อมทางครอบครัวเป็นปัจจัยที่สำคัญ
ในการพัฒนาพฤติกรรมความรุนแรงของวัยรุ่น ได้แก่ บิดามารดาผู้ปกครองขาดการดูแลเอาใจใส่
ควบคุมตั้งแต่เด็กและใช้การลงโทษเด็กด้วยความรุนแรงหยาบคายบิดามารดาที่มีความขัดแย้งกัน

ครอบครัวขาดความสัมพันธ์ความผูกพันกัน มักมีการทำร้ายกันในบ้าน มารดามีบุตรเมื่ออายุน้อย มีประสบการณ์ที่บิดามารดาแตกแยกหรือหย่าร้างกันตั้งแต่วัยเด็กคบเพื่อนไม่ดี

2.3.4 อิทธิพลที่มีผลต่อการแสดงออกพฤติกรรมความรุนแรง

1. อิทธิพลจากการคบเพื่อนมีส่วนสำคัญในการแสดงออกพฤติกรรมความรุนแรง เด็กจะให้ความสำคัญกับการคบเพื่อนในช่วงอายุ 13-18 ปี มักจะมีการจับกลุ่มกันสร้างความสัมพันธ์และความผูกพันได้เป็นอย่างมากในช่วงวัยนี้ และจะเริ่มลดความสนใจและใส่ใจบุคคลต่างวัยไม่ว่าจะเป็นผู้ใหญ่หรือเด็กกว่าตนเอง และเริ่มมีการตั้งกลุ่มรวมกลุ่มเพราะมีการร่วมทุกข์ร่วมสุขกันในกลุ่มและช่วยกันแก้ปัญหาตนเองซึ่งมองว่าดีกว่าปรึกษาจากคนต่างวัย จึงกล่าวได้ว่าการคบเพื่อนนั้นมีทั้งทั้งคุณและโทษ หากแต่ว่าเด็กกลุ่มของเพื่อนนั้นช่วยกันแนะนำทำในสิ่งที่มีประโยชน์และสร้างสรรค์ก็จะช่วยส่งเสริมให้เด็กเติบโตเป็นผู้ใหญ่ที่ดีในสังคม แต่หากเพื่อนชักนำไปในทางที่ไม่ดีหรือเสื่อมถอย เช่น เกเร ทะเลาะวิวาท ดิตยาเสพติด ฯลฯ อาจจะกลายเป็นปัญหาอาชญากรรมในปัจจุบัน

2. อิทธิพลจากสื่อเปรียบเสมือนดาบสองคม สื่อในปัจจุบันได้ความข่าวสารความรู้มากมาย แต่ในขณะเดียวกันก็แฝงไปด้วยเนื้อหาที่รุนแรง อาจจะได้จากภาพ การแสดง ความรุนแรงต่างๆ จากข่าวอาชญากรรม ความรุนแรงทางเพศ สื่อโทรทัศน์ละคร ซึ่งอาจมีเนื้อหาที่ไม่เหมาะสมทางการแสดงออกหรือการใช้ภาษา ทำให้เด็กนั้นอาจจดจำและนำไปเลียนแบบได้ เนื่องจากเด็กในวัยนี้ยังไม่สามารถที่แยกแยะในการรับสื่อต่างๆ ที่เข้ามาในชีวิตของตนเองได้ อีกทั้งอาจจะทำพฤติกรรมเลียนแบบบุคคลที่ตนพึงพอใจ จึงทำให้เกิดความสับสนในตนเองและสร้างกลไกปกป้องตนเอง หรือการแสดงออกแปลกๆ เพื่อฝันในเรื่องตนเองสนใจหรือกำลังเลียนแบบอยู่ โดยไม่รู้ตัวบ้างไม่รู้ตัวบ้าง ซึ่งในบางกรณีอาจจะทำให้ตกอยู่ในภาวะเครียด เกิดการพฤติกรรมผิดในด้านต่างๆ เช่น การประพฤติดิตยาเสพติด

3. อิทธิพลจากการเล่นเกมออนไลน์หรือการเล่นเกมที่มีความรุนแรงส่งผลโดยตรงกับเด็ก หากจะหาประโยชน์จากการเล่นเกมนั้นก็ยังมี เช่น เป็นการพักผ่อน การใช้ทักษะด้านความคิด การแก้ไขสถานการณ์เฉพาะหน้า แต่หากเล่นเกมเกินวัตถุประสงค์ผลที่ตามมาก็คือเสียเวลาโดยเปล่าประโยชน์ สุขภาพเสื่อมโทรมเนื่องจากหมกมุ่น ผลการเรียนตกต่ำ หรือพักผ่อนไม่เพียงพอ สัมพันธภาพในครอบครัว เพื่อน หรือสังคมลดลง ยิ่งโดยเฉพาะการเล่นเกมที่เนื้อหาความรุนแรงแล้วยังส่งผลให้เด็กมีความเสี่ยงที่จะมีพฤติกรรมก้าวร้าวภายหลังจากการเล่นเกม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เนื่องจากการบ่มเพาะของพฤติกรรมความรุนแรงในขณะที่เล่นเกม ต้องการเอาชนะ ทำร้าย เพราะเด็กยังไม่สามารถแยกแยะระหว่างโลกของเกมและโลกแห่งความเป็นจริงได้ ดังนั้น ผู้ปกครองจะต้องมีการดูแลอย่างใกล้ชิด

2.3.5 การจำแนกความรุนแรงออกเป็น 3 ประเภท

1. การทำร้ายทางกายเช่นเตะทุบตีทำร้ายอวัยวะต่างๆหรือจี้ ไฟจี้ น้ำร้อนราดเตา ริดนาบล่ามโซ่เอาไม้แขวนเสื้อตีจนผู้ถูกกระทำแขนขาหักฟกช้ำหรือพินหักเป็นต้น
2. การทำร้ายทางเพศ เช่นการล่วงละเมิดทางเพศ ทำอนาจารข่มขืนถ่ายภาพโป๊เอา
3. การทำร้ายทางจิตใจทางวาจาเช่นบังคับขู่เข็ญคำทอบังคับเหยียดหยามบังคับได้ ออกจากบ้านทอดทิ้ง

2.3.6 ระดับของความรุนแรง

1. ระดับที่ไม่รุนแรงมาก เป็นเหตุการณ์ไม่ใหญ่และไม่อันตรายมากเป็นอยู่ไม่นาน เช่นบาดแผลอยู่ในเนื้อเยื่อหรือเฉพาะที่ผิวๆตลอดจนรอยช้ำรอยฟกช้ำเล็กๆแผลไฟไหม้ไม่มากรอยเขียวช้ำ
2. ระดับที่รุนแรงปานกลาง เป็นการทารุณรุนแรงขึ้นแต่ยังไม่เป็นการคุมคามถึงชีวิตหรือมีแนวโน้มที่จะให้ผลไปอีกนาน
3. ระดับที่รุนแรงมาก การกระทำอย่างต่อเนื่องหรือมีอันตรายต่อร่างกายหรือจิตใจอย่างรุนแรง
4. ระดับปางตายเป็นอันตรายที่เป็นอยู่ยาวนานหรือรุนแรงต่อร่างกายหรือต่อจิตวิทยา ซึ่งอาจเป็นผลให้ลูกคามชีวิตบาดแผลที่เป็นอยู่ยาวนานปางตายหรือเกือบตาย

2.3.7 แนวทางป้องกันพฤติกรรมความรุนแรง

ปัญหาของความรุนแรงในสังคม เป็นปัญหาที่มีมาอย่างยาวนานในสังคม สร้างความเสียหายให้กับชีวิต ทรัพย์สิน ผลกระทบต่อสภาพจิตใจของผู้ถูกกระทำ และผู้ใกล้ชิดของคนรอบคร้ว ซึ่งไม่สามารถประเมินค่าได้จากความสูญเสียที่ตามมา โดยที่ผ่านมามีหลายๆหน่วยงาน ได้ทำการร่วมรณรงค์ต่างๆ อาทิเช่น การต่อต้านความรุนแรงต่อสตรี และเด็ก แต่ในความเป็นจริง ความรุนแรงที่เกิดขึ้นในสังคมก็ยังไม่ได้อัตราที่ลดลงไป ซึ่งแนวทางในการป้องกันปัญหาความรุนแรงมีดังต่อไปนี้

1. การป้องกันเกี่ยวกับตัวบุคคล มีจุดประสงค์มีด้วยกันสองประการ คือ ประการแรกจะมีการต้องสนับสนุนและสร้างเจตคติและพฤติกรรมที่ดีให้แก่เด็กและวัยรุ่นเป็นอันดับแรก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อเป็นการป้องกันก่อนที่จะเติบโตเป็นผู้ใหญ่ที่มีพฤติกรรมการใช้ความรุนแรง ประการที่สอง เปลี่ยนเจตคติและพฤติกรรมในบุคคลที่ก่อความรุนแรงหรือมีความเสี่ยงที่จะใช้ความรุนแรง

2. การป้องกันในด้านความสัมพันธ์บุคคล เป็นการป้องกันที่อยู่ในระดับครอบครัว คือ ส่งเสริมสัมพันธ์ของคนครอบครัว รวมทั้งและแก้ไขสัมพันธ์ที่ไม่ดีหรือไม่ถูกต้องที่เกิดขึ้นภายในครอบครัว อาจเป็นความสัมพันธ์ระหว่างผู้ถูกระทำและผู้กระทำ เช่น ความขัดแย้งในชีวิตของการแต่งงาน การขาดความใกล้ชิดระหว่างบิดามารดาผู้ปกครองและเด็ก หรือความสัมพันธ์ระหว่างเพื่อนที่มีอิทธิพลที่สามารถส่งถึงบุคคล

3. การป้องกันในด้านชุมชนเป็นการสร้างการรับรู้ตื่นตัวในเรื่องของความรุนแรงให้แก่ชุมชน ท้องถิ่น และสังคม จะต้องมีการรณรงค์กระตุ้นให้ชุมชนเข้ามาส่วนร่วม บทบาทในการแก้ไขปัญหา และส่งเสริมให้มีระบบติดตามดูแลเหยื่อที่ได้รับความรุนแรง

4. การป้องกันในสังคม คือการป้องกันที่มีเป้าหมายที่มีปัจจัยหลักด้านวัฒนธรรม สังคม และเศรษฐกิจของแต่ละประเทศ หากจะป้องกันก็เน้นไปแนวทางของการเปลี่ยนแปลงกฎหมาย นโยบายทางภาครัฐ สังคมโดยรวม และวัฒนธรรมความเชื่อ

2.4 ทฤษฎีการทำเหมืองข้อมูล

Information คือ ข้อมูลดิบหรือข้อเท็จจริงที่เราสนใจแล้วนำไปเข้าสู่กระบวนการเพื่อให้ได้ “สารสนเทศ” ที่เราต้องการ แล้วนำเอา สารสนเทศนั้นไปใช้ประโยชน์ ซึ่งองค์กรใดสามารถที่จะเก็บข้อมูลดิบต่างๆ ให้ได้มากที่สุดแล้วนำเอาข้อมูลเหล่านั้นมา สกัดเป็น สารสนเทศได้ตามต้องการอาจทำให้องค์กรนั้น ได้เปรียบในแง่ของการผลิตภัณฑ์ให้องค์กรนั้นมีโอกาสประสบความสำเร็จมากกว่าองค์กรอื่น

การเก็บรวบรวมข้อมูลไว้ในแหล่งเดียวกัน โดยที่อาจไม่สนใจลักษณะของข้อมูล แล้วเมื่อต้องการใช้งานข้อมูล จึงทำการดึงเอาข้อมูลนั้นๆมาผ่านกระบวนการใดๆ เพื่อให้ได้ สารสนเทศที่เราต้องการ คือ การทำเหมืองข้อมูล (Data Mining) ซึ่ง เหมืองข้อมูลถือเป็น “กุญแจหลัก” สำคัญขององค์กร นั้นๆ โดย การทำเหมืองข้อมูล เป็นกระบวนการของการวิเคราะห์ข้อมูลจากมุมมองที่แตกต่างกันและสรุปเป็นข้อมูลที่เป็นประโยชน์ เช่นข้อมูลที่สามารถนำมาใช้เพื่อวิเคราะห์ผู้บริโภคในการซื้อของหรือสินค้าในห้างสรรพสินค้า ว่ามักจะซื้อสินค้าใดพร้อมๆกันบ้าง เพื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ห้างสรรพสินค้านั้นจะได้นำเอาสารสนเทศนั้นไปวิเคราะห์แล้วนำมาใช้เป็นเครื่องมือในการช่วย การขาย อาทิเช่น จัดโปรโมชัน ชื่อสินค้า A และ สินค้า B พร้อมกันจะได้ของแถมใด เป็นต้น

2.4.1 ความหมายของเหมืองข้อมูล (Data Mining)

เป็นกระบวนการในการแตก Previously Unknown, Valid and actionable information จากฐานข้อมูลขนาดใหญ่ และนำสารสนเทศนั้นไปใช้ในการตัดสินใจทางธุรกิจ และ ลักษณะของ Information ที่ดึงออกมาจากแหล่งของข้อมูล หรือฐานข้อมูลขนาดใหญ่จะมีลักษณะ

- | | | |
|--------------|---------|-------------------------|
| - Unknown | หมายถึง | ไม่ทราบมาก่อน |
| - Valid | หมายถึง | ถูกต้อง |
| - Actionable | หมายถึง | เอาไปใช้ในทางปฏิบัติได้ |

ทั้ง 3 ส่วนนี้จะเป็นตัวบอกได้ว่าการทำ Data Mining สำเร็จมากหรือน้อยแค่ไหน

2.4.2 ขั้นตอนหลักในการทำเหมืองข้อมูล

1. การวิเคราะห์ความต้องการขององค์กร

หรือความต้องการของบุคคลนั้นๆเพื่อให้ได้ข้อสรุปที่ชัดเจน การวิเคราะห์ความต้องการของธุรกิจจะทำให้เกิดความเข้าใจถึงสถานะในปัจจุบันขององค์กรและทำให้ผู้บริหาร สามารถกำหนดเรื่องที่ต้องตัดสินใจได้ดียิ่งขึ้น

2. การวิเคราะห์ความต้องการข้อมูล

เนื่องจากคลังข้อมูลขนาดใหญ่ที่มีอยู่นั้นมีข้อมูลที่หลากหลายซึ่งมีทั้งข้อมูลที่จำเป็น และไม่จำเป็นต้องใช้ในการทำเหมืองข้อมูลกับข้อมูลอื่นซึ่งไม่เป็นที่ต้องการในขณะนี้

3. การแปลงข้อมูล

เมื่อกำหนดข้อมูลที่จะใช้ในการทำเหมืองได้แล้ว อาจต้องทำการแปลงหรือปรับเปลี่ยนข้อมูล ให้อยู่ในรูปแบบที่เหมาะสมกับวิธีการที่จะใช้ในการทำเหมืองข้อมูล

4. การดำเนินการทำเหมืองข้อมูล

เป็นขั้นตอนที่นำเอาวิธีการหรือเทคนิคการทำเหมืองข้อมูลตั้งแต่หนึ่งวิธีขึ้นไป มาทำการสกัดสาระสำคัญออกจากฐานข้อมูล

5. การแปลผล

จากผลลัพธ์ที่ผ่านกระบวนการทำเหมืองข้อมูลเป็นขั้นตอนที่ผู้บริหารหรือผู้ใช้งาน จะต้องแปลความหมายผลลัพธ์ที่ได้เพื่อนำไปพัฒนาองค์กร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.3 เทคนิคการทำเหมืองข้อมูล

1. กฎความสัมพันธ์ (Association Rule Discover)

กฎ แสดงความสัมพันธ์ของเหตุการณ์หรือสิ่งที่เกิดขึ้นพร้อมกัน สองชุดหรือมากกว่าสองชุดขึ้นไปแล้วนำมาวิเคราะห์ เช่น การวิเคราะห์ข้อมูลการขายสินค้าของร้านค้าแห่งหนึ่ง โดยเก็บข้อมูลจากระบบ ณ จุดขาย แล้วพิจารณาสินค้าที่ผู้ซื้อมักจะซื้อพร้อมกัน เช่น พบว่า ผู้บริโภคมักจะซื้อเบียร์ พร้อมกับ แครกเกอร์ ในวันศุกร์ ก็อาจจัด โปรโมชันหรือ จัดวางสินค้าให้สะดวกต่อการหยิบสินค้า เพื่อเพิ่มยอดขายให้กับทางร้านค้า

- การแบ่งกลุ่มข้อมูล (Data clustering)

เป็นการแบ่งข้อมูลที่มีลักษณะคล้ายกันออกเป็นกลุ่ม หรือ การแบ่งกลุ่มที่ต้องใช้ข้อมูลของทุกกลุ่มในการหาจุดเด่นของแต่ละกลุ่มออกมาให้เห็นอย่างชัดเจน เพื่อได้มาซึ่งลักษณะเฉพาะของกลุ่มนั้นๆ เช่น

- ภาพเสมือน (Visualization)

เป็นการนำเอาข้อมูลปัจจุบันมาทำเป็นแบบจำลองโดยใช้โปรแกรมในการสร้างสถานการณ์เพื่อทำนาย หรือคำนวณหาค่าความน่าจะเป็นในอนาคต

- การจำแนกประเภทข้อมูล (Data classification)

การจำแนกประเภทของข้อมูล เป็นการจัดแบ่งประเภทของข้อมูล โดยหาชุดต้นแบบหรือชุดของการทำงานที่อธิบายและแบ่งประเภทข้อมูล วัตถุประสงค์เพื่อให้สามารถใช้เป็นต้นแบบทำนายประเภทของวัตถุหรือข้อมูลที่ไม่มีการระบุประเภทหรือชนิดของข้อมูล ซึ่งต้นแบบสร้างจากการวิเคราะห์ชุดของข้อมูลฝึกสอน (Training Data) โดยอาจจะเป็นกลุ่มข้อมูลที่มีการระบุประเภทหรือกลุ่มเรียบร้อยแล้ว รูปแบบของต้นแบบแสดงได้หลายแบบเช่น Classification Rules Decision Trees และ Neural Networks เป็นต้น โดยตัวอย่างของการจำแนกประเภทข้อมูล เช่น หาความสัมพันธ์ระหว่างกลุ่มของผู้บริโภคที่กลุ่มแรกยังคงใช้บริการหรือยังคงจงรักภักดีกับบริษัท A และ อีกกลุ่มหนึ่งหันไปหาบริษัท B

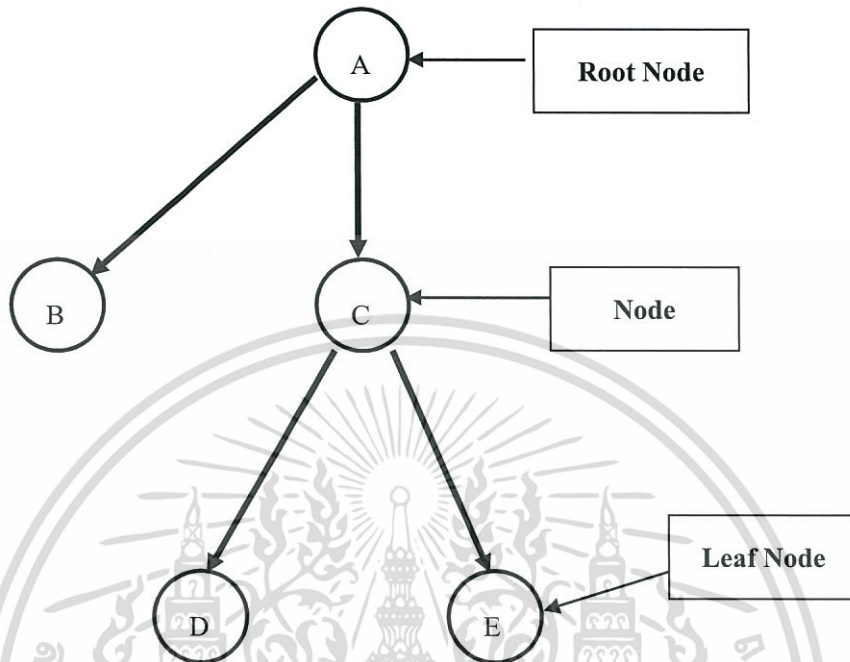
การศึกษานี้ผู้เขียนได้พิจารณานำเทคนิคที่เหมาะสมกับข้อมูลพื้นฐานที่ผู้เขียนได้นำมาประกอบการวิจัยนี้คือเทคนิคต้นไม้การตัดสินใจแก้ปัญหาเพื่อใช้ในการจำแนกแยกประเภทกลุ่มของผู้เรียนแล้วทำการทำนายพฤติกรรมของผู้เรียนนั้นๆ

2.4.4 ต้นไม้ตัดสินใจ (Decision Tree)

เป็นการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปของ Decision Tree ซึ่ง Decision Tree นั้นมีการทำงานแบบ Supervised Learning คือ สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จาก กลุ่มตัวอย่างของข้อมูลที่ได้กำหนดไว้ก่อนล่วงหน้า ที่เรียกว่า Training Set ได้อัตโนมัติ และสามารถพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่ได้ด้วย

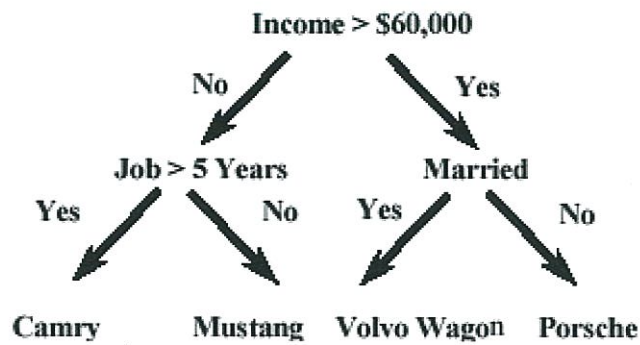


รูปที่ 2.1 แสดง โครงสร้างต้นไม้ตัดสินใจ (Decision Tree)

จากรูปที่ 2.1 เป็นโครงสร้างของเทคนิค ต้นไม้ตัดสินใจ โดยจะประกอบไปด้วย

- Root Node หมายถึง โหนดแม่ที่อยู่ส่วนบนสุดที่ทำหน้าที่เชื่อมโยงแตกกิ่งก้านกรณีออกเป็นส่วนๆ
- Node หมายถึง กิ่ง (Branch) แสดงค่าที่เป็นไปได้ของแต่ละ Attribute
- Leaf Node หมายถึง กลุ่มของผลลัพธ์ในการแยกแยะข้อมูล โดยสามารถแสดงส่วนประกอบของต้นไม้ตัดสินใจ

รูปแบบของ Tree จะประกอบด้วย Node แรกสุดที่เรียกว่า Root Node จาก Root Node ก็จะแตกออกเป็น Node ลูก และที่ Node ลูกก็จะมีลูกของตัวเองซึ่ง Node ที่ระดับสุดท้ายจะเรียกว่า Leaf Node



รูปที่ 2.2 ตัวอย่างต้นไม้ตัดสินใจ จำแนกกลุ่มลูกค้าใช้งานรถยนต์แต่ละยี่ห้อ

จากรูปที่ 2.2 เป็นตัวอย่างของการทำต้นไม้ตัดสินใจจำแนกกลุ่มลูกค้าที่เลือกซื้อรถยนต์ หรือจะซื้อรถยนต์ในแต่ละยี่ห้อ โดย เมื่อพิจารณาสามารถ แบ่งกลุ่ม คนจากรายได้ที่เข้ามาได้ สองกลุ่ม คือ กลุ่มคน ที่รายได้ เข้ามาตั้งแต่ \$60,000 ขึ้นไป และมีรายได้ น้อยกว่า \$60,000 ลงมา โดย กลุ่มคนที่มีรายได้ \$60,000 ขึ้นไป แบ่งเป็นสองกลุ่ม ที่เลือกซื้อ รถ 2 ยี่ห้อ โดยสรุปได้ว่า

- คนที่แต่งงานแล้ว เลือกที่จะซื้อรถยนต์ Volvo Wagon เนื่องจากจำนวนที่นั่งและสมรรถนะที่รองรับความสะดวกสบายสำหรับผู้โดยสารและมีความปลอดภัยสูงซึ่งเป็นปรกติของคนที่ต้องการให้การเดินทางแต่ละครั้งกับครอบครัวนั้นปลอดภัย และสะดวกสบายสูงสุดให้กับครอบครัว

- คนที่ยังไม่แต่งงานมักจะเลือกซื้อรถยนต์ Porsche ที่มีลักษณะ สปอร์ต ปราดเปรียว และต้องการความคล่องตัวและทำความเร็วได้สูงและโดย ปรกติ คนที่ยังไม่แต่งงานมักชอบที่จะอิสระจึงเลือกรถที่สามารถตอบสนองในเรื่องของความหรูหรา และเร็วโดยยังไม่คำนึงถึงจำนวนผู้โดยสารเท่าไรนักแต่ก็คำนึงถึงสมรรถนะของรถเป็นหลัก

กลุ่มคนที่มีรายได้น้อยกว่า 60,000 \$ ลงมา แบ่งเป็น สองกลุ่ม โดย

- คนที่ทำงานมาแล้ว 5 ปีขึ้นไป จะเลือกซื้อรถรุ่น Camry โดยที่ รถรุ่นนี้เป็นรถประเภทครอบครัว และมีพื้นที่ภายในค่อนข้างกว้าง

- คนที่ทำงาน ไม่ถึง 5 ปี มักจะเลือกซื้อรถยนต์รุ่น Mustang โดยรถรุ่นนี้เป็นรถแนวสปอร์ต ตอบสนองเรื่องความเร็วและความสวยงามและปราดเปรียว ตัวอย่างการจำแนกกลุ่มคนใช้งานรถยนต์ในแต่ละแบบ อาจสรุปได้ว่า คนที่แต่งงานแล้ว และอาจทำงานมาแล้วมากกว่า 5 ปีขึ้นไปมักมองหารถยนต์ที่มุ่งเน้นเรื่องของความกว้างขวาง มีจำนวนที่นั่งมากๆและมีเนื้อที่ภายในกว้างขวาง ไม่เน้นความรวดเร็ว และคำนึงถึงความสะดวกสบายเป็นหลักมากกว่า และคนที่เพิ่งเริ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทำงานและ อาจมีเงินเดือนมากกว่าและยังไม่ได้แต่งงานมักชอบอิสระ จึงชอบรถที่สามารถทำความเร็วและปราดเปรี้ยว มาใช้งาน

ขั้นตอนการทำต้นไม้ตัดสินใจจาก Data set

ตารางที่ 2.1 ตารางตัวอย่างข้อมูลการให้บริการ

Name	Tenure	Service	Attrition
A	>2.5	>3	Stay
B	>2.5	<3	Stay
C	<2.5	<3	Leave
D	<2.5	>3	Stay

จากตารางที่ 2.1 สามารถอธิบาย Attribute ในแต่ละช่องได้ว่า

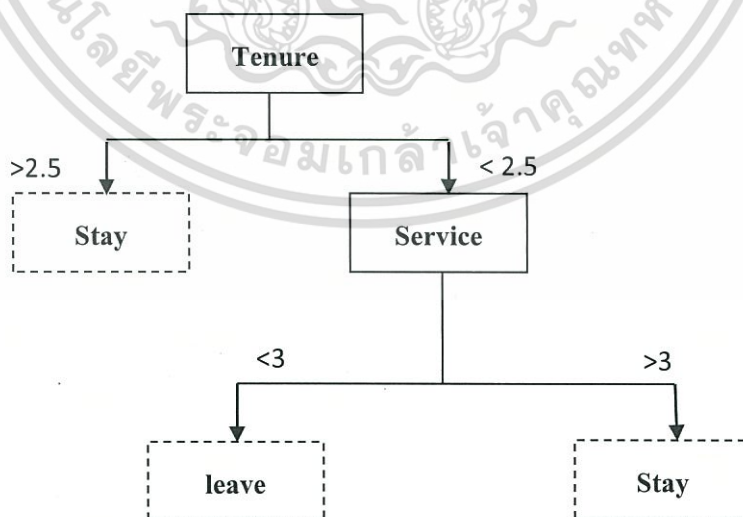
Name หมายถึง คุณลักษณะที่แทนรายชื่อลูกค้า

Tenure หมายถึง คุณลักษณะที่แทน วาระการดำรงตำแหน่ง

Service หมายถึง คุณลักษณะที่แทน จำนวนการให้บริการ

Attrition หมายถึง คุณลักษณะที่แทนสถานการให้บริการว่ายังคงรับการบริการจัดดีหรือไม่

โดยเมื่อทำการนำเอาค่าคุณลักษณะของตารางมาทำการเขียนให้อยู่ในรูปของต้นไม้ตัดสินใจได้ตามรูปที่ 2.3



รูปที่ 2.3 โครงสร้างต้นไม้ตัดสินใจ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. อัลกอริทึมในการสร้าง ต้นไม้ตัดสินใจ

อัลกอริทึมพื้นฐานในการสร้างต้นไม้ตัดสินใจนั้นมีด้วยกันหลายวิธีได้แก่

- CART (classification and regression tree) เป็นAlgorithm แรกๆในการสร้าง Tree
- CHAID (Chi-square Automatic Interaction Detection) มักจะใช้ในทางสถิติ
- ID 3 (Iterative Dichotomiser3) เป็นตัวที่จะใช้เป็นหลัก ง่าย และเป็นพื้นฐานของ

หลาย Algorithm C4.5, C5 พัฒนามาจาก ID3

2. แนวคิดการเรียนรู้ต้นไม้ตัดสินใจ

- หาAttributeที่สำคัญที่สุดที่มีอิทธิพลต่อการทำนายมากที่สุดมาเป็น Root Node
- จาก Root Node ให้พิจารณาค่าที่เป็นไปได้มีกี่ค่า และทำการแตกกิ่งออกมาตามค่าที่เป็นไปได้
- ตัวอย่างการฝึกอบรมจะถูกจัดเรียงเพื่อ โหนดลูกหลานที่เหมาะสม
- กระบวนการทั้งหมดจะถูกทำซ้ำ โดยการใช้ตัวอย่างการฝึกอบรมที่เกี่ยวข้องกับ แต่ละโหนดลูกหลานเพื่อเลือกคุณลักษณะที่ดีที่สุดในการทดสอบ

3. อัลกอริทึม ID 3 (Iterative Dichotomiser3)

เป็นขั้นตอนวิธีที่ใช้ในการสร้างต้นไม้ตัดสินใจโดยนำหลักการของทฤษฎีข่าวสารมาใช้ค่าที่วัดได้จะนำมาตัดสินใจว่าจะใช้ตัวแปรใดในการแบ่งข้อมูลโดยวิธีการกำหนดโครงสร้างต้นไม้ตัดสินใจจะเป็นการเลือกข้อมูลตามลำดับของตัวชี้วัดหรือค่าเกณฑ์(Gain)สูงที่สุดเป็นข้อมูลเริ่มต้นและข้อมูลถัดไปที่มีค่าลดหลั่นกันตามลำดับ โดยสูตรการหาค่า Information Gain ตามสมการที่ 2.1

$$Gain(S, A) = Entropy(s) - \sum_{v \in value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2.1)$$

และค่าการสุ่มตัวอย่าง(Entropy) สามารถหาได้จากสมการที่ 2.2

$$Entropy(S) = \sum_{i=1}^e -p_i \log 2p_i \quad (2.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ID 3 Algorithms

```

If Example is empty

    Then below this new branch add a leaf node

    Else below this new branch add the subtree

        ID3(Examplesvi, target_attribute, attributes -{A})

End

Return Root
  
```

4. อัลกอริทึม C4.5

เป็นอัลกอริทึมที่ใช้ในการสร้างต้นไม้ตัดสินใจพัฒนาโดยรอสส์ควินเลน. C4.5 อัลกอริทึม เป็นส่วนขยายของ ID3 หรือพัฒนาจาก C4.5 การตัดสินใจของต้นไม้ที่สร้างขึ้น โดย C4.5 สามารถใช้สำหรับการจำแนกข้อมูล และด้วยเหตุผลนี้ C4.5 มักจะถูกเรียกว่าเป็นลักษณะนาม สติติ

C4.5 สร้างต้นไม้ตัดสินใจจากชุดของข้อมูลการฝึกอบรมในลักษณะเดียวกับ ID3 โดยใช้แนวคิดของเอนโทรปี ข้อมูลการฝึกอบรมเป็นชุด $S = \{S_1, S_2, \dots\}$ ของกลุ่มตัวอย่างจำแนกแล้ว S_i แต่ละตัวอย่างประกอบด้วยเวกเตอร์ P-มิติ $(X_{\{1,i\}}, X_{\{2,i\}}, \dots, X_{\{p,i\}})$ ถูก X_p แสดงคุณลักษณะหรือคุณสมบัติของตัวอย่างเช่น ดีเป็นชั้นที่ S_i น้ำตกในแต่ละโหนดของต้นไม้ C4.5 เลือกแอตทริบิวต์ของข้อมูลที่มีประสิทธิภาพมากที่สุดแยกชุดของกลุ่มตัวอย่างออกเป็น ส่วนย่อยอุดมในระดับหนึ่งหรืออื่น ๆ เกณฑ์แยกเป็นกำไรจากข้อมูลปกติ (ความแตกต่างในเอนโทรปี) แอตทริบิวต์ที่มีกำไรปกติข้อมูลสูงสุดได้รับเลือกในการตัดสินใจอัลกอริทึม C4.5 แล้วเกิดขึ้นอีกในรายการย่อยขนาดเล็ก อัลกอริทึมนี้มีไม้ก็กรณีฐานตัวอย่างทั้งหมดในรายการที่อยู่ในระดับเดียวกัน เมื่อเกิดเหตุการณ์นี้มันก็สร้างโหนดใบสำหรับต้นไม้ตัดสินใจบอกว่าจะเลือกชั้นเรียนที่ไม่มีคุณสมบัติให้ได้รับข้อมูลใด ๆ ในกรณีนี้ C4.5 สร้างโหนดการตัดสินใจสูงขึ้นไปบนต้นไม้โดยใช้ค่าที่คาดหวังของการเรียนรู้

5. ขั้นตอนการทำต้นไม้ตัดสินใจ

ผู้เขียนขอยกตัวอย่างข้อมูลสภาพภูมิอากาศที่มีผลต่อการเล่นกีฬาเทนนิส

ตารางที่ 2.2 ตัวอย่างข้อมูลสภาพภูมิอากาศที่มีผลต่อการเล่นกีฬาเทนนิส

Day	Outlook	Temperature	Humidity	Wind	Play tennis
Day1	Sunny	Hot	High	Weak	No
Day2	Sunny	Hot	High	Strong	No
Day3	Overcast	Hot	High	Weak	Yes
Day4	Rain	Mild	High	Weak	Yes
Day5	Rain	Cool	Normal	Weak	Yes
Day6	Rain	Cool	Normal	Strong	No
Day7	Overcast	Cool	Normal	Strong	Yes
Day8	Sunny	Mild	High	Weak	No
Day9	Sunny	Cool	Normal	Weak	Yes
Day10	Rain	Mild	Normal	Weak	Yes
Day11	Sunny	Mild	Normal	Strong	Yes
Day12	Overcast	Mild	High	Strong	Yes
Day13	Overcast	Hot	Normal	Weak	Yes
Day14	Rain	Mild	High	Strong	No

จากตัวอย่างตารางที่ 2.2 นี้สามารถอธิบายได้ว่า

- Day** คือ วันที่ทำการบันทึกสภาพอากาศ
- Outlook** คือ สภาพอากาศ ว่า
- Temperature** คืออุณหภูมิที่สามารถวัดได้
- Humidity** คือสภาพความชื้นของอากาศโดยแทนด้วยปกติ (Normal) และความชื้นสูง (High)
- Wind** คือสภาพความแรงของลม

Play tennis คือสถานะมีการเล่นเทนนิสหรือไม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ในพิธีการพิเศษเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยสามารถสรุปข้อมูลจากตารางที่ 2.2 และทำการคำนวณค่าต่างๆ เพื่อทำการสร้างต้นไม้การตัดสินใจโดยใช้เทคนิค ID3 ดังนี้

1. มีจำนวน Record จำนวน 14 Record
2. เมื่อต้องการทราบว่า จะเล่น Tennis หรือไม่ ดังนั้น Play Tennis เป็น Target Attribute
3. จำนวนหาค่า Pop ของ Target Attribute มี No = 5, Yes = 9, All Data = 14
4. จำนวนหา Entropy

$$Entropy(S) = -\frac{9}{14} \log_2 \frac{9}{14} - \frac{5}{14} \log_2 \frac{5}{14} = 0.94 \text{ Attribute} \quad (2.3)$$

5. จำนวนหา Information Gain ของ Attribute อื่น ๆ เพื่อหาว่า Attribute ไหนสำคัญกว่า

$$Gain(S, Wind) \text{ All Data} = 14$$

$$\text{Weak} = 8 \text{ ซึ่งมี Yes} = 6, \text{ No} = 2$$

$$\text{Strong} = 6 \text{ ซึ่งมี Yes} = 3, \text{ No} = 3$$

$$Gain(S, Wind) = 0.94 - \left(\frac{6}{14} \left(-\frac{3}{6} \log_2 \frac{3}{6} - \frac{3}{6} \log_2 \frac{3}{6} \right) + \frac{8}{14} \left(-\frac{6}{8} \log_2 \frac{6}{8} - \frac{2}{8} \log_2 \frac{2}{8} \right) \right) \quad (2.4)$$

$$= 0.94 - [(0.5714((-0.75 \times -0.4150) + (-0.25 \times -2))) + (0.4286((-0.5 \times -1) + (-0.5 \times -1)))]$$

$$= 0.94 - [(0.4635 + 0.4286)]$$

$$= 0.0479 = 0.048$$

6. เมื่อคำนวณหาค่า Information Gain ครบแล้วจะได้คำตอบดังนี้

$$- Gain(S, Outlook) = 0.246$$

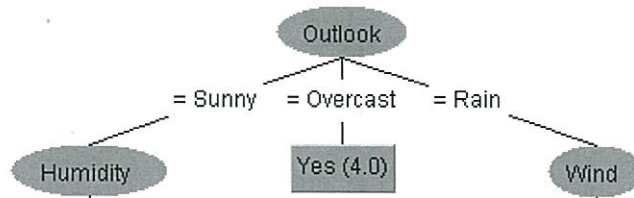
$$- Gain(S, Temperature) = 0.029$$

$$- Gain(S, Humidity) = 0.151$$

$$- Gain(S, Wind) = 0.048$$

7. เลือก Root Node โดยพิจารณาจากค่า Information Gain ที่สูงที่สุดคือ Outlook แสดงว่า Outlook มีความสำคัญสูงสุดดังแสดงในรูปที่ 2.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



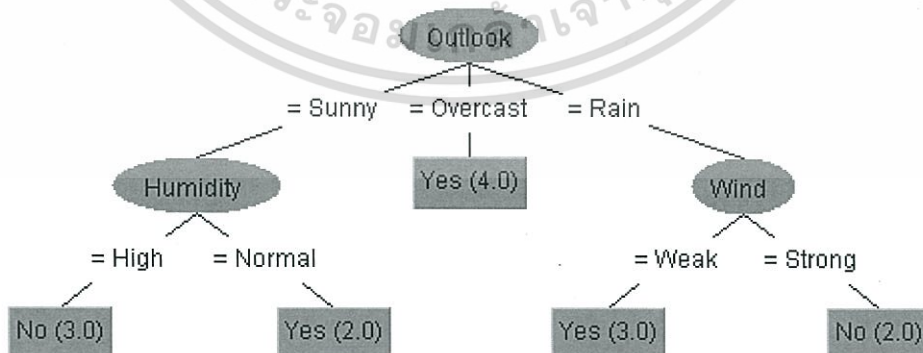
รูปที่ 2.4 การเลือก Root Node

8. พิจารณาว่าต้องแตก Tree ต่อหรือไม่ จะพบว่า Overcast ไม่ต้องแตกต่อแล้วเพราะได้คำตอบแล้ว ส่วน Sunny และ Rain จำเป็นต้องแตกต่อ ซึ่งหลักการในการแตกต่อก็เหมือนเดิม

- หาค่า Pop ของ Target Attribute ในที่นี้คือ Sunny มี All Data = 5 Yes = 2, No = 3
- หา Entropy ของ Sunny

$$Entropy(S) = -\frac{2}{5} \log_2 \frac{2}{5} - \frac{3}{5} \log_2 \frac{3}{5} = 0.97 \quad (2.5)$$

- หา Information Gain ของ Attribute อื่น ๆ นอกเหนือจาก Outlook และ Play Tennis
 - Gain(S, Temperature) = 0.057
 - Gain(S, Humidity) = 0.97
 - Gain(S, Wind) = 0.01
- จะเห็นว่า Humidity มีค่าสูงสุดจึงเลือกเอามาใช้ใน Tree ดังรูปที่ 2.5



รูปที่ 2.5 ภาพ Tree ที่สมบูรณ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

9. ทำการกำหนด Classification Rules จากผลลัพธ์ต้นไม้ตัดสินใจ โดยมีเงื่อนไขดังนี้

- การนำไปใช้คือจะแปลงเป็นกฎ IF THEN ออกมา
- 1 กฎต่อ 1 Path จาก Root Node ถึง Leaf Node
- ที่ Leaf Node จะเป็นการทำนาย การเกิดขึ้นของ Class

จากรูปด้านบนจะสามารถสร้างกฎได้ 5 กฎตาม Leaf Node ดังนี้

1. If Outlook = Sunny and Humidity = Normal Then Play Tennis = Yes
2. If Outlook = Sunny and Humidity = High Then Play Tennis = No
3. If Outlook = Overcast Then Play Tennis = Yes
4. If Outlook = Rain and Wind = Weak Then Play Tennis = Yes
5. If Outlook = Rain and Wind = Strong Then Play Tennis = No

2.5 งานวิจัยที่เกี่ยวข้อง

2.5.1 การศึกษาวิจัย เรื่อง ความรุนแรงต่อเด็กในประเทศไทย

ของสำนักงานส่งเสริมสวัสดิภาพและพิทักษ์เด็ก เยาวชน ผู้ด้อยโอกาสและผู้สูงอายุ กระทรวงพัฒนาสังคมและความมั่นคงของมนุษย์ได้ทำการวิจัยเกี่ยวกับเรื่องความรุนแรงต่อเด็กในประเทศไทยซึ่งความรุนแรง ประสบการณ์ที่ได้รับที่เกิดขึ้นนั้นอาจส่งผลต่อเด็กในอนาคตโดยก่อให้เกิดปัญหาทางด้านจิตใจและการวิจัยนี้ได้ทำการเก็บรวบรวมข้อมูลความรุนแรงที่เกิดขึ้นกับเด็กโดยพิจารณา ซึ่งการเก็บข้อมูลดังกล่าวใช้วิธีการเก็บข้อมูลแบบกลุ่มตัวอย่างตามความเหมาะสมโดยพิจารณาความรุนแรงที่เกิดขึ้น

1. ทางกาย
2. ทางวาจา
3. ทางเพศ

โดยจะเก็บข้อมูลที่เกิดจากเหตุการณ์ที่เกิดขึ้น

1. จากเด็กที่ถูกเพื่อนกระทำความรุนแรง
2. จากเด็กเป็นผู้กระทำความรุนแรงต่อเพื่อน
3. จากครอบครัว
4. จากโรงเรียน โดยครู

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และพิจารณา แบ่งตาม ลักษณะของสถานศึกษา ,ภูมิภาค เป็นการสรุปวิเคราะห์ในเชิงสถิติ การเก็บข้อมูลแบ่งเป็น 2 แบบ

1. แบบสอบถามใช้โปรแกรม Statistical Package for the Social Sciences (SPSS) โดยจะคิดออกมาเป็นอัตราร้อยละ

2. แบบสัมภาษณ์ใช้วิธีการวิเคราะห์เนื้อหาและจัดเก็บเป็นหมวดหมู่ ผู้เขียนเห็นว่างานวิจัยทางด้านจิตวิทยานั้นมีประโยชน์มากเนื่องจากข้อมูลที่กลุ่มวิจัยได้จัดทำมานั้นสามารถที่จะนำไปต่อยอดในการพัฒนาเครื่องมือที่จะนำไปใช้กับเยาวชนของชาติโดยจะนำเอาข้อมูลทางสถิติทางด้านพฤติกรรม อารมณ์ ดังกล่าวมาประกอบเป็นฐานข้อมูลเพื่อพัฒนาเป็นเครื่องมือที่อาจช่วยปรับปรุงหรืออาจเปลี่ยนแปลงพฤติกรรมที่ไม่ดีได้

2.5.2 การพัฒนาระบบทำนายแนวโน้มพฤติกรรมและผลสัมฤทธิ์ทางการศึกษาของนักเรียน

โดย ศราวุธ ชนะบำรุง ได้ทำวิจัยเกี่ยวกับผลสัมฤทธิ์ทางการศึกษา โดยพิจารณาพฤติกรรมทั้งหมด 5 ด้าน ได้แก่ พฤติกรรมด้านอารมณ์ พฤติกรรมเกร พฤติกรรมไม่อยู่นิ่ง พฤติกรรมสัมพันธ์กับเพื่อน และพฤติกรรมสัมพันธ์ทางสังคม โดยมีการประเมินผลของพฤติกรรมที่มีผลต่อการเรียนของนักเรียนด้วยกัน 3 ระดับ

1. พฤติกรรมปกติ
2. พฤติกรรมเสี่ยง
3. พฤติกรรมปัญหา

การวิจัยนี้ใช้วิธีการแบ่งกลุ่มของข้อมูลโดยวิธี กฎความสัมพันธ์ (Association Rule) ในการระบุกลุ่มของข้อมูลที่ใช้ในการวิเคราะห์กลุ่มของพฤติกรรมดังกล่าว

บทที่ 3

วิธีการดำเนินการศึกษา

ในการศึกษาโครงการนี้เป็นการนำทฤษฎีการศึกษาพฤติกรรมของมนุษย์แลทฤษฎีต้นไม้ตัดสินใจเป็นเครื่องมือช่วยในการจำแนกพฤติกรรมของกลุ่มตัวอย่างที่ได้มาจากการทำแบบสอบถามพฤติกรรมของกลุ่มนักเรียนในระดับประถมศึกษาถึงมัธยมศึกษาเป็นต้นแบบ

เมื่อนำเอาข้อมูลพฤติกรรมจากผู้ถูกประเมินที่ได้จากแบบสอบถามเข้าสู่เครื่องมือหรือโมเดลจะทำให้สามารถที่จะทำนายพฤติกรรมของผู้ถูกประเมินได้ว่าจะมีแนวโน้มพฤติกรรมเป็นอย่างไรในอนาคต

แบบประเมินพฤติกรรมที่ได้ทำการคัดเลือกข้อความที่เป็นประโยชน์ในการเก็บข้อมูลของผู้ถูกประเมินถูกจัดเก็บในรูปแบบเอกสารประเภท Microsoft Excel โดยผู้เขียนสังเกตเห็นว่าข้อมูลดังกล่าวสามารถที่จะนำมาเป็นข้อมูลพื้นฐานในการสร้างต้นแบบผู้เขียนจึงนำเอาแบบประเมินพฤติกรรมดังกล่าวมาทำการคัดเลือกคำถามและจัดเก็บในรูปแบบเอกสาร Excel และจำเป็นต่อการออกแบบ โมเดล โดยใช้ทฤษฎีต้นไม้ตัดสินใจ (Decision Tree)

3.1 ข้อมูลที่ใช้กับระบบ

โมเดลที่จัดทำในโครงการนี้จะรับข้อมูลเข้าในลักษณะของ Excel File ประกอบไปด้วยคุณลักษณะต่างๆของชุดข้อมูลที่จะนำมาใช้โดยแต่ละคุณลักษณะจะเป็นข้อมูลชนิดตัวอักษรและผลของคำตอบจะจัดเก็บเป็นตัวเลขดังรูปที่ 3.1

แถว/คอล	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20			
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
3	15	16	17	16	16	15	15	16	15	15	16	16	16	16	16	17	16	15	17	16	15			
4	21	7	7	1	13	30	13	19	29	28	28	12	16	2	2	20	28	1	19	6				
5	11	11			3	1	2	6	5	1	8	10	9	4	3	6	11	12	12	3				
6	37	36	36	37	35	35	35	37	37	38	36	36	36	37	36	37	36	37	35	36	35			
7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6			
8	2	5	5	1	1	9	5	2	1	2	3	1	1	5	1	1	1	1	1	1	6			
9	5	9	9	2	9	9	9	9	3	3	6	9	6	9	9	9	9	9	6	9	3	9		
10	3	9	9	4	9	9	9	4	3	4	6	9	4	9	5	9	4	9	3	9				
11	2	1	2	2	1	1	1	1	2	2	1	2	1	1	1	1	1	1	1	1	3	2		
12	4	4	1	5	6	7	7	8	4	6	5	7	6	5	4	10	10	6	4					
13	2	1	17	17	1	1	1	1	1	1	4	1	1	5	1	1	1	1	1	17	11			
14	3	12	3	3				2	4	8	4	13	2	2	9	2	15	2	2					
15	4																							
16	1	1	1	1	1	1	1	1	1	1	2	2	2	3	2	3	2	1	1	1	2	1		
17	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1		
18	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1		
19	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1		
20	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1		

รูปที่ 3.1 ตัวอย่างข้อมูลต้นแบบที่นำมาใช้ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับนักเรียนที่สนใจศึกษาข้อมูลเพิ่มเติมเกี่ยวกับทฤษฎีต้นไม้ตัดสินใจและการนำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.1 เป็นข้อมูลดิบต้นแบบที่ได้ทำการบันทึกมาจากข้อคำถาม สามารถสรุปลักษณะของข้อมูลที่น่ามาใช้ได้ ดังนี้

หมายเลข 1 คือ ลำดับของข้อมูล

หมายเลข 2 คือ ตารางรหัสนักเรียน โดยค่าถูกแทนด้วยตัวเลขประด้วยตัวเลข 5 หลัก

หมายเลข 3 คือ ตารางวัน/เดือน/ปี เกิดของนักเรียน

หมายเลข 4 คือ ตารางเก็บข้อมูลของนักเรียนประกอบไปด้วย

- สถานะของบิดาและมารดา
- การศึกษาของบิดา
- การศึกษาของมารดา
- คือ มีพี่น้องร่วมสายโลหิตกี่คน

หมายเลข 5 คือ ตารางเก็บข้อมูลคำตอบจากชุดแบบสอบถาม

ส่วนที่ ๑ (เพื่อน)				
คำอธิบาย: ๑. ขอให้ตอบทุกข้อ โดยกากบาท (X) ถ้าตอบในแต่ละข้อ				
๒. เพื่อน หมายถึง เพื่อนร่วมห้อง เพื่อนรุ่นพี่ เพื่อนรุ่นน้อง เพื่อนร่วมโรงเรียน และเพื่อนที่เคยเรียนด้วยกัน				
๓. คำว่า "บ่อย" หรือ "ไม่บ่อย" ให้ตอบตามความรู้สึกของตัวเอง				
๑. เท่าที่จำได้ มีเพื่อนเคยทำแบบที่กับเราหรือไม่? ขอให้ตอบทุกข้อ				
ข้อ	ข้อความ	ไม่เคย	เคย แต่ไม่บ่อย	เคยบ่อย
๑.๑	เขกหัว หรือเอามือฟาดหัว (เบ็ดเตล็ด)			
๑.๒	กระชาก			
๑.๓	ตบหรือตีแขนหรือหัว			
๑.๔	ตบหน้าหรือท้ายทอย			
๑.๕	จับหัวโยก			
๑.๖	เตะ			
๑.๗	ชกต่อย			
๑.๘	ปาดหรือข่วนของใส่			
๑.๙	กระแทบ			
๑.๑๐	รุมเตะ ต่อย กระแทบ			
๑.๑๑	เอาของอันตรายหรือสกปรก ใส่กระเป๋าทิ้งไว้ให้เรียน (เช่น ถูน้ำลายใส่ ใส่ขยะ ใส่ขี้ดินสอดใส่ตะปูล ฯลฯ)			
๑.๑๒	โยนบก โยนน้ำ			
๑.๑๓	แบ่งพรรคแบ่งพวก หรือไม่ให้เข้าร่วมกลุ่ม			
๑.๑๔	บิบบคองหนายใจไม่ออก			
๑.๑๕	บังคับให้กินของสกปรก หรือสิ่งที่ไม่ใช่อาหารที่อาจเป็นอันตราย			
๑.๑๖	เอาอาหารร้อน (เช่น ถ้วยเดียว ข้าวต้ม ฯลฯ) หรือน้ำร้อน สาดใส่			

รูปที่ 3.2 ตัวอย่างแบบสอบถามที่ใช้ประเมินพฤติกรรมที่เป็นผู้ถูกระทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษเท่านั้น ไม่อนุญาตให้เผยแพร่โฆษณาเป็นการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.2 เป็นตารางตัวอย่างการเก็บข้อมูลของผู้ตอบคำถามในเชิงพฤติกรรมที่ผู้ตอบแบบสอบถามจะป้อนข้อมูลพฤติกรรมลงไปโดยในคำถามจะประกอบด้วย

- ชุดคำถาม 22 คำถาม
- แต่ละคำถามมี 3 คำตอบประกอบด้วย ไม่เคย,เคยแต่ไม่บ่อยและ เคยบ่อยๆ

โดยชุดคำถามจะสอบถามในเชิงพฤติกรรมที่อาจเกิดขึ้นกับตัวผู้ตอบแบบสอบถามเอง หรือผู้อื่น กระทำกับผู้ตอบแบบสอบถามซึ่งในการค้นคว้านี้ผู้เขียนเองได้สนใจประเด็นพฤติกรรมโอกาสที่ผู้ตอบแบบสอบถามอาจก่อความรุนแรงต่อผู้อื่นและส่งผลให้ผู้อื่นเกิดความเดือดร้อนในอนาคตได้ ผู้เขียนจึงหยิบยกเอาข้อมูลที่เกี่ยวข้องกับตัวผู้ตอบแบบสอบถามมาทำการวิเคราะห์และทำการศึกษาโรคครงการนี้

แบบ สบป.

๔. เมื่อเพื่อนทับกับเราตามข้อ ๑.๑ - ๑.๒๓ แล้ว เราได้บอกใครบ้าง?
(เลือกตอบข้อ ๔.๑ หรือ ข้อ ๔.๒ ตามความเป็นจริง)

๔.๑ ไม่บอกใครเลยตามข้อ ๑.๑ - ๑.๒๓ เพราะ... (ตอบได้มากกว่า ๑ คำตอบ)

(1) เป็นเรื่องธรรมดา (2) เป็นเรื่องน่าอาย (3) ไม่อยากให้ใครรู้

(4) ลังบอกก็ไม่มีการเรื่อง (5) กลัวเพื่อน (6) กลัวเพราะถูกขู่

(7) รู้สึกว่าเราผิดเอง (8) อื่น ๆ (ช่วยบอกด้วย.....)

(ตอบถึงตรงนี้แล้ว ข้ามไปตอบข้อ ๕)

๔.๒ บอกบางเรื่อง...ตอบด้วยว่าบอกใคร (ตอบได้มากกว่า ๑ คำตอบ)

(01) บอกเพื่อน (02) บอกแฟน (03) บอกพ่อ

(04) บอกแม่ (05) บอกครู (06) บอกพี่น้อง

(07) บอกรุ่นพี่ (08) บอกญาติ (09) บอกพ่อแม่เพื่อน

(10) บอกญาติของเพื่อน (11) บอกเพื่อนบ้าน (12) บอกหมอ

(13) บอกตำรวจ (14) บอกนักสังคมสงเคราะห์

(15) บอกพระหรือผู้นำทางศาสนา (16) แจ้งสายด่วนหรืออินเตอร์เน็ต

(17) บอกคนอื่น คือ.....

๔.๒.๑ ที่บอกผู้อื่น เพราะอะไร? (ตอบได้มากกว่า ๑ คำตอบ)

(1) ต้องการให้ผู้อื่นทราบการกระทำของเพื่อน (2) รู้สึกว่าเพื่อนทำรุนแรง

(3) ต้องการความช่วยเหลือ (4) ต้องการความเห็นใจ

(5) ต้องการให้ผู้อื่นบอกเพื่อนไม่ให้ทำอีก (6) ต้องการให้ผู้อื่นแก้แค้นแทน

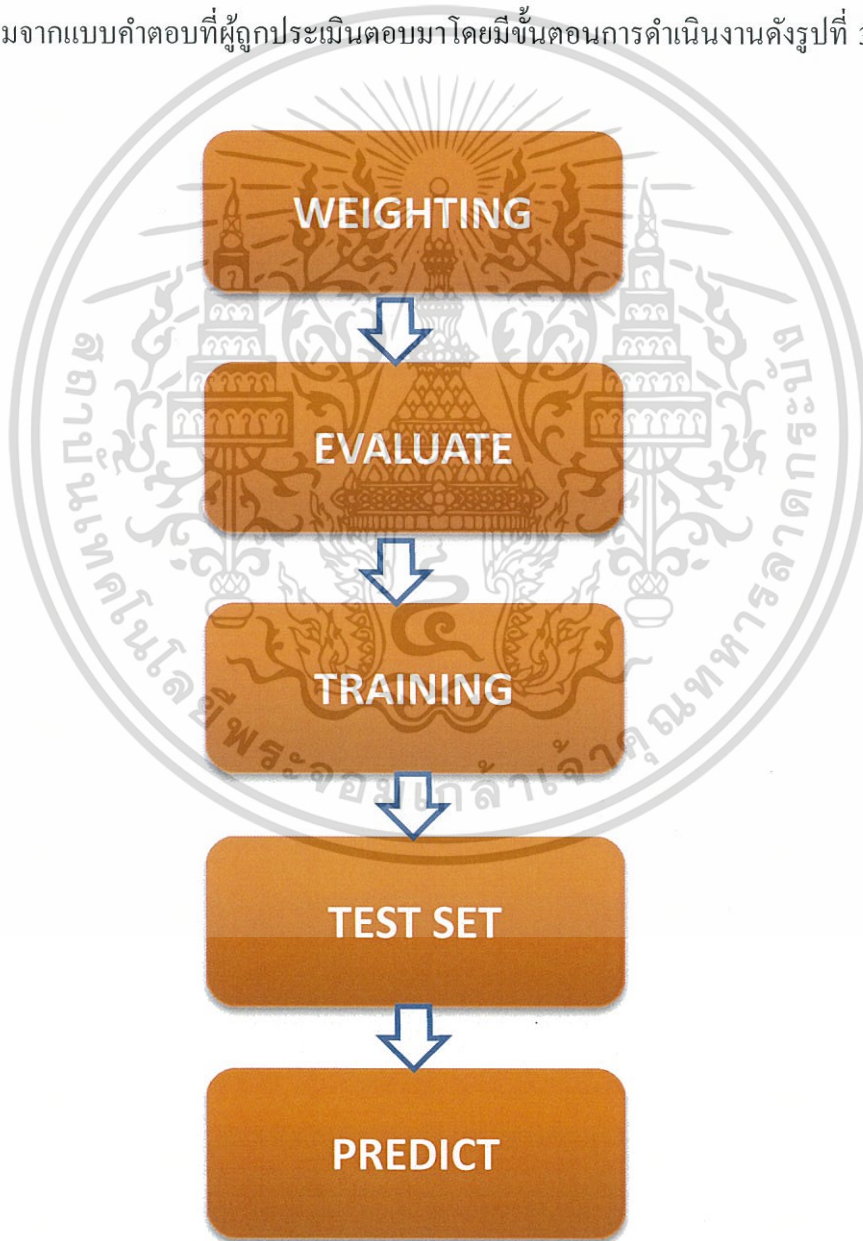
(7) อื่น ๆ (ช่วยบอกด้วย.....)

เอกสารนี้เป็นเอกสารรูปที่ 3.3 ตัวอย่างแบบสอบถามที่ใช้ประเมินพฤติกรรมด้านความรู้สึกระยะโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 3.3 เป็นข้อคำถามที่ใช้เก็บข้อมูลในแง่มุมมองความรู้สึกและสิ่งที่ต้องการจะกระทำต่อหลังจากที่ผู้ตอบแบบสอบถามเป็นผู้ถูกกระทำ

3.2 ขั้นตอนการดำเนินงาน

การศึกษาโครงการนี้จะดำเนินการศึกษาทฤษฎีพฤติกรรมมนุษย์ที่เกี่ยวข้องกับการมีโอกาสที่จะมีแนวโน้มเป็นอาชญากร โดยอาศัยการวิเคราะห์ข้อมูลจากแบบสอบถามที่ให้ผู้ถูกประเมินทำการตอบคำถามแล้วใช้เทคนิค ต้นไม้ตัดสินใจ ในการจำแนกหมวดหมู่ของข้อมูล เพื่อทำนายข้อมูลพฤติกรรมจากแบบคำตอบที่ผู้ถูกประเมินตอบมา โดยมีขั้นตอนการดำเนินงานดังรูปที่ 3.4



รูปที่ 3.4 ขั้นตอนการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ขั้นตอนการวิเคราะห์ข้อมูลชุดคำตอบจากข้อมูลของหน่วยงานที่ทำการเก็บค่าจากกลุ่มประชากรทั่วประเทศ โดยการวิจัยนี้เลือกเอามาข้อมูลที่จะทำการศึกษามาจากกลุ่มตัวอย่างของนักเรียนชั้นมัธยมศึกษาตอนต้นและตอนปลายเป็นจำนวนทั้งสิ้น 2,895 คน

2. ขั้นตอนการวิเคราะห์ชุดคำถาม-ตอบเป็นขั้นตอนวิเคราะห์ชุดคำถาม-ตอบโดยชุดคำถามมีจำนวนข้อ 22 ข้อแต่ละคำถามมี คำตอบ 3 คำตอบ ประกอบด้วย

ไม่เคย

เคยไม่บ่อย

เคยบ่อย

ดังนั้นความเป็นไปได้ของคำตอบ $22^3 = 10,648$ คำตอบ

3. ขั้นตอนการกำหนดค่าน้ำหนักคะแนนให้กับแบบสอบถาม ขั้นตอนการกำหนดค่าความสำคัญหรือค่าน้ำหนักให้กับแต่ละคำถาม และค่าน้ำหนักของ คำตอบโดยทำการคำนวณหาจำนวนกลุ่มประชากรตัวอย่างจากประชากรทั้งหมด 2,895 คน เพื่อนำเอามาคำนวณหาจำนวนคนในการที่จะไปเก็บข้อมูลนำมาประเมินค่าน้ำหนักของแต่ละคำถาม

$$N = \frac{n}{1 + Ne^2} \quad (3.1)$$

โดย

n = จำนวนกลุ่มตัวอย่าง

N = จำนวนรวมทั้งหมดของประชากรที่ใช้ในการศึกษา

e = ความผิดพลาดที่ยอมรับได้ (โดยกำหนดให้เท่ากับ 0.05)

$$N = \frac{2,895}{1 + (2,895 * 0.05^2)} \quad (3.2)$$

$$N \approx 352 \text{ คน}$$

สรุปได้ว่าจำนวนกลุ่มตัวอย่างที่จะนำมาทำเป็นกลุ่มตัวอย่างในการกำหนดค่าน้ำหนักให้กับแต่ละข้อคำถามเป็นจำนวนทั้งสิ้น 352 คน

4. ทำการกำหนดค่าน้ำหนักให้กับแต่ละข้อคำถามโดยใช้วิธีการนำเอาชุดคำถามไปให้กลุ่มตัวอย่าง ให้น้ำหนักแต่ละคำตอบ โดยแบ่งกลุ่มตัวอย่างเป็น 3 กลุ่ม

- ค่าคะแนนที่ได้มาจากกลุ่มนักเรียนตัวอย่าง 117 คน
- ค่าคะแนนที่ได้มาจากกลุ่มครูตัวอย่าง 117 คน
- ค่าคะแนนที่ได้มาจากกลุ่มผู้ปกครองตัวอย่าง 117 คน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.1 ตารางเก็บค่าน้ำหนักคะแนน

ข้อความ	ระดับค่าน้ำหนัก				
	5	4	3	2	1
1	8	16	111	121	101
2	22	64	97	63	111
3	10	55	116	83	93
4	71	136	61	37	52
5	254	66	18	12	7
6	75	137	82	27	36
7	131	111	32	21	62
8	61	167	60	37	32
9	263	56	9	10	19
10	242	86	16	11	2
11	100	109	57	23	68
12	97	121	78	19	42
13	10	14	27	48	258
14	272	48	14	8	15
15	279	44	10	4	20
16	179	71	21	12	74
17	269	60	20	7	1
18	114	116	40	21	66
19	306	40	5	3	3
20	319	30	1	5	2
21	311	35	3	6	2
22	331	9	9	5	3

จากตารางที่ 3.1 เป็นตารางที่เก็บค่าน้ำหนักของแต่ละข้อความจากกลุ่มตัวอย่าง 3 กลุ่ม (352 คน)

5. หาค่าเฉลี่ย (\bar{X}) ของแต่ละคำถามเพื่อหาค่าความสำคัญหรือค่าน้ำหนักของแต่ละคำถาม โดยค่าเฉลี่ย หมายถึงการหารผลรวมของข้อมูลทั้งหมดด้วยจำนวนข้อมูลทั้งหมดเพื่อหาค่าที่มี

เอกสารแนบเข้าสู่ส่วนกลางหรือ อาจเรียกได้ว่าหาค่ากลางนั่นเองโดย อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\bar{X} = \frac{\sum x}{n} \quad (2.8)$$

เมื่อ \bar{X} คือ ค่าเฉลี่ยของเลขคณิต
 $\sum x$ คือ ผลบวกของข้อมูลทุกค่า
 n คือ จำนวนข้อมูลทั้งหมด

ตารางที่ 3.2 ตารางค่าน้ำหนักของแต่ละคำถาม

ข้อคำถาม	\bar{X}	ค่าน้ำหนัก
1	2.17	2
2	2.50	2
3	2.45	2
4	3.39	3
5	4.56	5
6	3.53	4
7	3.65	4
8	3.53	4
9	4.52	5
10	4.58	5
11	3.43	3
12	3.60	4
13	1.49	1
14	4.57	5
15	4.59	5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 (ต่อ)

ข้อคำถาม	\bar{X}	ค่าน้ำหนัก
16	3.76	4
17	4.67	5
18	3.45	4
19	4.83	3
20	4.87	5
21	4.84	5
22	4.88	5
Summary		87

จากตารางสรุปได้ว่า

กลุ่มของค่าน้ำหนักคะแนน 5 คะแนนได้แก่ข้อคำถามข้อที่ 5,9,10,14,15,17,19,20,21,22

กลุ่มของค่าน้ำหนักคะแนน 4 คะแนนได้แก่ข้อคำถามข้อที่ 6,7,8,12,16,18

กลุ่มของค่าน้ำหนักคะแนน 3 คะแนนได้แก่ข้อคำถามข้อที่ 4,11

กลุ่มของค่าน้ำหนักคะแนน 2 คะแนนได้แก่ข้อคำถามข้อที่ 1,2,3

จากข้อมูลการแบ่งระดับค่าน้ำหนักคะแนน ออกเป็น 5 ระดับผู้เขียนได้นำเอาหลักวิธีการแบ่งช่วงคะแนนตามหลักวิธีการของ บุญชม ศรีสะอาด โดยแบ่งเป็น 5 ระดับ

ค่าเฉลี่ยระหว่าง 4.50 - 5.00 หมายถึง ระดับมากที่สุด

ค่าเฉลี่ยระหว่าง 3.50 - 4.49 หมายถึง ระดับมาก

ค่าเฉลี่ยระหว่าง 2.50 - 3.49 หมายถึง ระดับปานกลาง

ค่าเฉลี่ยระหว่าง 1.50 - 2.49 หมายถึง ระดับน้อย

ค่าเฉลี่ยระหว่าง 1.00 - 1.49 หมายถึง ระดับน้อยที่สุด

6. นำเอาค่าน้ำหนักคะแนนไปกำหนดให้ข้อคำถาม

7. ขั้นตอนการวิเคราะห์ชุดคำตอบจากข้อมูลที่เก็บมาจากกลุ่มนักเรียนตาม โรงเรียนต่างๆ

โดยผู้เขียนนำมาเพียงส่วนหนึ่งเพื่อใช้ในการศึกษาใน โครงการนี้โดยผู้เขียนสนใจในการศึกษา

ปัจจัย ที่อาจส่งผลให้เกิดแนวโน้มพฤติกรรมความรุนแรงขึ้นได้

8. พิจารณาข้อมูลต้นฉบับเพื่อเตรียมข้อมูลให้เหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาวิจัย ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.5 ตารางเก็บข้อมูลต้นฉบับ

จากรูปที่ 3.5 เป็นตารางการเก็บข้อมูลคำตอบของผู้ทำแบบสอบถามจัดเก็บในรูปแบบ Excel โดยตารางจะเก็บข้อมูลในรายละเอียดต่างๆเช่น อายุ เพศ ระดับชั้น สถานะของ พ่อ แม่ การศึกษา ทั้งนี้ ผู้เขียนสนใจในเนื้อหา ข้อมูลคำตอบในเรื่องพฤติกรรมการแสดงออกของผู้ตอบคำถาม จึงหยิบยกส่วนที่สนใจมาวิเคราะห์โดยจากรูปที่ 3.5 อธิบายถึง

A หมายถึง ลำดับของตัวผู้ตอบแบบสอบถาม

B หมายถึง ค่าคำตอบที่ผู้ตอบแบบสอบถามเลือกมามีค่าตั้งแต่ 1-3 โดย

1 แทน ความหมาย ไม่เคย

2 แทน ความหมาย เคยแต่ไม่บ่อย

3 แทน ความหมาย เคยบ่อยๆ

C หมายถึง ลำดับของข้อคำถาม โดยมีจำนวนข้อคำถาม 22 ข้อ

ทำการจัดเรียงข้อมูลใหม่ เพื่อให้ง่ายต่อการเตรียมข้อมูลพื้นฐานเพื่อจัดทำเป็นโมเดลต้นแบบ โดยข้อมูลจะถูกจัดเก็บในรูปแบบ Excel ดังรูปที่ 3.5

ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Point	
1	0	1	0	1	0	0	0	1	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0.184
2	1	1	2	0	1	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1.736
3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.046
4	1	0	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0.253
5	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0.218
6	1	0	0	0	0	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0.333
7	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.057
8	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.207
9	1	1	2	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0.391

รูปที่ 3.6 รูปตัวอย่างตารางเก็บข้อมูลคำตอบของกลุ่มนักเรียนตัวอย่าง

จากรูปที่ 3.6 เป็นรูปตารางเก็บค่าคำตอบจากนักเรียนที่ทำแบบสอบถามโดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

A หมายถึง ลำดับผู้ตอบแบบสอบถาม

B หมายถึง ข้อคำถามมีทั้งหมด 22 ข้อ โดยแต่ละคำถามประกอบด้วยคำตอบ 3 คำ

- 0 แทนด้วย ไม่เคย
- 1 แทนด้วย เคยแต่ไม่บ่อย
- 2 แทนด้วย เคยบ่อย

C หมายถึง ค่าคะแนนได้จากการคำนวณ

ทำการจัดกลุ่มผู้ตอบแบบสอบถามจากค่าที่ได้จากการคำนวณใน โดย

ผู้เขียนใช้วิธีการจัดกลุ่มเป็น 3 กลุ่มด้วยกันประกอบด้วย

- กลุ่มคนที่ไม่โอกาสเสี่ยงที่จะก่อความรุนแรง แทนค่าด้วย A
- กลุ่มคนที่มีโอกาสเสี่ยงที่จะก่อความรุนแรงปานกลาง แทนค่าด้วย B
- กลุ่มคนที่มีโอกาสเสี่ยงที่จะก่อความรุนแรงมาก แทนค่าด้วย C

ID	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Point	Grade	
1	0	1	0	1	0	0	0	1	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0.184	A
2	1	1	2	0	1	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1.736	C
3	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.046	A
4	1	0	1	0	0	1	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0.253	A
5	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0.218	A
6	1	0	1	1	0	1	1	1	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0.333	A
7	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.057	A
8	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.207	A
9	1	1	2	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0.391	A

รูปที่ 3.7 ตารางคำตอบที่ทำการแบ่งกลุ่มผู้ตอบตามระดับความรุนแรง

9. ทำการบันทึกเพิ่มเอกสารเป็น นามสกุล .CSV เพื่อเตรียมนำเข้าสู่โปรแกรมประยุกต์ WEKA เพื่อเข้าสู่กระบวนการ Classification Decision tree ต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

บทนี้จะกล่าวถึงรายละเอียดของขั้นตอนการใช้งาน โปรแกรมประยุกต์ การแบ่งข้อมูลที่จะนำมาฝึกสอนระบบ, ทดสอบระบบ และการทำงาน โดยแบ่งการอธิบายออกเป็นส่วนต่าง ๆ ดังนี้

1. การใช้งานโปรแกรมประยุกต์ WEKA เพื่อการทำงาน
2. ข้อมูลที่ใช้กับโปรแกรมประยุกต์
3. ผลการทดลองและการวิเคราะห์

4.1 การใช้งานโปรแกรมการประยุกต์

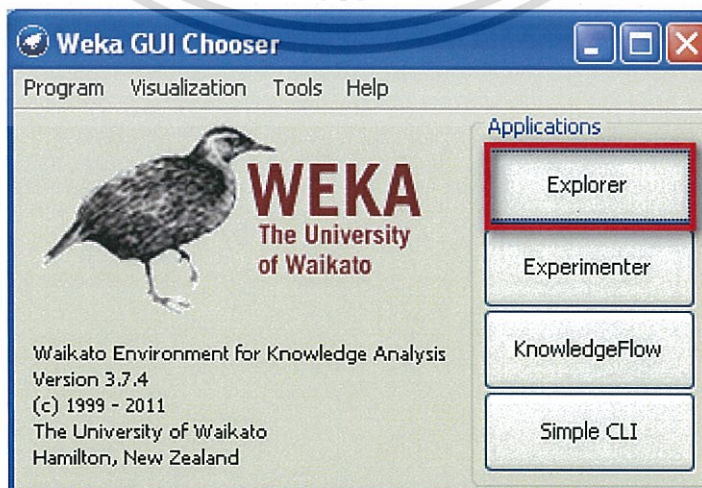
1. ทำการจัดเตรียมเพิ่มเอกสาร Excel ให้เป็นนามสกุล .CSV เพื่อนำเข้าสู่โปรแกรมประยุกต์ WEKA โดยมีขั้นตอนดังนี้

1.1 แบ่งข้อมูลแต่ละชุดคำตอบออกเป็น 3 เพิ่มเอกสารประกอบด้วย

- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Training Set มีข้อมูล 70 %
- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Test Set มีข้อมูล 20 %
- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Validation Set มีข้อมูล 10 %

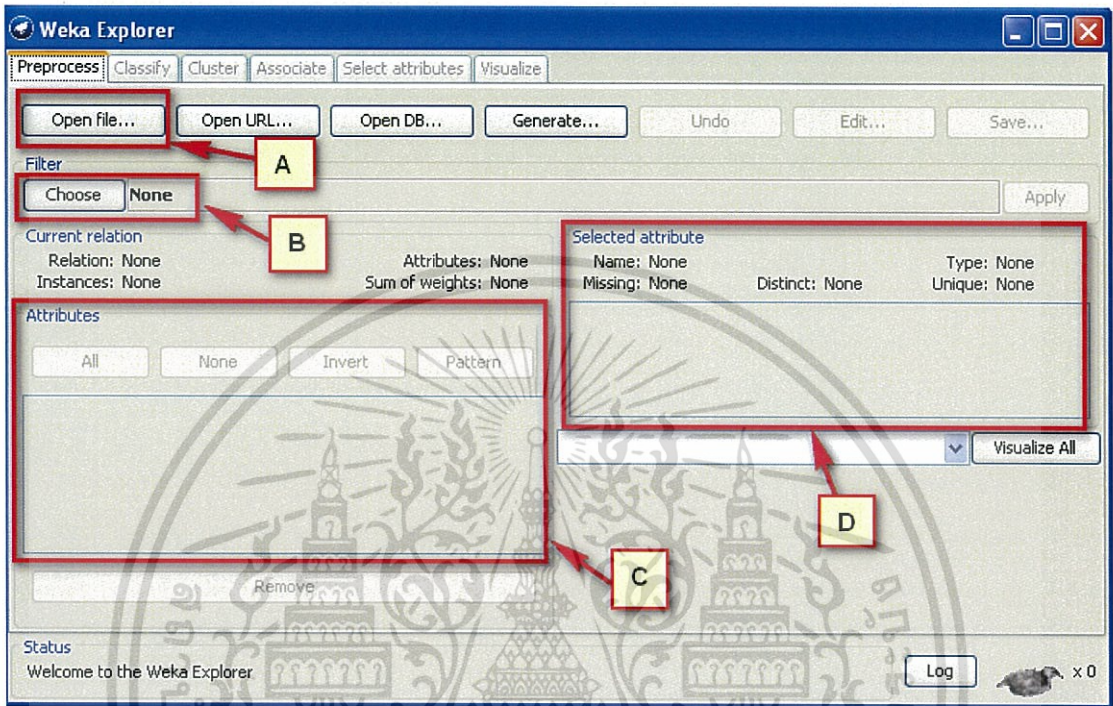
1.2 ทำการ กำจัด Missing Value และ Error ต่างๆ ในแต่ละเพิ่มข้อมูลในตารางที่เก็บข้อมูล

1.3 นำเอาข้อมูลทำการเข้าสู่โปรแกรม WEKA (ผู้เขียนใช้ WEKA รุ่น 3.7.4)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้รูปที่ 4.1 โปรแกรม WEKA อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.1 เป็นหน้าต่างโปรแกรมประยุกต์ WEKA กดเลือกเมนูปุ่ม Explorer เพื่อเข้าสู่การใช้งานโปรแกรม WEKA

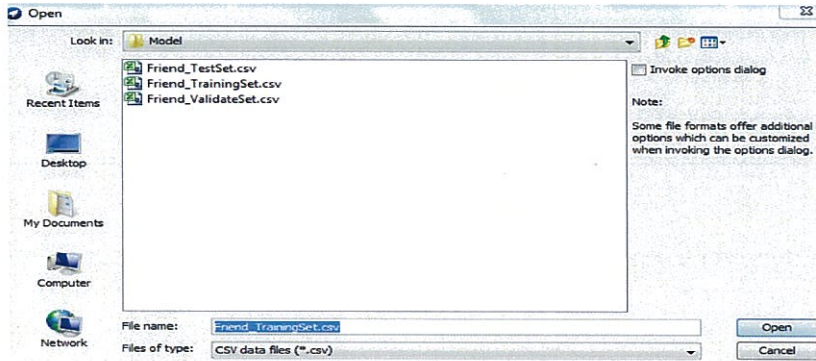


รูปที่ 4.2 โปรแกรม WEKA ในรูปแบบ Explorer

รูปที่ 4.2 คือหน้าต่าง WEKA สำหรับจัดการเตรียมข้อมูลเบื้องต้น โดยมีรายละเอียดดังนี้

- A คือ ปุ่มสำหรับการเลือกไฟล์ที่จะนำเข้าสู่โปรแกรม WEKA โดยข้อมูลในการใช้งานสำหรับการศึกษาโครงการนี้ผู้เขียนใช้ข้อมูลเพิ่มเอกสารประเภท Excel นามสกุล CSV
- B คือ เครื่องมือสำหรับการจัดการคัดกรองเพิ่มเอกสารในกรณีที่มีลักษณะการจัดเก็บเนื้อหาภายในผิดปกติ เช่น คำว่างในตาราง,ค่าซ้ำซ้อน อื่นๆ เพื่อเตรียมความพร้อมก่อนการเข้าสู่กระบวนการทำ Data Mining
- C คือ ช่องแสดงรายละเอียดของข้อมูลหลังจากที่นำเพิ่มเอกสารข้อมูลเข้าสู่โปรแกรม WEKA แล้วในรูปแบบ ข้อความ
- D คือ ช่องแสดงรายละเอียดของข้อมูลหลังจากที่นำเพิ่มเอกสารข้อมูลเข้าสู่โปรแกรม WEKA แล้วในรูปแบบกราฟรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 การเลือกเพิ่มเอกสารที่ต้องการนำเข้าสู่ WEKA

รูปที่ 4.3 เป็นขั้นตอนการนำเอาข้อมูลเข้าสู่โปรแกรม WEKA เพื่อทำการจำแนกข้อมูลเป็นกลุ่มโดยจำแนกตามประเภท การถามดังนี้

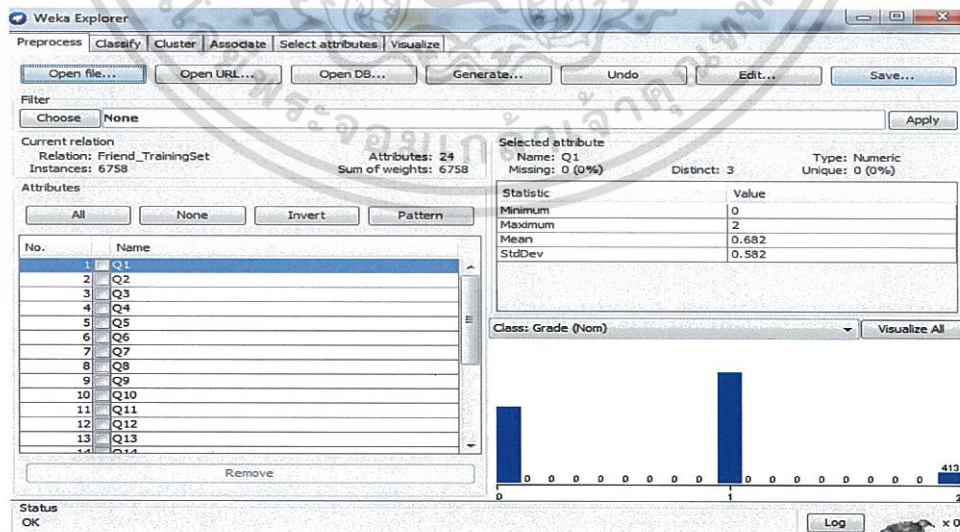
กลุ่มชุดคำตอบประกอบด้วยเพิ่มเอกสาร

- Friend_TrainingSet.csv สำหรับให้โปรแกรมเรียนรู้ข้อมูล
- Friend_TestSet.csv สำหรับใช้ทดสอบอัลก
- อัลกอริทึมที่ได้เลือกมาใช้งาน
- Friend_PredictionSet.csv สำหรับใช้ในการเป็นข้อมูลให้โปรแกรม

ทำนายผลลัพธ์ออกมา

2. จำแนกกลุ่มของผู้ตอบแบบสอบถามโดย WEKA

2.1 ทำการนำเพิ่มข้อมูล Friend_TrainingSet.csv เข้าสู่โปรแกรม WEKA



รูปที่ 4.4 รูปการแสดงผลข้อมูลหลังจากนำข้อมูลเข้า WEKA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.4 เป็นการแสดงเนื้อหาหลังจากที่นำเอาเพิ่มเอกสารเข้าสู่ WEKA โดยจะแสดงรายละเอียดต่างๆของเพิ่มเอกสาร โดยสามารถสรุปได้ว่า

จำนวนคำถาม(Attribute) = 23

จำนวนผู้ตอบแบบสอบถาม(Instances) = 6,758 คน

2.2 ทำการคลิกเลือก อัลกอริทึม ในการจำแนกกลุ่มของข้อมูลในที่นี้ผู้เขียนเลือกใช้การจำแนกข้อมูลแบบ Classification Tree โดยใช้วิธี J48 เนื่องจากผลลัพธ์ในการคำนวณคะแนนของแต่ละผู้ตอบแบบสอบถามเป็นตัวอักษร ซึ่ง J48 เหมาะกับค่าดังกล่าว

Classifier output

Correctly Classified Instances	6758	100 %
Incorrectly Classified Instances	0	0 %
Kappa statistic	1	
Mean absolute error	0	
Root mean squared error	0	
Relative absolute error	0 %	
Root relative squared error	0 %	
Coverage of cases (0.95 level)	100 %	
Mean rel. region size (0.95 level)	33.3333 %	
Total Number of Instances	6758	

--- Detailed Accuracy By Class ---

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	1	A
1	0	1	1	1	1	1	C
1	0	1	1	1	1	1	B
Weighted Avg.	1	0	1	1	1	1	

--- Confusion Matrix ---

a	b	c	<-- classified as
6661	0	0	a = A
0	12	0	b = C
0	0	85	c = E

รูปที่ 4.5 ผลลัพธ์จากการเลือกอัลกอริทึมในการจำแนกข้อมูล

รูปที่ 4.5 เป็นผลลัพธ์ของอัลกอริทึม เพื่อการจำแนกข้อมูลโดยใช้วิธี Classification Tree โดยใช้วิธี J48 โดยสามารถอธิบายได้ว่า

A คือ เมนูสำหรับการเลือกวิธีในการจำแนกข้อมูล

B คือเมนูเลือกอัลกอริทึมในการจำแนกข้อมูล โดยเลือกอัลกอริทึม J48

C คือเมนูเลือกสำหรับการทำ Training Set ให้กับระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

D คือผลลัพธ์หลังจากทำการประมวลผลโดยสามารถทำการบันทึกผลลัพธ์เป็นแฟ้มเอกสารได้

E คือผลลัพธ์หลังจากทำการประมวลผลโดยแสดงค่าในการจำแนกข้อมูลเป็นตัวอักษร

```

==== Output for Training Set ====
Scheme:   weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Friend_TrainingSet
Instances: 6758
Attributes: 23
== Classifier model (full training set) ==
J48 pruned tree
-----
Point <= 0.655: A (6661.0)
Point > 0.655
| Point <= 1.276: B (85.0)
| Point > 1.276: C (12.0)
Number of Leaves :      3
Size of the tree :      5
Time taken to build model: 0.07 seconds
==== Evaluation on training set ====
==== Summary ====
Correctly Classified Instances 6758      100 %
Incorrectly Classified Instances 0        0 %
Kappa statistic 1
Mean absolute error 0
Root mean squared error 0
Relative absolute error 0 %
Root relative squared error 0 %
Coverage of cases (0.95 level) 100 %
Mean rel. region size (0.95 level) 33.3333 %
Total Number of Instances 6758
==== Detailed Accuracy By Class ====
TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
1         0         1         1         1         1         A
1         0         1         1         1         1         C
1         0         1         1         1         1         B
Weighted Avg. 1         0         1         1         1         1
==== Confusion Matrix ====
a  b  c  <-- classified as
6661  0  0 | a = A
0  12  0 | b = C
0  0  85 | c = B

```

รูปที่ 4.6 ผลลัพธ์จากการนำข้อมูลเข้า WEKA

จากรูปที่ 4.6 เป็นข้อมูลผลลัพธ์ที่ได้้นำแฟ้มเอกสารฐานข้อมูลที่ได้จัดทำไว้นำเข้าสู่โปรแกรม WEKA โดยอธิบายได้ว่า

โปรแกรมประยุกต์ทำการประมวลผลโดยพิจารณาที่ Attribute ที่ยี่สิบสองของตารางแล้วทำการกักระดับของคะแนนออกเป็น 3 ช่วงคะแนนโดย

Point <= 0.655 : A (6,661.0) คะแนนที่เท่ากับหรือน้อยกว่า 0.655 คะแนนจะถูกจัดอยู่ในกลุ่ม A

Point <= 1.276 : B (85.0) คะแนนที่เท่ากับหรือน้อยกว่า 1.276 คะแนนจะถูกจัดอยู่ในกลุ่ม B

Point > 1.276 : C (12.0) คะแนนที่มากกว่า 1.276 คะแนนจะถูกจัดอยู่ในกลุ่ม C

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Correctly Classified Instances 6,758 100 %

หมายถึงความถูกต้อง ในการจำแนกจำนวนผู้ตอบแบบสอบถามทั้งหมด 6758 คนมีความถูกต้องคิดเป็น 100%

- Incorrectly Classified Instances 0 0 %

หมายถึงจำนวนและความไม่ถูกต้องในการจำแนกข้อมูลโดย ข้อมูลในการสอนให้กับระบบไม่พบความผิดปกติ

- Confusion Matrix

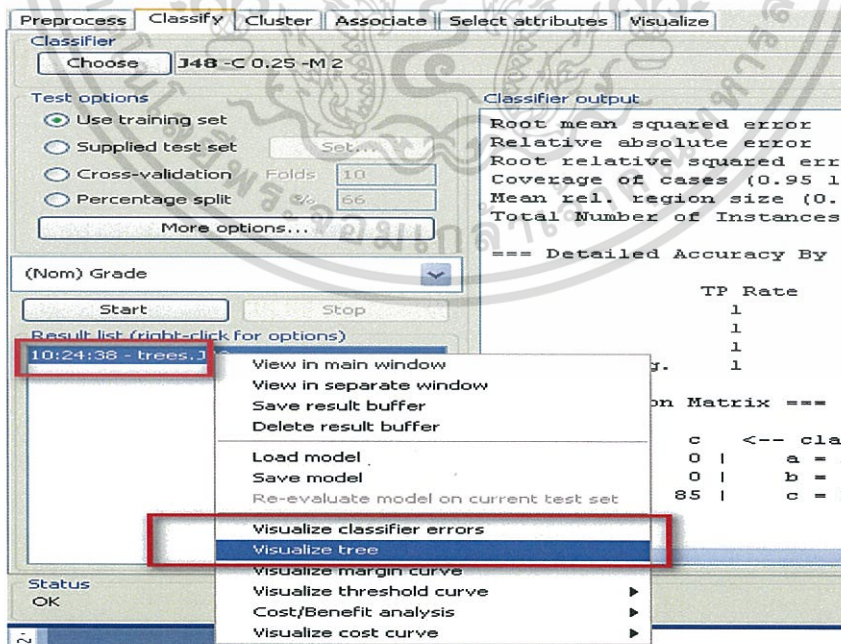
หมายถึง ตารางการจำแนกข้อมูลที่นำเข้าไปรแกรมประยุกต์และ โปรแกรมประยุกต์ได้ทำการแบ่งกลุ่มมาให้โดย

กลุ่ม A มีจำนวน 6,661 คน

กลุ่ม B มีจำนวน 12 คน

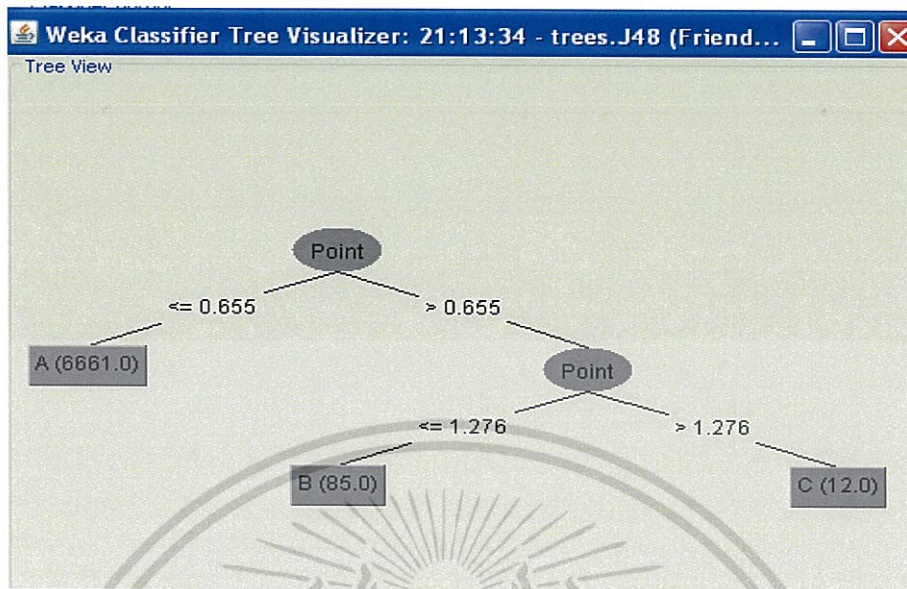
กลุ่ม C มีจำนวน 85 คน

2.3 โปรแกรมประยุกต์ WEKA สามารถที่จะแสดงผลลัพธ์ในรูปแบบรูปภาพได้ โดยแสดงออกมาในรูปแบบของต้นไม้ตัดสินใจ โดยทำการนำเอาผลลัพธ์ที่โมเดลแล้วเลือกการแสดงผลแบบ Visualize Tree ดังรูปที่ 4.7



รูปที่ 4.7 การเลือกให้ WEKA แสดงผลลัพธ์ออกมาเป็นรูปภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



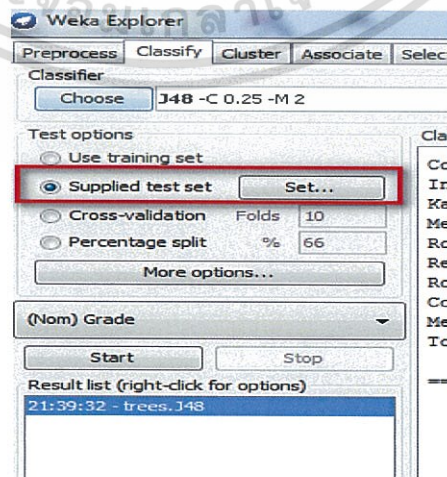
รูปที่ 4.8 ผลลัพธ์ที่ได้จาก โมเดลต้นไม้แสดงแบบ Tree

พิจารณาจากรูป 4.8 สามารถอธิบายได้ว่า

- ผู้ที่ถูกจัดอยู่ในกลุ่ม A มีจำนวน 6,661 คน โดยมีคะแนนพฤติกรรมเท่ากับหรือน้อยกว่า 0.665 คะแนนลงมา
- ผู้ที่ถูกจัดอยู่ในกลุ่ม B มีจำนวน 85 คน โดยมีคะแนนพฤติกรรมเท่ากับหรือน้อยกว่า 1.276 คะแนนลงมา
- ผู้ที่ถูกจัดอยู่ในกลุ่ม C มีจำนวน 12 คน โดยมีคะแนนพฤติกรรมมากกว่า 0.665 คะแนนขึ้นไป

2.4 ทำการนำข้อมูลเพิ่มเอกสาร Friend_TestSet.csv เข้าสู่ WEKA เพื่อเข้าสู่

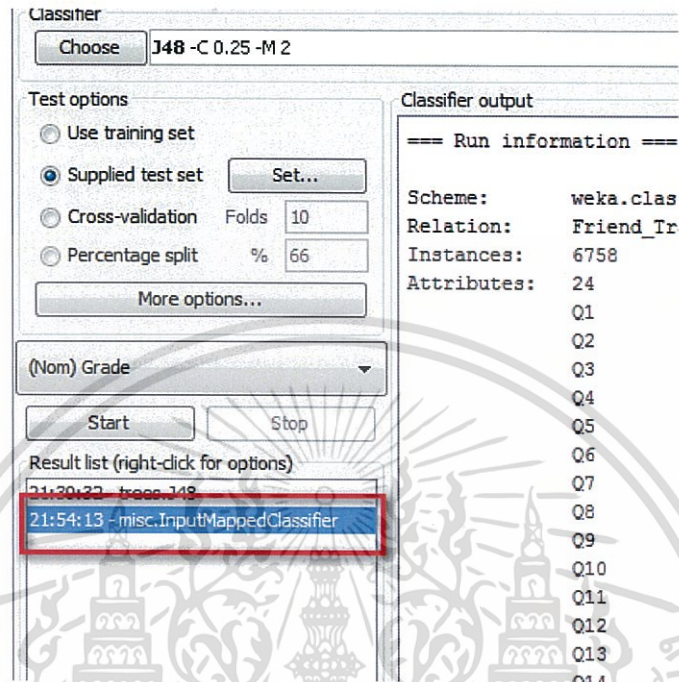
กระบวนการ Test โดย



รูปที่ 4.9 เลือกเมนูการ Test

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.9 ทำการเลือกเมนู Test เพื่อเข้าสู่กระบวนการทดสอบ โมเดลดังกล่าว แล้วเลือกเพิ่มเอกสาร Friend_TestSet.csv แล้วคลิกเมนู Start เพื่อให้ WEKA เริ่มทำการประมวลผล



รูปที่ 4.10 ผลลัพธ์ที่ได้หลังจาก ให้ WEKA ประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

==== Run information ====
Scheme: weka.classifiers.misc.InputMappedClassifier -I -trim -W
       weka.classifiers.trees.J48 -- -C 0.25 -M 2
Relation: Friend_TrainingSet
Instances: 6758
Attributes: 24
          Q1
          Q2
          Q3
          Q4
          Q5
          Q6
          Q7
          Q8
          Q9
          Q10
          Q11
          Q12
          Q13
          Q14
          Q15
          Q16
          Q17
          Q18
          Q19
          Q20
          Q21
          Q22
          Point
          Grade
Test mode: user supplied test set: 2899 instances
==== Classifier model (full training set) ====
InputMappedClassifier:
J48 pruned tree
-----
Point <= 0.655: A (6661.0)
Point > 0.655
| Point <= 1.276: B (85.0)
| Point > 1.276: C (12.0)
Number of Leaves :      3
Size of the tree :      5
Attribute mappings:
Model attributes      Incoming attributes
-----
(numeric) Q1         --> 1 (numeric) Q1
(numeric) Q2         --> 2 (numeric) Q2
(numeric) Q3         --> 3 (numeric) Q3
(numeric) Q4         --> 4 (numeric) Q4

```

รูปที่ 4.11 ข้อมูลผลลัพธ์ Test Set

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

(numeric) Q4 --> 4 (numeric) Q4
 (numeric) Q5 --> 5 (numeric) Q5
 (numeric) Q6 --> 6 (numeric) Q6
 (numeric) Q7 --> 7 (numeric) Q7
 (numeric) Q8 --> 8 (numeric) Q8
 (numeric) Q9 --> 9 (numeric) Q9
 (numeric) Q10 --> 10 (numeric) Q10
 (numeric) Q11 --> 11 (numeric) Q11
 (numeric) Q12 --> 12 (numeric) Q12
 (numeric) Q13 --> 13 (numeric) Q13
 (numeric) Q14 --> 14 (numeric) Q14
 (numeric) Q15 --> 15 (numeric) Q15
 (numeric) Q16 --> 16 (numeric) Q16
 (numeric) Q17 --> 17 (numeric) Q17
 (numeric) Q18 --> 18 (numeric) Q18
 (numeric) Q19 --> 19 (numeric) Q19
 (numeric) Q20 --> 20 (numeric) Q20
 (numeric) Q21 --> 21 (numeric) Q21
 (numeric) Q22 --> 22 (numeric) Q22
 (numeric) Point --> 23 (numeric) Point
 (nominal) Grade --> 24 (nominal) Grade

Time taken to build model: 0.09 seconds

== Predictions on test set ==

inst#	actual	predicted	error	prediction
1	1:A	1:A	1	
2	1:A	1:A	1	
3	1:A	1:A	1	
4	1:A	1:A	1	
5	1:A	1:A	1	
6	1:A	1:A	1	
7	1:A	1:A	1	
8	1:A	1:A	1	
9	1:A	1:A	1	
2889	1:A	1:A	1	
2890	1:A	1:A	1	
2891	1:A	1:A	1	
2892	1:A	1:A	1	

รูปที่ 4.11 ข้อมูลผลลัพธ์ Test Set (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2893	1:A	1:A	1
2894	1:A	1:A	1
2895	1:A	1:A	1
2896	1:A	1:A	1
2897	1:A	1:A	1
2898	1:A	1:A	1
2899	1:A	1:A	1
=== Evaluation on test set ===			
=== Summary ===			
Correctly Classified Instances	2895	99.931 %	
Incorrectly Classified Instances	2	0.069 %	
Kappa statistic	0.9801		
Mean absolute error	0.0005		
Root mean squared error	0.0215		
Relative absolute error	2.1723 %		
Root relative squared error	19.9605 %		
Coverage of cases (0.95 level)	99.931 %		
Mean rel. region size (0.95 level)	33.3333 %		
Total Number of Instances	2897		
Ignored Class Unknown Instances	2		
=== Detailed Accuracy By Class ===			
TP Rate	FP Rate	Precision	Recall F-Measure ROC Area Class
1	0	1	1 1 A
1	0.001	0.667	1 0.8 0.999 C
0.957	0	1	0.957 0.978 0.979 B
Weighted Avg.	0.999	0	1 0.999 0.999 1
=== Confusion Matrix ===			
a	b	c	<-- classified as
2846	0	0	a = A
0	4	0	b = C
0	2	45	c = B

รูปที่ 4.11 ข้อมูลผลลัพธ์ Test Set (ต่อ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Correctly Classified Instances 2,895 99.931 %

หมายถึงความถูกต้อง ในการจำแนกจำนวนผู้ตอบแบบสอบถามทั้งหมด 6,758 คนมีความถูกต้องคิดเป็น 99.931 %

- Incorrectly Classified Instances 2 0.069 %

หมายถึงจำนวนและความไม่ถูกต้องในการจำแนกข้อมูล โดย ข้อมูลในการสอนให้กับระบบมีความผิดพลาด 2 คน คิดเป็นอัตราส่วน 0.069 %

- Confusion Matrix

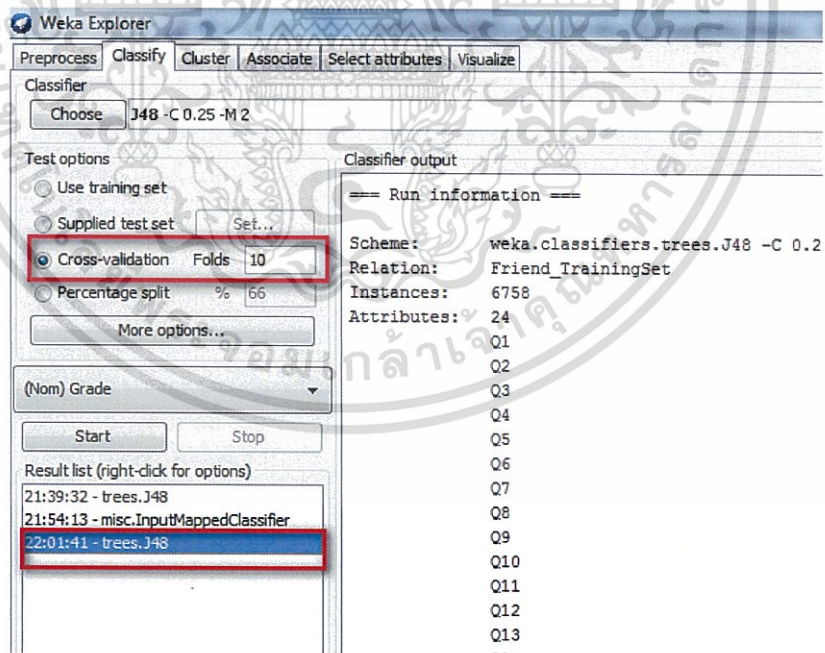
หมายถึง ตารางการจำแนกข้อมูลที่นำเข้าไปโปรแกรมประยุกต์และโปรแกรมประยุกต์ได้ทำการแบ่งกลุ่มมาให้โดย

กลุ่ม A มีจำนวน 2,845 คน

กลุ่ม B มีจำนวน 4 คน

กลุ่ม C มีจำนวน 2 คน

2.5 ทำการกำหนดให้ WEKA ให้ทำการ Cross - Validation เป็นวิธีการวัดประสิทธิภาพในการทำนายตัวอย่างของโมเดลดังรูปที่ 4.12



รูปที่ 4.12 ขั้นตอนการเลือกการวัดประสิทธิภาพในการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

==== Run information ====
Scheme:   weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: Friend_TrainingSet
Instances: 6758
Attributes: 24
Test mode: 10-fold cross-validation
==== Classifier model (full training set) ====
J48 pruned tree
-----
Point <= 0.655: A (6661.0)
Point > 0.655
| Point <= 1.276: B (85.0)
| Point > 1.276: C (12.0)
Number of Leaves :    3
Size of the tree :    5
Time taken to build model: 0.2 seconds
==== Stratified cross-validation ====
== Summary ==
Correctly Classified Instances   6757      99.9852 %
Incorrectly Classified Instances    1      0.0148 %
Kappa statistic                  0.9948
Mean absolute error               0.0001
Root mean squared error           0.0099
Relative absolute error           0.5164 %
Root relative squared error       10.219 %
Coverage of cases (0.95 level)   99.9852 %
Mean rel. region size (0.95 level) 33.3333 %
Total Number of Instances        6758
==== Detailed Accuracy By Class ====
   TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
   1.000    0.000    1.000    1.000    1.000    1.000    A
   0.988    0.000    0.923    0.96    0.94    0.994    C
   0.000    0.988    0.000    0.000    0.000    0.000    B
Weighted Avg. 1.000  0.000  1.000  1.000  1.000  1.000
==== Confusion Matrix ====
 a  b  c  <-- classified as
6661  0  0 | a = A
  0 12  0 | b = C
  0  1 84 | c = B

```

รูปที่ 4.13 ผลลัพธ์จากการ Cross Validation

จากรูปที่ 4.13 เป็นขั้นตอนการตรวจสอบความถูกต้องของโมเดลต้นไม้ โดยการใช้อัลกอริทึม classifiers.trees.J48 -C 0.25 -M 2

- Instances: 6,757 หมายถึง จำนวนตัวอย่างผู้ตอบแบบสอบถาม 6,757 คน
- Attributes: 24 หมายถึง จำนวน Attribute ที่เป็นข้อความและค่าคะแนนและระดับความรุนแรงของผู้ตอบแบบสอบถาม
- Test mode: 10 - fold cross-validation จำนวนรอบในการตรวจสอบความถูกต้อง
- **Correctly Classified Instances** 6,757 99.9852 %
- หมายถึงความถูกต้อง ในการจำแนกจำนวนผู้ตอบแบบสอบถามทั้งหมด 6,757 คน มีความถูกต้องคิดเป็น 99.9852 %
- **Incorrectly Classified Instances** 1 0.0148 %

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หมายถึงจำนวนและความไม่ถูกต้องในการจำแนกข้อมูลโดย ข้อมูลในการสอนให้กับระบบมีความ ผิดปรกติ 1คน คิดเป็นอัตราส่วน 0.0148 %

- **Confusion Matrix**

หมายถึง ตารางการจำแนกข้อมูลที่นำเข้าไปรแกรมประยุกต์และ โปรแกรมประยุกต์ได้ทำการ แบ่งกลุ่มมาให้โดย

กลุ่ม A มีจำนวน 6,661 คน

กลุ่ม B มีจำนวน 84 คน

กลุ่ม C มีจำนวน 12 คน

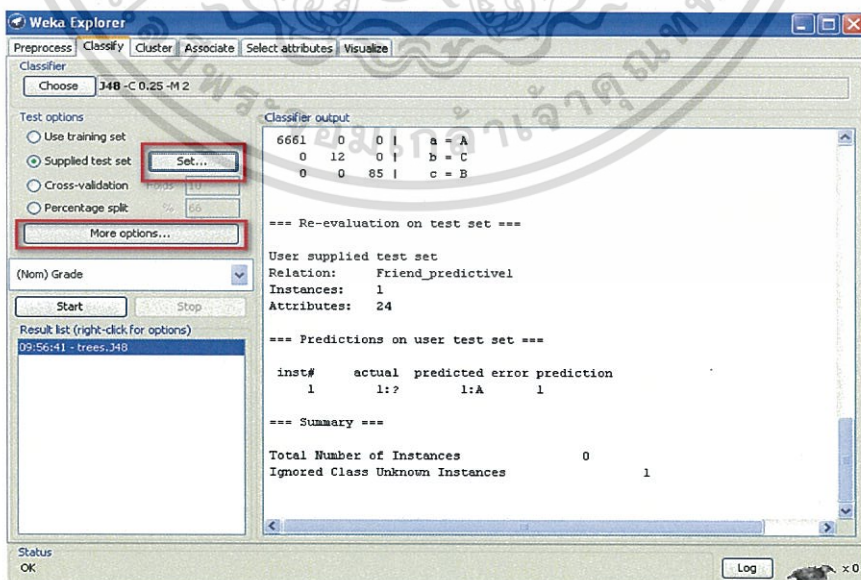
และมีผู้ตอบแบบสอบถาม 1 คน มีอาจถูกจัดอยู่ในระหว่างกลุ่ม B และ กลุ่ม C ด้วย เนื่องจากค่า ROC ของ B มีความน่าเชื่อถือมากกว่า C

2.6 ทำการนำข้อมูลเพิ่มเอกสารที่จะทำการทำนายโดยข้อมูลที่ต้องการทำนายว่า ผู้ทำแบบสอบถามเป็นบุคคลที่อยู่ในกลุ่มใด

- เตรียมข้อมูลเพิ่มเอกสาร โดยไม่ระบุค่า ใน Attribute Grade ดังรูป 4.14

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Point	Grade	
0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0.03	

รูปที่ 4.14 ตารางข้อมูลเพื่อการทำนาย



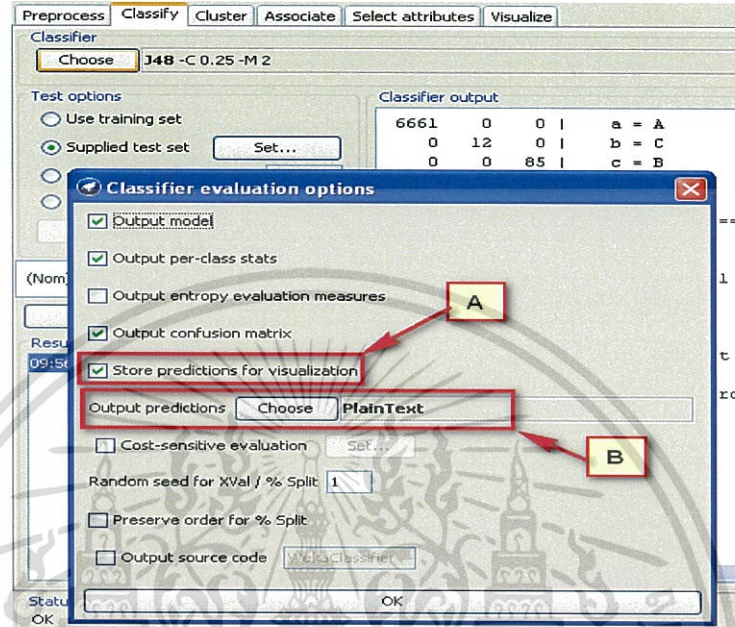
รูปที่ 4.15 ขั้นตอนการเลือกเพิ่มเอกสารที่จะต้องการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- คลิกเลือกที่ปุ่ม Set ดังรูปที่ 4.15 เพื่อทำการเลือกเพิ่มเอกสารที่จะต้องการ

ทำนาย

- เมื่อทำการเลือกเพิ่มเอกสารให้คลิก ปุ่ม More option



รูปที่ 4.16 การกำหนด สำหรับการทำนาย

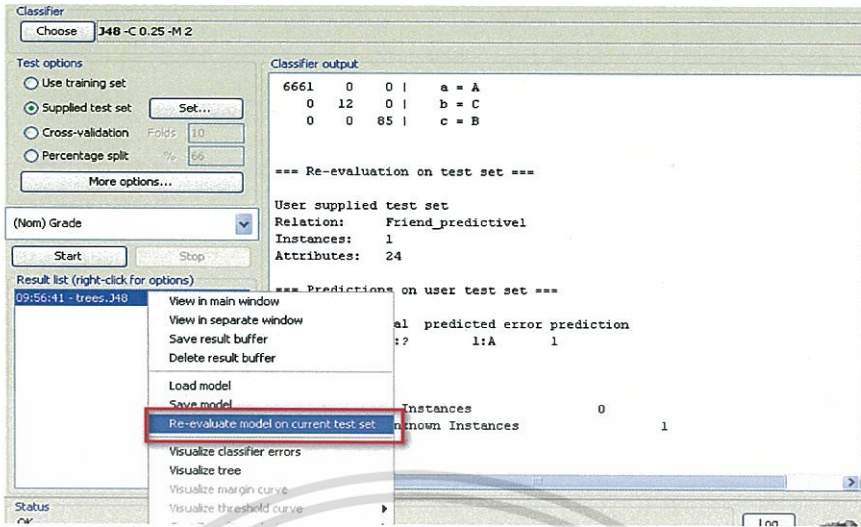
จากรูปที่ 4.16 เป็นขั้นตอนการกำหนดค่าต่างๆเพื่อกำหนดให้โปรแกรมประยุกต์ WEKA แสดงผลออกมาตามที่เรต้องการโดย

A หมายถึง ให้ WEKA ทำการเก็บข้อมูลที่ได้จากการทำนาย

B หมายถึง กำหนดให้ WEKA แสดงผลออกมาในรูปแบบตัวอักษร

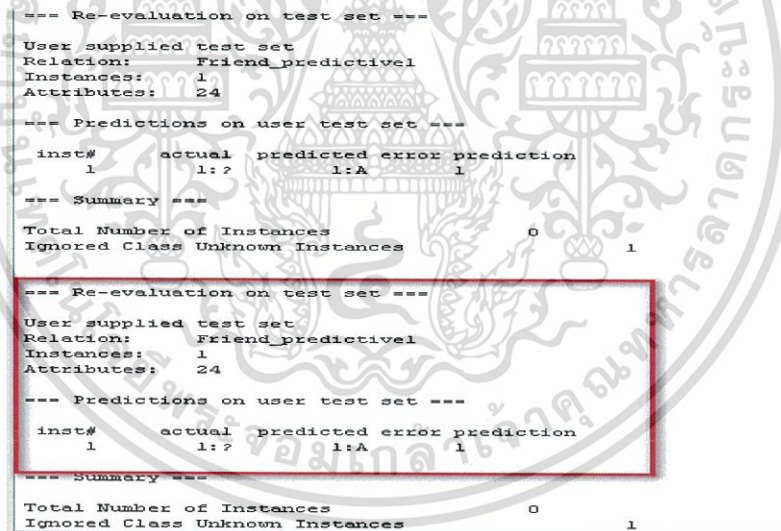
- ทำการคลิกเมาส์ขวาที่ โมเดล ที่ได้จากการ Training ดังรูปที่ 4.16 แล้วเลือก Re-evaluate on current test set และ คลิกปุ่ม Start

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.17 ขั้นตอนการเลือกโมเดลต้นแบบในการทำนาย

- หลังจากทำ WEKA ทำการประมวลผลจะแสดงข้อมูลออกมาดังรูปที่ 4.17



รูปที่ 4.18 ผลลัพธ์ที่ได้จากการทำนาย

จากรูปที่ 4.18 คือผลลัพธ์ที่ได้จากการทำนายโดยอธิบายได้ว่า

Relation: Friend_predictive1 หมายถึง ชื่อแฟ้มเอกสารที่ใช้ในการทำนาย

Instances: 1 หมายถึง จำนวนผู้ตอบแบบสอบถาม

Attributes: 24 หมายถึง จำนวน Attribute ที่เป็นข้อคำถาม และ Grade

โดยการทำนายนี้ ข้อมูลของผู้ตอบแบบสอบถามถูกจัดอยู่ในกลุ่ม A คือ ไม่เสี่ยงต่อการที่จะเกิด

พฤติกรรมความรุนแรง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ข้อมูลที่ใช้กับโปรแกรมประยุกต์

ข้อมูลที่นำมาใช้ในการทดสอบกับโปรแกรมประยุกต์ จะถูกจัดเก็บอยู่ในรูปของ Excel File ประกอบไปด้วยกลุ่มของข้อมูล

- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Training Set
- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Test Set
- เพิ่มเอกสารที่ใช้สำหรับเป็นส่วนของการ Validation Set
- เพิ่มเอกสารที่ใช้สำหรับให้โปรแกรมประยุกต์ทำนาย Predictive Set

4.3 ผลการทดลองและวิเคราะห์

ผู้เขียนได้นำเอาระบบดังกล่าวไปใช้กับนักเรียนระดับชั้นมัธยมศึกษาตอนต้นและตอนปลาย โดยทำการทดสอบกับจำนวนผู้ทดสอบ 300 ตัวอย่าง

ตารางที่ 4.1 ตารางเก็บข้อมูลประวัติการทดสอบระบบ

จำนวนตัวอย่าง	300
จำนวนความถูกต้อง	285
จำนวนตัวอย่างที่ระบบไม่ทราบ	15
เปอร์เซ็นต์ความถูกต้อง	95

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการศึกษาและข้อเสนอแนะ

จากการศึกษาการจำแนกข้อมูล,การทำนายผลลัพธ์ที่เราอาจไม่ทราบค่าโดยใช้วิธีการจำแนกกลุ่มข้อมูล Decision Tree J48 ในโครงการนี้สามารถสรุปผลการดำเนินงานและสรุปผลการทดลองและรวมถึงข้อเสนอนี้

5.1 สรุปผลการศึกษา

การศึกษาในโครงการนี้สามารถสรุปผลดำเนินงานได้ดังนี้

ในการศึกษาข้อมูลพฤติกรรมความรุนแรงที่อาจเกิดขึ้นในสังคมทำให้ทราบสาเหตุที่อาจส่งผลให้คนมีแนวโน้มที่อาจก่อความรุนแรงกับผู้อื่นได้ และถ้าเราสามารถจำแนกกลุ่มคนดังกล่าวได้ อาจช่วยพัฒนาทางด้านความคิด อารมณ์ และอาจช่วยลดอัตราการเกิดความรุนแรงต่อบุคคลลงได้ ในการนำข้อมูลมาใช้งานนั้นจำเป็นต้องทำการจัดเตรียมข้อมูลและแปลงค่าข้อมูลให้เหมาะสมต่อโปรแกรมประยุกต์

จากผลทดลองโดยนำเอาโมเดลดังกล่าวไปใช้งานจริง พบปัญหาหลังจากที่โมเดลได้ทำนายพฤติกรรมออกมาอาจไม่สอดคล้องกับบุคคลทั้งหมด เพราะเนื่องจากการทำนายพฤติกรรมจะต้องอาศัยองค์ประกอบอื่นๆร่วมด้วย เช่น เจตคติ ทัศนคติ อุปนิสัย ครอบครัวยุคสมัย ฯลฯ เนื่องจากเป็นเรื่องที่ละเอียดซับซ้อนและยากต่อการชี้เฉพาะเจาะจงว่าผู้ถูกทำนายนั้นจะเป็นคนที่มีโอกาสเสี่ยงที่จะก่อความรุนแรงหรือไม่

โมเดลนี้สามารถที่จะทำนายพฤติกรรมได้จริง แต่ยังต้องการข้อมูลที่จำเป็นอื่นๆเพื่อมา Validation ให้กับโมเดล เช่น ข้อมูลการกระทำความผิดของกลุ่มอาชญากร ทั้งนี้เพื่อทำให้การทำนายมีประสิทธิภาพมากขึ้น

5.2 ประโยชน์ที่ได้รับจากการศึกษาและพัฒนาโมเดลต้นแบบ

1. การศึกษาโครงการนี้ทำให้เข้าใจถึงทฤษฎีพฤติกรรมมนุษย์ และพฤติกรรมความรุนแรงที่อาจเกิดขึ้นในวัยรุ่น

เอกสารนี้เป็นเอกสาร 2. ทำให้เข้าใจกระบวนการในการดำเนินโครงการตลอดจนขั้นตอนการเตรียมโครงการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ให้สำเร็จ

3. ทำให้ได้โมเดลต้นแบบที่จะช่วยในการทำนายแนวโน้มพฤติกรรมความรุนแรงที่อาจจะเกิดขึ้น

5.3 ข้อเสนอแนะ

1. เพื่อปรับปรุงโมเดลให้มีประสิทธิภาพในการจำแนกข้อมูลเพื่อการทำนายให้ได้ผลลัพธ์ที่มีความแม่นยำและให้ง่ายต่อการใช้งานต่อไปผู้เขียนจึงขอเสนอให้มีการดำเนินงานดังนี้
2. อัลกอริทึมที่จะใช้ในการสร้างโมเดลการทำนายผู้เขียนเล็งเห็นว่าอาจมีโมเดลที่ให้ความแม่นยำในการทำนายเช่น นิวรอนเน็ตเวค (Artificial Neural Network) ครัสเตอร์ริง (Clustering) แอสโซซิเอชันรู (Association Rule)
3. เพิ่มข้อมูลอื่นๆเช่นข้อมูลจากทางครอบครัว จากเพื่อน เพื่อช่วยให้การทำนายเกิดประสิทธิภาพยิ่งขึ้น
4. ผลลัพธ์การทำนายอาจเพิ่มอาจเพิ่มข้อมูลว่าความรุนแรงที่อาจเกิดขึ้นนั้น สาเหตุมาจากสิ่งเร้าใดเป็นเหตุ

บรรณานุกรม

ชูทิพย์ ปานปรีชา. 2551. จิตวิทยาทั่วไปหน่วยที่ 9. พิมพ์ครั้งที่ 21. กรุงเทพฯ: โรงพิมพ์มหาวิทยาลัย
สุโขทัยธรรมมาธิราช.

ทีมข่าวรายงานพิเศษ. 2555. วิเคราะห์จิตลึกลับ"อาชญากรเรียนเก่ง"ผู้โดดเดี่ยว. [Online].

เข้าถึงได้จาก : <http://www.komchadluek.net/detail/20120730/136441/วิเคราะห์จิตลึกลับ'อาชญากรเรียนเก่ง'ผู้โดดเดี่ยว.html>

บุญชม ศรีสะอาด. 2553. วิจัยเบื้องต้น. พิมพ์ครั้งที่ 8. กรุงเทพฯ : สุวีริยาสาส์น.

ประภาเพ็ญ สุวรรณ. 2538. ทักษะการวัดการเปลี่ยนแปลงและพฤติกรรมอนามัย. กรุงเทพฯ:
ไทยวัฒนาพานิช.

แพรภัทร ยอดแก้ว. ม.ป.ป. กลไกป้องกันตัวเอง (Defense Mechanism). [Online].

เข้าถึงได้จาก:<http://gotohnow.org/blog/theories/28371>

ม.ป.ป. ความรุนแรง. [Online]. เข้าถึงได้จาก: <http://dict.longdo.com/search/รุนแรง>

ราชบัณฑิตยสถาน. 2554. ความหมายของพฤติกรรม. [Online]. เข้าถึงได้จาก:

<http://www.royin.go.th/dictionary/search.php>

เรียม ศรีทอง. 2542. พฤติกรรมมนุษย์กับการพัฒนาตน. กรุงเทพฯ : เวิร์ดเวฟเอ็ดดูเคชั่น.

วรารักษ์ ตระกูลสุภะดีและคณะ. 2543. จิตวิทยาการปรับตัว. พิมพ์ครั้งที่ 1. กรุงเทพฯ: สำนักพิมพ์
ศูนย์ส่งเสริมวิชาการ.

วิไลวรรณ ศรีสงครามและคณะ. 2549. จิตวิทยาทั่วไป. พิมพ์ครั้งที่ 1. กรุงเทพฯ: ทริปเฟลด์
เอ็ดดูเคชั่น.

อรอนงค์ อินทรวิจิตร และ นรินทร์ กรินชัย. 2541. ผู้หญิงและเด็ก : เหยื่อความรุนแรงใน
ครอบครัว. กรุงเทพฯ: ปกเกล้าการพิมพ์.

Jiawei Han and Micheline Kamber . **Data Mining Concept and Techniques** . Second Edition .
Elsevier Inc .2006

World Health Organization. (2002). **World report on violence and health** . Geneva : World
Health Organization.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาวจรรุภา ไชยนุรักษ์
วันเดือนปีเกิด	30 กรกฎาคม 2530
ที่อยู่	130/176 ม.1 ต.ลำผักกูด อ.ธัญบุรี จ.ปทุมธานี 12110
ประวัติการศึกษา	ครุศาสตรบัณฑิต สาขาการประถมศึกษา มหาวิทยาลัยเทคโนโลยีราชมงคลธัญบุรี
ประสบการณ์ทำงาน	ตำแหน่งงานปัจจุบัน ข้าราชการครู



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้