การจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณ โดยเทคนิคการทำเหมืองข้อความ

ปรางค์ปราณี ล้ำเลิศพระคุณ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมเว็บ วิทยาลัยครีเอทีฟดีไซน์ แอนด์
เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์
พ.ศ. 2561

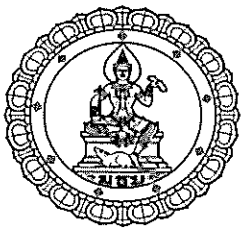
Classification of Helping Abused Children by using Text Mining Techniques

Prangpranee Lamlertphrakhun

A Thesis Submitted in Partial Fulfillment of Requirements for The Degree of Master of Science Program in Web Engineering, College of Creative Design and Entertainment Technology

Dhurakij Pandit University

2018



ใบรับรองวิทยานิพนธ์

วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์

ปริญญา วิทยาศาสตรมหาบัณฑิต

หัวข้อวิทยานิพนธ์ การจำแนกแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณ โดยการประยุกต์ใช้เทคนิคการทำเหมืองข้อความ

เสนอโดย ปรางค์ปราณี ล้ำเลิศพระคุณ

สาขาวิชา วิศวกรรมเว็บ

อาจารย์ที่ปรึกษาวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.อรวรรณ อิ่มสมบัติ

ได้พิจารณาเห็นชอบโดยคณะกรรมการสอบวิทยานิพนธ์แล้ว

.....ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.พิจิตรา จอมศรี)

..... กรรมการและอาจารย์ที่ปรึกษาวิทยานิพนธ์

(ผู้ช่วยศาสตราจารย์ ดร.อรวรรณ อิ่มสมบัติ

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ธนภัทร์ อนุศาสน์อมรกุล)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.วิลาวัลย์ อินทร์ชำนาญ)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.วรสิทธิ์ ชูชัยวัฒนา)

วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอ็นเตอร์เทนเมนต์เทคโนโลยี

..... คณบดี

(ผู้ช่วยศาสตราจารย์ ดร.วรสิทธิ์ ชูชัยวัฒนา)

วันที่ 24 เดือน กรกฎาคม พ.ศ. 2561

หัวข้อวิทยานิพนธ์	การจำแนกการให้ความช่วยเหลือเด็กที่ถูกทารุณ โดยเทคนิคการทำเหมืองข้อความ
ชื่อผู้เขียน	ปรางค์ปราณี ล้ำเลิศพระคุณ
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.อรวรรณ อิ่มสมบัติ
สาขา	วิศวกรรมเว็บ
ปีการศึกษา	2560

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเทคนิคการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมด้วยเทคนิคการทำเหมืองข้อความ โดยการวิเคราะห์ข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็ก ซึ่งอยู่ในรูปของการบันทึกข้อความ ประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมสามารถแบ่งได้ 2 ประเภทใหญ่ๆ คือ การสงเคราะห์เด็ก กับการคุ้มครองสวัสดิภาพเด็ก โดยเอกสารบันทึกเหตุการณ์จะถูกนำมาผ่านการประมวลผลภาษาธรรมชาติก่อนเพื่อสกัดคุณลักษณะซึ่งเป็นคำที่เกี่ยวกับการทำร้ายร่างกาย การล่วงละเมิดทางเพศ การทำร้ายจิตใจ อาวุธบาดแผลและพฤติกรรมที่เกี่ยวข้อง เทคนิคในการจำแนกประเภทที่นำมาทดสอบ มี 3 เทคนิค คือ Decision Tree, Naïve Bayes และ Support Vector Machine จากผลการประเมินพบว่า Decision Tree, Support Vector Machine โดยให้ค่า Precision 92.31%, ค่า Recall 98.34% และ F-score 98.39% ตามลำดับ

Thesis Title	Classification of Helping Abused Children by using Text Mining Techniques
Author	Prangpranee Lamlertphrakhun
Thesis Advisor	Assistant Professor Dr.Aurawan Imsombut
Academic Program	Web Engineering
Academic Year	2018

ABSTRACT

The purpose of this research was to study the classification techniques of helping abused children by using text mining techniques. The data of this research were the records of children abused cases which were in memorandum format. The helping abused children was classified into 2 types: child assistance and child welfare protection. Natural Language Process (NLP) was used for features extraction from case records. The features are the words concerning to physical abuse, sexual abuse, emotional abuse, weapon, wound and related behavior. Three classification techniques used in this study were Decision Tree, Naïve Bayes and Support Vector Machine. Decision Tree and Support Vector Machine showed the best performance with achieved 92.31% precision 98.34% recall and 98.39% F-score.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จลุล่วงได้อย่างสมบูรณ์ โดยได้ได้รับความอนุเคราะห์และการสนับสนุนการดำเนินการวิจัยจาก ผู้ช่วยศาสตราจารย์ ดร. อรวรรณ อิ่มสมบัติ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่กรุณาให้คำแนะนำ ความรู้ และข้อคิดเห็นต่างๆ ที่เป็นประโยชน์ต่อการศึกษาวิจัยในครั้งนี้

ขอกราบขอบพระคุณคณาจารย์สาขาวิศวกรรมเว็บ วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอนเตอร์เทนเมนต์เทคโนโลยี มหาวิทยาลัยธุรกิจบัณฑิตย์ทุกท่าน ที่กรุณาถ่ายทอดความรู้อันเป็นประโยชน์ตลอดการศึกษา

ขอกราบขอบพระคุณ คุณจิตติญา ดุลยวรนันท์ นักสังคมสงเคราะห์ งานสังคมสงเคราะห์ โรงพยาบาลรามาธิบดี ที่ให้ความช่วยเหลือในการติดต่อประสานงานเพื่อเก็บรวบรวมข้อมูลที่ใช้ในการศึกษาวิจัย และคอยให้คำแนะนำที่เป็นประโยชน์ต่อการศึกษาวิจัย

ขอกราบขอบพระคุณเจ้าหน้าที่นักสังคมสงเคราะห์ นักจิตวิทยา และกุมารแพทย์ โรงพยาบาลรามาธิบดีทุกท่าน ที่กรุณาสละเวลาอันมีค่าในการให้ข้อมูล ความรู้ ความคิดเห็น และคำแนะนำในการศึกษาวิจัย

ขอขอบคุณพี่ๆ เพื่อนๆ ทุกคนที่คอยเป็นกำลังใจ และให้ความช่วยเหลือในด้านต่างๆ ในท้ายที่สุดนี้ ขอขอบคุณบิดา มารดาที่ให้กำเนิด เลี้ยงดู ให้การศึกษา และเป็นกำลังใจอันยิ่งใหญ่ที่คอยผลักดันให้การทำวิทยานิพนธ์เล่มนี้สำเร็จลุล่วงไปด้วยดี

ปรางค์ปราณี ล้ำเลิศพระคุณ

สารบัญ

หน้า

บทคัดย่อภาษาไทย..... ฆ

บทคัดย่อภาษาอังกฤษ..... ง

กิตติกรรมประกาศ..... จ

สารบัญตาราง..... ซ

สารบัญภาพ..... ฌ

บทที่

1. บทนำ..... 1

1.1 ที่มาและความสำคัญของปัญหา..... 1

1.2 วัตถุประสงค์ของการวิจัย..... 3

1.3 ขอบเขตการวิจัย..... 3

1.4 สมมติฐานการวิจัย..... 4

1.5 ประโยชน์ที่คาดว่าจะได้รับ..... 4

1.6 นิยามคำศัพท์..... 4

2. แนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง..... 6

2.1 ความรู้เกี่ยวกับการทารุณเด็ก..... 6

2.2 พระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546..... 9

2.3 เหมืองข้อมูล (Data Mining)..... 10

2.4 เทคนิคการจำแนกประเภท (Classification)..... 11

2.5 เหมืองข้อความ (Text Mining)..... 12

2.6 งานวิจัยที่เกี่ยวข้อง..... 17

3. ระเบียบวิธีวิจัย..... 22

3.1 ขั้นตอนการดำเนินงานวิจัย..... 22

3.2 สถาปัตยกรรมการทำงานของระบบ..... 25

3.3 การพัฒนาระบบ Child Protection Clinic..... 37

3.4 เครื่องมือที่ใช้ในการดำเนินงานวิจัย..... 39

.

สารบัญ (ต่อ)

บทที่	หน้า
4. ผลการวิจัย…..…….……	40
4.1 ข้อมูลที่นำมาใช้ในการทดสอบ.....……..……..	40
4.2 ผลการวิจัย.....……..…….……	40
4.3 การวิเคราะห์ผลการทดลอง.....……..…….……	42
4.4 ระบบจำแนกการให้ความช่วยเหลือเด็กที่ถูกทารุณ……..…….….....	42
5. สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ.....……..……	47
5.1 สรุปผลการวิจัย.....……..……..……..	47
5.2 ปัญหาอุปสรรคและข้อจำกัดของงานวิจัย……..….....	47
5.3 ข้อเสนอแนะ……..……..….....	48
บรรณานุกรม……..……..….....	49
ภาคผนวก……..	52
ก. กลุ่มคุณลักษณะและคำที่แสดงคุณลักษณะ……..……....	53
ข. กฎที่ใช้ในการวิเคราะห์ประเภทการให้ความช่วยเหลือ……..….....	82
ค. บทความการประชุมวิชาการ……..…….....	91
ประวัติผู้เขียน……....……....……....…….	98

สารบัญตาราง

ตารางที่ **หน้า**

2.1 ตัวอย่างข้อมูลลูกค้า…….……..…….. 10

2.2 ตัวอย่างการตัดคำ……...……....……..…… 15

2.3 ตัวอย่างการกำจัดคำหยุด…….…...…….. 16

2.4 ตัวอย่างคำพ้องความในภาษาไทย……...……. 16

2.5 ตัวอย่างการวิเคราะห์ชนิดของคำ (Part of speech)……..…… 17

3.1 ตัวอย่างการตัดคำ…….……...…… 26

3.2 ตัวอย่างการกำจัดคำที่ไม่จำเป็นในประโยค (Stop word)……... 26

3.3 ตัวอย่างการหาคำพ้องความ……...…….…… 27

3.4 ชนิดของคำ (POS)……. 28

3.5 ตัวอย่างการกำกับหน้าที่ของคำ……. 30

3.6 กลุ่มหลัก (Main feature groups)…….…….. 31

3.7 กลุ่มย่อย (feature groups)……....…….……....…… 31

3.8 ตัวอย่างการพิจารณาคุณลักษณะร่วมกับคำปฏิเสธ.....…… 35

3.9 ตัวอย่างกฎที่ใช้ในการจำแนกประเภท……..... 36

4.1 ผลการทดสอบประสิทธิภาพของตัวแบบ…….. 40

4.2 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Decision Tree..... 41

4.3 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Naïve Bayes..... 41

4.4 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Support Vector Machine…….. 41

สารบัญภาพ

ภาพที่ **หน้า**

2.1 ตัวอย่างแผนภาพเพื่อการตัดสินใจ..... 12
2.2 ตัวอย่าง Linear model ของ SVM ในการแบ่งข้อมูลออกเป็น 2 กลุ่ม..... 14
3.1 สถาปัตยกรรมการทำงานของระบบ..... 25
3.2 ขั้นตอนการเตรียมข้อมูล..... 25
3.3 การออกแบบหน้าจอ Login..... 37
3.4 การออกแบบหน้าบันทึกข้อมูลทะเบียนผู้ป่วย..... 38
3.5 การออกแบบหน้าบันทึกเคส..... 38
3.6 การออกแบบหน้าบันสรุปผลการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรม..... 39
4.1 หน้าจอ Login..... 43
4.2 หน้าจอหลัก Login..... 43
4.3 หน้าจอแสดงรายการผู้ป่วย..... 43
4.4 หน้าจอการเพิ่มข้อมูลใหม่..... 44
4.5 หน้าจอบันทึกเหตุการณ์..... 44
4.6 หน้าจอประวัติผู้ป่วย..... 44
4.7 หน้าจอข้อมูลครอบครัว/ผู้ดูแล..... 45
4.8 หน้าจอข้อมูลโรงเรียน/ชุมชน..... 45
4.9 หน้าจอข้อมูล Abuser..... 45
4.10 หน้าจอผลการวิเคราะห์ประเภทการให้ความช่วยเหลือ..... 46
4.11 หน้าจอแสดงรายละเอียดการประเมิน..... 46

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การทารุณกรรมเด็ก (Child Abuse) เป็นปัญหาที่เกิดขึ้นทั่วโลก รวมถึงในประเทศไทย เป็นปัญหาใหญ่ของสังคม และถือเป็นปัญหาครอบครัว จำเป็นต้องเร่งหาทางแก้ไขและแนวทางป้องกัน การทารุณกรรมส่งผลกระทบต่อสภาพร่างกายและจิตใจของเด็กเป็นอย่างมาก อาจก่อให้เกิดการพิการ เสียชีวิต หรือปัญหาทางสุขภาพจิตรวมถึงพฤติกรรมเบี่ยงเบนต่างๆ และจะส่งผลต่อการเติบโตเป็นผู้ใหญ่ การใช้ชีวิตในสังคม การอยู่ร่วมกับผู้อื่นและการปฏิบัติต่อผู้อื่น (กาย จิต สังคม) “เด็ก” หมายความว่า บุคคลซึ่งมีอายุต่ำกว่าสิบแปดปีบริบูรณ์ แต่ไม่รวมถึงผู้ที่บรรลุนิติภาวะด้วยการสมรส “การทารุณกรรม” หมายความว่า การกระทำหรือละเว้นการกระทำด้วยประการใดๆ จนเป็นเหตุให้เด็กเสื่อมเสียสเถียรภาพหรือเกิดอันตรายแก่ร่างกายหรือจิตใจ การกระทำผิดทางเพศต่อเด็ก การใช้เด็กให้กระทำหรือประพฤติในลักษณะที่น่าจะเป็นอันตรายแก่ร่างกายหรือจิตใจ หรือขัดต่อกฎหมายหรือศีลธรรมอันดี ทั้งนี้ไม่ว่าเด็กจะยินยอมหรือไม่ก็ตาม (พระราชบัญญัติคุ้มครองเด็ก พ.ศ.2546, 2546) รูปแบบความรุนแรงที่เด็กได้รับส่วนมากคือ การถูกทำให้บาดเจ็บทางกาย รองลงมาคือความรุนแรงทางเพศ ทางจิตใจและการถูกปล่อยปละละเลยตามลำดับ (นันทกา สวัสดิพานิช, 2555: 25) เนื่องจากเด็กที่ถูกทารุณจะได้รับผลกระทบทั้งทางร่างกาย จิตใจ และการใช้ชีวิตในสังคม ดังนั้นจึงต้องอาศัยความร่วมมือจากบุคลากรหลายฝ่ายในการเยียวยารักษา รวมถึงการให้ความช่วยเหลือในด้านต่างๆ ซึ่งเป็นการทำงานในลักษณะสหวิชาชีพ

กฎหมายเป็นเครื่องมือสำคัญในการกำหนดแนวทางการให้ความช่วยเหลือต่อเด็กที่ถูกกระทำทารุณ และเป็นมาตฐานในการดำเนินการลงโทษผู้กระทำผิด และเป็นข้อปฏิบัติสำหรับผู้ปกครอง ผู้ดูแลเด็ก รวมถึงเจ้าหน้าที่ทุกคนที่เกี่ยวข้องในการปฏิบัติได้อย่างถูกต้องในทิศทางเดียวกัน เพื่อประโยชน์สูงสุดของเด็ก ดังนั้นนักสังคมสงเคราะห์ และทีมสหวิชาชีพที่เกี่ยวข้อง

จำเป็นต้องใช้กฎหมายเป็นพื้นฐานในการดำเนินการ ตัดสินใจในการให้การสงเคราะห์หรือคุ้มครอง รวมถึงการให้ความช่วยเหลือในด้านต่างๆ

พระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 เป็นกฎหมายหนึ่งที่เป็นพื้นฐานหลักในการดำเนินการหาแนวทางการให้ความช่วยเหลือต่างๆ โดยจะกระทำตามกรอบกฎหมายและคำนึงถึงประโยชน์ของเด็กเป็นสำคัญ นักสังคมสงเคราะห์และทีมสหวิชาชีพจำเป็นต้องรู้จักข้อกฎหมายเพื่อประโยชน์ต่อการดำเนินการ และให้ความช่วยเหลือที่รวดเร็ว

จากระบบการทำงานในปัจจุบัน ยังไม่มีระบบสารสนเทศเข้ามาช่วยจัดเก็บข้อมูล ทำให้การทำงานขึ้นอยู่กับตัวบุคคล คือเจ้าหน้าที่นักสังคมสงเคราะห์ที่เป็นผู้ดูแลจะจดจำข้อมูลของเด็กที่ดูแลได้ แต่หากมีบุคคลอื่นเข้ามาปฏิบัติงานต่ออาจจะต้องมาค้นหาข้อมูลใหม่ และวิธีการดำเนินการตัดสินใจต่างๆ อาจเปลี่ยนไปตามตัวบุคคล การจดบันทึกข้อมูลลงในกระดาษโดยไม่ได้มีการจัดเก็บอย่างเป็นระบบทำให้เกิดข้อมูลสูญหาย และจะใช้ความจำจากคนในการทำงานแทน ผู้วิจัยจึงเห็นว่าหากนำเทคโนโลยีเว็บแอปพลิเคชันเข้ามาช่วยในการจัดเก็บข้อมูลจะช่วยป้องกันการสูญหายของข้อมูลได้ และยังช่วยให้เกิดความสะดวกในการติดตามข้อมูล การค้นหาข้อมูล รวมถึงการติดตามการปฏิบัติงานของเจ้าหน้าที่ผู้รับผิดชอบ และเก็บข้อมูลได้อย่างครบถ้วนสมบูรณ์

ข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็ก เช่นการรับแจ้งเหตุ การสัมภาษณ์เด็กการสัมภาษณ์ผู้ปกครอง ความเห็นของแพทย์ และอื่นๆ ยังมีความสำคัญในการใช้ประกอบการพิจารณาการให้ความช่วยเหลือว่าจะเป็นไปในแนวทางสงเคราะห์หรือคุ้มครอง เนื่องจากกระบวนการให้ความช่วยเหลือเด็กตาม พรบ. คุ้มครองเด็ก พ.ศ. 2546 มีข้อกำหนดเรื่องอำนาจของเจ้าหน้าที่การให้ความช่วยเหลือ การดำเนินการทางกฎหมายที่แตกต่างกันโดยแบ่งตามประเภทของเด็กที่พึงได้รับความสงเคราะห์ กับเด็กที่พึงได้รับความคุ้มครอง ดังนั้นการวิเคราะห์เอกสารจากบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็กจึงเป็นข้อมูลสำคัญที่ช่วยในการตัดสินใจของเจ้าหน้าที่ และผู้เกี่ยวข้องทุกฝ่าย ในการตัดสินใจดำเนินการให้ความช่วยเหลือ หรือดำเนินการตามกฎหมาย ทีมสหวิชาชีพจะร่วมกันปรึกษาเสนอความคิดเห็นเพื่อหาแนวทางให้ความช่วยเหลือ โดยดูจากลักษณะของเด็กว่าควรสงเคราะห์ หรือคุ้มครองเป็นอันดับแรก ซึ่งบ่อยครั้งที่มีความเห็นต่างกันตามประสบการณ์ และมุมมองของแต่ละคน ทำให้บางครั้งการดำเนินการเกิดความล่าช้า เช่นในบางกรณีที่ต้องคุ้มครองเด็ก และต้องเร่งจับกุมตัวผู้กระทำผิด แต่เนื่องจากยังไม่สามารถตัดสินใจได้ว่า

ควรเป็นการคุ้มครองหรือไม่ ทีมงานและผู้เกี่ยวข้องจึงไม่สามารถเข้าแจ้งความและดำเนินคดีได้ ทำให้ผู้กระทำผิดยังไม่ถูกจับกุม ก่อให้เกิดความหวาดกลัวต่อเด็ก ส่งผลให้เด็กต้องถูกแยกจากครอบครัว และส่งผลต่อสภาพจิตใจของเด็ก รวมถึงผู้ก่อเหตุอาจหลบหนีไปจนไม่สามารถจับกุมได้

ผู้วิจัยเห็นว่าหากสามารถนำข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องมาใช้ประมวลผลได้จะเกิดประโยชน์ต่อการพิจารณาดำเนินการให้ความช่วยเด็กเป็นอย่างมาก จึงได้มองเห็นถึงประโยชน์ของเหมืองข้อความที่สามารถสกัดเอาความรู้ หรือสารสนเทศที่สำคัญมาใช้ประโยชน์ในด้านจำเพาะเจาะจง ประกอบกับการบันทึกข้อมูลของเด็กที่ถูกทารุณจะอยู่ในรูปแบบของบันทึกข้อความซึ่งเป็นเอกสารที่ไม่มีโครงสร้าง การค้นหา การวิเคราะห์ข้อมูลต่างๆ ทำได้ยากและเสียเวลา จึงเห็นว่าหากนำเทคนิคเหมืองข้อความมาประยุกต์ใช้ จะทำให้เกิดประโยชน์ในด้านนี้มาก รวมทั้งเทคโนโลยีเว็บแอปพลิเคชันซึ่งเป็นการเขียนโปรแกรมบนเว็บที่ช่วยให้การจัดเก็บข้อมูลเอกสารที่เกี่ยวข้องกับการทารุณกรรมเด็ก มีการจัดการข้อมูลที่เป็นระบบ มีมาตรฐานมากขึ้น ป้องกันการสูญหายของข้อมูล ผู้ใช้งานที่เกี่ยวข้องหลายๆ ฝ่ายสามารถเข้าใช้งานได้ทุกที่ ทุกเวลา ผู้วิจัยจึงคิดพัฒนาการประยุกต์ใช้เทคนิคเหมืองข้อความสำหรับการวิเคราะห์หาแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณ

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อวิเคราะห์เอกสารที่เกี่ยวข้องกับการทารุณเด็กที่ได้รับแจ้ง และจำแนกประเภทตามลักษณะการช่วยเหลือ
2. เพื่อพัฒนาระบบที่ช่วยในการวิเคราะห์และจำแนกข้อมูลบันทึกที่เกี่ยวข้องกับเด็กที่ถูกทารุณ

1.3 ขอบเขตการวิจัย

1. ศึกษาเอกสารที่เกี่ยวข้องกับการทารุณกรรมเด็ก อายุ 0-18 ปี ซึ่งแบ่งประเภทการทารุณเป็น 2 ประเภท ได้แก่ การทำร้ายร่างกาย (Physical Abuse) จำนวน 30 คน และการทารุณทางเพศ (Sexual Abuse) จำนวน 30 คน
2. ข้อมูลที่นำมาใช้ในการศึกษาเป็นข้อมูลผู้ป่วยเด็กของคณะแพทย์ศาสตร์โรงพยาบาลรามาธิบดีมหาวิทยาลัยมหิดลที่อยู่ในความดูแลของนักสังคมสงเคราะห์ของโรงพยาบาล

3. กฎหมายที่นำใช้ในการศึกษาเป็นพระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 หมวด 3 เกี่ยวกับการสงเคราะห์เด็ก และหมวด 4 เกี่ยวกับการคุ้มครองสวัสดิภาพเด็ก เนื่องจากหมวดที่ 3 และ 4 ได้มีการกำหนดลักษณะของเด็กที่ควรได้รับการสงเคราะห์กับการคุ้มครอง หน้าที่ของเจ้าหน้าที่ผู้เกี่ยวข้องในการดำเนินการให้ความช่วยเหลือ ข้อปฏิบัติ และข้อห้ามต่างๆ

1.4 สมมติฐานการวิจัย

การวิเคราะห์หาแนวทางให้ความช่วยเหลือเด็กที่ถูกทารุณโดยการประยุกต์ใช้เทคนิคเหมืองข้อความ สามารถเป็นเครื่องมือช่วยให้การทำงานของนักสังคมสงเคราะห์กลุ่มเด็กที่ถูกทารุณและทีมสหวิชาชีพในการวิเคราะห์ และสามารถใช้งานได้สะดวก

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นเครื่องมือช่วยในการทำงานของนักสังคมสงเคราะห์

2. สามารถวิเคราะห์เอกสารที่เกี่ยวข้องกับการทารุณเด็กเพื่อนำมาใช้เป็นแนวทางในการให้ความช่วยเหลือ และการดำเนินการต่างๆ เพื่อให้เกิดความสะดวก รวดเร็วในการตัดสินใจ

1.6 นิยามคำศัพท์

การทำร้ายร่างกาย (Physical Abuse) หมายถึง การกระทำให้เกิดบาดแผลบนร่างกายเด็กโดยความตั้งใจ เช่นการทุบตี หรือการกระทำที่อาจจะก่อให้เกิดอันตรายต่อเด็ก เช่นการกักขัง (พระราชบัญญัติคุ้มครองเด็ก พ.ศ.2546, 2546)

การทารุณทางเพศ (Sexual Abuse) หมายถึง การกระทำกิจกรรมทางเพศต่อเด็ก โดยที่เด็กไม่สามารถให้ความยินยอมพร้อมใจหรือมีวุฒิภาวะเพียงพอที่จะเข้าใจการกระเหล่านั้น เป็นการกระทำที่ตอบสนองความพึงพอใจของผู้กระทำ (พระราชบัญญัติคุ้มครองเด็ก พ.ศ.2546, 2546)

ทีมสหวิชาชีพ หมายถึง บุคคลาที่มีความรู้ความชำนาญในแต่ละสาขาวิชาชีพ สำหรับการวิเคราะห์หาแนวทางให้ความช่วยเหลือเด็กที่ถูกทารุณ เป็นการทำงานที่ต้องอาศัยบุคคลากรจากหลายสาขาอาชีพร่วมกันในการดำเนินการและตัดสินใจ ซึ่งเป็นการทำงานร่วมกันของบุคคลากรจากหลายสาขาวิชาชีพ

การสงเคราะห์เด็ก หมายถึง การให้ความช่วยเหลือ อุปการะเลี้ยงดู ให้การศึกษาแต่เด็กที่ครอบครัวหรือผู้ดูแลไม่สามารถเลี้ยงดูได้ หรือเด็กกำพร้า เด็กเร่ร่อน เด็กพิการ

การคุ้มครองสวัสดิภาพเด็ก หมายถึง การแยกตัวเด็กออกจากครอบครัว หรือบ้านที่อาศัยที่ไม่เหมาะสม โดยไม่จำเป็นต้องได้รับการยินยอมจากผู้ปกครอง การสืบเสาะพินิจเด็กและครอบครัว และการดำเนินคดีความในบางกรณี

บทที่ 2
แนวคิด ทฤษฎี และผลงานวิจัยที่เกี่ยวข้อง

การวิจัยนี้เป็นการนำเทคนิคเหมืองข้อความมาใช้เป็นเครื่องมือในการวิเคราะห์หาแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณโดยประเมินจาก ประเภทการทารุณกรรมที่เกิดกับเด็ก ลักษณะที่ถูกทำร้าย ความสัมพันธ์ของเด็กกับผู้กระทำความรุนแรง (Abuser) และข้อกฎหมายที่เกี่ยวข้อง เพื่อเป็นแนวทางในการประเมินวิธีการให้ความช่วยเหลือตามพระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 โดยผู้วิจัยได้ศึกษาแนวคิดทฤษฎี เอกสาร และงานวิจัยที่เกี่ยวข้อง เพื่อนำมาประยุกต์เป็นแนวทางในการวิจัย โดยครอบคลุมหัวข้อต่างๆ มีรายละเอียดดังนี้

2.1 ความรู้เกี่ยวกับการทารุณเด็ก

2.1.1 ประเภทของการทารุณเด็ก

1. การทำร้ายร่างกาย (Physical abuse) คือ การกระทำที่ก่อให้เกิดอันตราย หรืออาจก่อให้เกิดอันตรายต่อเด็ก เช่น เตะ ต่อย ตี การจับทารกมาเขย่าตัว

2. ความรุนแรงทางเพศ (Sexual abuse) หรือการล่วงละเมิดทางเพศ คือการนำเด็กไปร่วมกิจกรรมทางเพศ หรือบังคับให้เด็กทำอนาจาร โดยที่เด็กยังไม่มีความเข้าใจหรือยังไม่มีวุฒิภาวะเพียงพอที่จะสามารถตัดสินใจได้อย่างเหมาะสม

3. การทำร้ายจิตใจ (Emotion abuse) คือ การกระทำที่ทำให้เกิดความกระทบกระเทือนหรือเกิดบาดแผลทางจิตใจของเด็ก โดยประเมินจากผลกระทบทางจิตใจและพฤติกรรมที่ผิดปกติของเด็ก เช่น เกิดความรู้สึกหวาดกลัวผู้กระทำ มีความโกรธและอยากทำร้ายผู้กระทำ มองว่าตัวเองไร้ค่า ทำร้ายตัวเอง นอนไม่หลับ ฝันร้าย ซึ่งบาดแผลทางจิตใจจะรบกวนต่อการพัฒนาทางร่างกาย จิตใจ จิตวิญญาณ คุณธรรมจริยธรรม ตลอดจนพัฒนาการทางด้านสังคมของเด็ก เป็นการทารุณที่ส่งผลเสียในระยะยาว และประเมินผลได้ยากที่สุด

4. การทอดทิ้ง ปล่อยปละละเลย (Neglect) คือ การที่ครอบครัวผู้ปกครอง หรือผู้ดูแลเด็กปล่อยปละละเลยไม่ได้ดูแลเด็กตามมาตฐานที่เด็กควรได้รับ เช่น ด้านปัจจัยสี่ การศึกษา การไม่ได้รับการสนองทางอารมณ์อย่างเหมาะสม ไม่มีความปลอดภัยในชีวิต ซึ่งอาจทำให้เกิดอันตรายหรือเสี่ยงต่ออันตราย ทั้งทางด้านร่างกาย จิตใจ พัฒนาการ

2.1.2 ผลกระทบจากการที่เด็กถูกทารุณ

เด็กที่ถูกทารุณจะได้รับผลกระทบทั้งด้านกาย จิต สังคม ซึ่งผลกระทบบางอย่างสามารถรักษาให้หายได้ แต่บางอย่างจะเป็นไปตลอดชีวิต ผลกระทบที่เกิดขึ้นกับเด็กที่ถูกทารุณมีดังนี้

1. ผลกระทบด้านกาย เช่น เด็กที่ถูกทำร้ายร่างกาย จะทำให้เด็กมีอาการบาดเจ็บ หากไม่รุนแรงเกินไปสามารถรักษาให้หายได้หรือมีรอยแผลเป็น ไปจนถึงการพิการ หรือถึงขั้นเสียชีวิต นอกจากการบาดเจ็บที่เกิดกับร่างกายแล้ว ยังรวมถึงความผิดปกติที่อาจเกิดกับสมองหากสมองได้รับความกระทบกระเทือนอย่างหนัก จะส่งผลต่อกระบวนการคิด การรับรู้ การตีความการสื่อสาร และการตอบสนองต่อเหตุการณ์ต่างๆ สำหรับเด็กที่ถูกล่วงละเมิดทางเพศ อาจนำไปสู่การตั้งครรภ์ที่ไม่พึงประสงค์และโรคติดต่อต่างๆ

2. ผลกระทบด้านจิต เด็กที่ถูกทารุณจะเกิดผลกระทบทั้งด้านจิตใจและพฤติกรรม เป็นผลกระทบที่ส่งผลต่อเด็กในระยะยาวจนถึงวัยผู้ใหญ่และอาจกลายเป็นพฤติกรรมติดตัวเมื่อโตขึ้น อาจส่งผลต่อการเสื่อมถอยทางศีลธรรม จริยธรรมของเด็ก และอาจนำไปสู่ปัญหายาเสพติดต่อไป ในส่วนของผลกระทบด้านจิตใจเด็กจะไม่สามารถมีความผูกพันที่มั่นคงหรือไว้วางใจใครได้ ทำให้มีปัญหาในการสร้างสัมพันธ์กับคนอื่นตั้งแต่วัยเด็กไปจนถึงวัยผู้ใหญ่ เช่นไม่สามารถเข้าสังคมได้ หวาดระแวง เกิดความรู้สึกในแง่ลบต่อตัวเอง มีความยอมรับนับถือตนเองต่ำ ขาดความเชื่อมั่นในตนเอง การขาดความรัก คิดว่าตนเองมีมลทินอารมณ์แปรปรวน ปราศจากความหวังในชีวิต เกิดภาวะซึมเศร้า และการคิดอยากฆ่าตัวตาย ในด้านของพฤติกรรม เด็กอาจมีพฤติกรรมที่โตเกินวัย การยั่วยวนหรือเสพติดทางเพศหรือล่วงเกินทางเพศผู้อื่น การซ้ำเติมตนเอง ก้าวร้าว และอาจมีแนวโน้มที่จะกลายเป็นคนที่ใช้ความรุนแรงกับผู้อื่นเมื่อโตเป็นผู้ใหญ่

3. ผลกระทบด้านสังคม เป็นผลกระทบต่อการใช้ชีวิตในสังคมของเด็กทุกด้าน เช่น การเข้าสังคมของเด็ก เนื่องจากเด็กจะมีการหวาดระแวง อาจทำให้เกิดการเก็บตัว แยกตัวออกจากสังคม รวมถึงไม่สามารถใช้ชีวิตในสังคมได้อย่างปกติ อาจเจอกับปัญหาการไม่เป็นที่ยอมรับ เป็นต้น

2.1.3 การรวบรวมข้อมูลเกี่ยวกับการทารุณเด็ก

เป็นการรวบรวมข้อมูลเพื่อนำไปสู่วิธีการให้ความช่วยเหลือ โดยการสืบค้นข้อมูลจากหลายทาง เช่น การสัมภาษณ์ผู้เกี่ยวข้อง การสอบถามจากตัวเด็กโดยตรง การเยี่ยมบ้าน การสังเกตุพฤติกรรมผู้ดูแลหรือผู้ที่เกี่ยวข้อง การสำรวจสภาพแวดล้อมกับสภาพสังคมของเด็ก ข้อมูลจากผลการตรวจร่องรอยการถูกทำร้ายบนร่างกาย ผลการตรวจสภาพจิต การตรวจทางห้องปฏิบัติการ และจากแหล่งอื่นๆ

2.1.4 แนวทางการให้ความช่วยเหลือ

การให้ความช่วยเหลือเด็กที่ถูกทารุณมีเป้าหมายเพื่อป้องกันมิให้ถูกกระทำซ้ำให้ความช่วยเหลือตามความเหมาะสม บำบัดผลกระทบที่เกิดขึ้น และตอบสนองต่อความต้องการทางอารมณ์และสังคมของเด็กในระยะยาว โดยแบ่งตามกระบวนการได้ดังนี้

1. กระบวนการคุ้มครองเด็ก (Child protection process) กระบวนการคุ้มครองเด็กเป็นการดำเนินการสงเคราะห์หรือให้ความคุ้มครองเด็กตามความเหมาะรวมไปถึงการสืบค้นข้อมูลเพิ่มเติมการประเมินต่างๆ การวางแผนการให้ความช่วยเหลือในระยะยาว ตลอดจนการดำเนินการทางกฎหมาย ที่ต้องดำเนินงานโดยสหวิชาชีพและกระทำตามพื้นฐานของพระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 มีใจความโดยสรุปว่า เมื่อพบหรือสงสัยว่าเด็กถูกกระทำทารุณ ต้องเร่งให้ความช่วยเหลือเพื่อให้เด็กได้อยู่ในที่ปลอดภัยรวมถึงต้องดำเนินการโดยผู้เชี่ยวชาญเฉพาะทางหรือที่เรียกว่าทีมสหวิชาชีพ

2. การบำบัดเยียวยาทางจิตใจ (Psychological treatment) เป็นการให้ความช่วยเหลือทางด้านจิตใจเพื่อลดผลกระทบด้านจิตใจที่จะเกิดกับเด็กให้มากที่สุด เพื่อให้เด็กหลุดพ้นจากสภาพความหวาดกลัว ความวิตกกังวลและความเครียดต่างๆ ที่ส่งผลต่อการพัฒนา การเรียนเรียนรู้และการใช้ชีวิตของเด็กในสังคม เป็นกระบวนการที่ต้องใช้ความชำนาญโดยผู้เชี่ยวชาญเนื่องจากช่วงอายุของเด็กอยู่ระหว่าง 0-18 ปี ทำให้มีความแตกต่างในช่วงวัย ดังนั้นการเยียวยาทางจิตใจจึงมีหลากหลายวิธี ต่างกันตามช่วงวัยของเด็กโดยผู้บำบัดจะต้องพิจารณาถึงความเหมาะสมต่อความต้องการตามวัยของเด็ก ศักยภาพของผู้ดูแล ความสัมพันธ์ระหว่างผู้กระทำความรุนแรงกับตัวของเด็ก และคนอื่นๆ ที่เกี่ยวข้องกับเด็กด้วย

2.1.5 ปัจจัยเสี่ยงที่ก่อให้เกิดการทารุณเด็ก

นอกจากตัวผู้กระทำแล้ว ยังมีปัจจัยเสี่ยงที่อาจเป็นตัวกระตุ้นให้เด็กถูกทารุณได้ ซึ่งไม่ใช่ปัจจัยจากตัวผู้กระทำเท่านั้น แต่ยังมีปัจจัยอื่นๆร่วมด้วย การให้ความช่วยเหลือหรือการดำเนินการต่างๆ จึงจำเป็นต้องคำนึงถึงปัจจัยเสี่ยงที่เป็นไปได้ทั้งหมดในทุกด้านเพื่อหาแนวทางลดอันตรายจากปัจจัยเสี่ยงลงและนำไปสู่การแก้ปัญหาได้ถูกจุด สามารถแบ่งได้ 4 กลุ่มดังนี้

1. ปัจจัยจากตัวเด็ก เช่น เป็นเด็กสมาธิสั้น (ADHD) ซึ่งเด็กจะมีอาการซนมากกว่าเด็กปกติ ไม่สามารถอยู่นิ่งได้ ทำให้ผู้เลี้ยงดูที่ขาดทักษะหรือความเข้าใจในตัวเด็กเกิดความรำคาญโมโห และอาจพลั้งมือทำร้ายเด็กได้

2. ปัจจัยจากพ่อแม่และผู้เลี้ยงดู เช่น พ่อแม่หรือผู้เลี้ยงดูที่มีปัญหายาเสพติด ส่งผลให้เกิดความผิดปกติทางอารมณ์ หงุดหงิด โมโหง่าย ก้าวร้าวและอารมณ์ร้ายกว่าคนปกติ หรือผู้ที่มีปัญหาทางสุขภาพจิตที่ไม่สามารถควบคุมตัวเองได้

3. ปัจจัยจากความสัมพันธ์ เช่นการขาดความผูกพันธ์ระหว่างเด็กกับผู้เลี้ยงดู มีปัญหาครอบครัว มีความรุนแรงในครอบครัว

4. ปัจจัยด้านสังคม เช่น ค่านิยมในสังคมที่ก่อให้เกิดความไม่เท่าเทียมกัน สังคมที่เต็มไปด้วยการก่ออาชญากรรมหรือเป็นแหล่งของสิ่งเสพติด อาจทำให้เด็กถูกใช้ให้ทำสิ่งที่ผิดต่อศีลธรรม การค้าประเวณี ต่างๆ ที่ส่งผลให้เด็กอาจถูกล่วงละเมิดทางเพศได้ เป็นต้น

2.2 พระราชบัญญัติ คุ้มครองเด็ก พ.ศ. 2546

พระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 เป็นกฎหมายไทย ประเภทพระราชบัญญัติ มีวัตถุประสงค์เพื่อ ปกป้อง คุ้มครองสวัสดิภาพเด็ก เพื่อประโยชน์สูงสุดของเด็ก โดยมีเนื้อหาเกี่ยวกับคณะกรรมการคุ้มครองเด็ก ข้อกำหนดในเรื่องการปฏิบัติต่อเด็ก การอบรมเลี้ยงดูที่ควรทำ บทบาทหน้าที่ของผู้ปกครองเด็ก รวมถึงผู้ที่เกี่ยวข้องกับเด็ก ผู้นำในชุมชน และเจ้าหน้าที่ต่างๆ บทกำหนดโทษ โดยได้ให้นิยามของคำว่า เด็ก หมายถึงบุคคลผู้ซึ่งมีอายุต่ำกว่าสิบแปดปีบริบูรณ์ แต่ไม่รวมถึงผู้ที่บรรลุนิติภาวะด้วยการสมรส กำหนดแนวทางการให้ความช่วยเหลือเด็กแบ่งตามลักษณะของเด็กที่พึงได้รับการสงเคราะห์กับเด็กที่พึงได้รับความคุ้มครองโดยได้กำหนดรายละเอียดไว้ในหมวดที่ 3 กับหมวดที่ 4 ดังนี้

1. เด็กเร่ร่อนหรือกำพร้า
2. เด็กที่ถูกทอดทิ้งหรือพลัดหลง
3. เด็กที่ผู้ปกครองไม่สามารถเลี้ยงดูได้ด้วยเหตุผลใดๆ
4. เด็กที่ผู้ปกครองประกอบอาชีพไม่เหมาะสมและอาจส่งผลต่อการพัฒนาทางร่างกายหรือจิตใจของเด็ก
5. เด็กที่ได้รับการเลี้ยงดูโดยมิชอบ
6. เด็กพิการ
7. เด็กที่อยู่ในสภาพยากลำบาก
8. เด็กที่จำต้องได้รับการสงเคราะห์ตามที่กำหนดในกระทรวง

<u>หมวด 4</u> มีเนื้อหาเกี่ยวกับลักษณะของเด็กที่พึงได้รับการคุ้มครองสวัสดิภาพ อำนาจของเจ้าหน้าที่ในการดำเนินการให้ความช่วยเหลือและการดำเนินการทางกฎหมาย อำนาจในการเข้าตรวจค้นและแยกตัวเด็กออกจากครอบครัวเพื่อคุ้มครองสวัสดิภาพโดยเร็วที่สุด อำนาจของศาลในการสั่งการไม่ให้ผู้กระทำทารุณเด็กเข้าใกล้ตัวเด็กหรือเข้าไปในสถานที่ที่เด็กอยู่ และข้อห้ามสำหรับผู้ปกครองหรือผู้ดูแลเด็กที่ห้ามปฏิบัติต่อเด็ก เด็กที่พึงได้รับการคุ้มครองมีดังนี้

1. เด็กที่ถูกทารุณกรรม
2. เด็กที่เสี่ยงต่อการกระทำผิด
3. เด็กที่อยู่ในสภาพที่จำต้องได้รับการคุ้มครองสวัสดิภาพตามที่กำหนดไว้ในกฎกระทรวง

ตามหลักการของพระราชบัญญัติคุ้มครองเด็ก พ.ศ. 2546 มุ่งเน้นการคุ้มครอง การทำความเข้าใจ มากกว่าการลงโทษตามกระบวนกฎหมาย ดังนั้น เจ้าหน้าที่และผู้เกี่ยวข้องจะเน้นการช่วยเหลือเด็กและประสานความเข้าใจระหว่างครอบกับตัวเด็กเป็นอันดับแรก โดยคำนึงถึงผลประโยชน์ของเด็ก ความปลอดภัย และมาตรฐานการเลี้ยงดูขั้นต่ำที่เด็กพึงได้รับ เพื่อให้เด็กสามารถใช้ชีวิตในสังคมได้อย่างปกติและอยู่ร่วมกับครอบครัวและผู้อื่นได้

2.3 เหมืองข้อมูล (Data Mining)

เหมืองข้อมูลคือการวิเคราะห์ข้อมูลเพื่อหาความสัมพันธ์จากฐานข้อมูลขนาดใหญ่ เป็นข้อมูลที่มีโครงสร้าง (Structured data) ในการค้นหาความสัมพันธ์ หาแนวโน้มใหม่ๆ ที่มีอยู่ในข้อมูลทำให้ความรู้ที่มีประโยชน์ถูกสกัดออกมาและสามารถนำไปใช้ประโยชน์ได้ โดยสิ่งที่ถูกสกัดออกมาเป็นสิ่งที่เราไม่เคยทราบมาก่อน (Unknown data) ซึ่งเป็นข้อมูลที่มีขนาดใหญ่และมาจากหลากหลายแหล่งจึงจำเป็นต้องมีการเตรียมข้อมูลก่อนเนื่องจากข้อมูลที่มีอาจจะล้าสมัย สูญหาย รูปแบบยังไม่เหมาะสม หรือเป็นข้อมูลมีความผิดปกติ มีขั้นตอนการทำเหมืองข้อมูลดังนี้

Data integration คือการรวมข้อมูลที่ได้มาจากแหล่งต่างๆ ให้อยู่ในรูปแบบเดียวกัน เนื่องจากข้อมูลที่รวบรวมอาจมาจากแหล่งต่างๆ ในแต่ละแหล่งที่มาของข้อมูลอาจมีรูปแบบที่ต่างกัน จึงต้องจัดการกับรูปแบบที่แตกต่างกันให้มีรูปแบบเดียวกันทั้งหมด

Data cleaning คือการกลั่นกรองข้อมูล เป็นขั้นตอนที่จัดการกับข้อมูลที่มีค่าผิดปกติหรือข้อมูลสูญหาย เพื่อให้ข้อมูลมีความถูกต้องและป้องกันความผิดพลาดในการนำข้อมูลไปใช้ทำเหมืองข้อมูล เนื่องจากข้อมูลที่เราเลือกมาอาจมีค่าที่ไม่ถูกต้องหรือมีข้อมูลที่สูญหายไป

ตารางที่ 2.1 ตัวอย่างข้อมูลลูกค้า

รหัสลูกค้า	รายได้	เพศ	อายุ
0001	45 000	M	41
0002	-20 000	F	24
0003	40 000	F	0
0004		M	33

จากตารางที่ 2.1 เป็นข้อมูลของลูกค้า แสดง รหัสลูกค้า รายได้ เพศ และอายุ จากตารางจะเห็นว่าในช่อง รายได้ มีข้อมูลที่มีค่าผิดปกติกับข้อมูลที่สูญหาย ส่วนในช่อง อายุ มีข้อมูลที่มีค่าผิดปกติอยู่ ข้อมูลที่ไม่ถูกต้องเหล่านี้ต้องถูกจัดการเสียก่อน

Data mining คือขั้นตอนการหารูปแบบที่เหมาะสมในการประมวลผลข้อมูลโดยใช้อัลกอริทึมที่เหมาะสมกับลักษณะของข้อมูลมาสร้างเป็นตัวแบบ (Model)

Pattern evaluation การประเมินคุณภาพของตัวแบบที่ใช้ในการทำเหมืองข้อมูล โดยจะประเมินตัวแบบว่ามีความแม่นยำ (Accuracy) ในการใช้ทำนายผลมากน้อยแค่ไหน และเลือกตัวแบบที่ดีที่สุดไปใช้

Knowledge representation เป็นขั้นตอนการนำเสนอความรู้ที่ค้นพบ

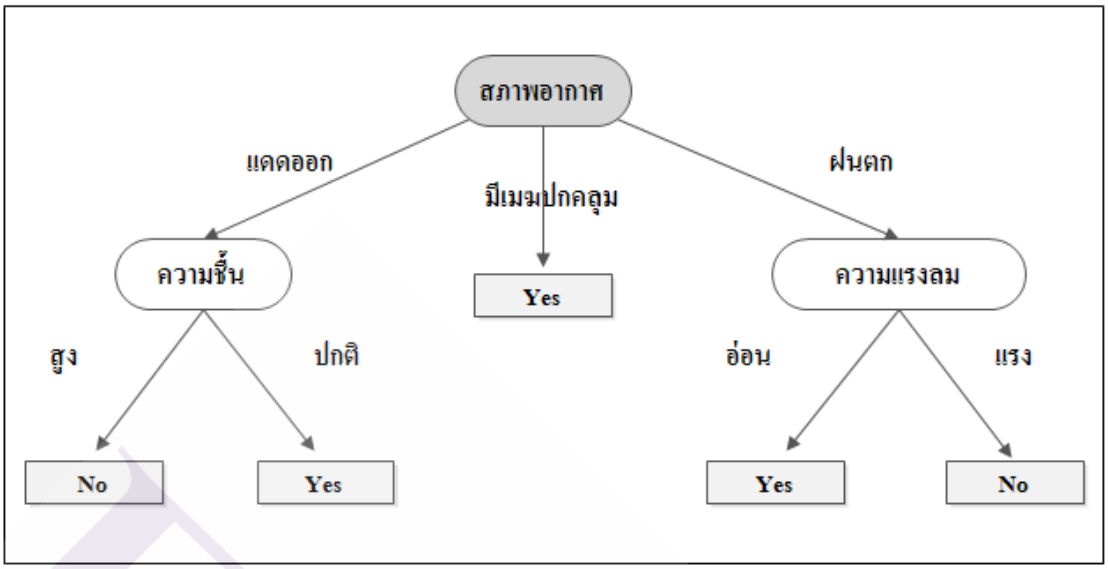
เทคนิคเหมืองข้อมูลสามารถนำไปประยุกต์ใช้กับงานได้หลากหลายด้าน เช่นด้านธุรกิจการเงิน การแพทย์ และวิทยาศาสตร์ เทคนิคในการทำเหมืองข้อมูลได้แก่ การหากฎความสัมพันธ์ (Association Rule Mining) การจำแนกประเภท (Classification) และการจัดกลุ่ม (Clustering)

2.4 เทคนิคการจำแนกประเภท (Classification)

เป็นกระบวนการสร้างตัวแบบ (Model) เพื่อจัดการข้อมูลให้อยู่ในกลุ่ม ที่เรียกว่า class ที่กำหนดให้ เพื่อใช้ในการทำนายข้อมูลต่อไป เช่นการจัดกลุ่มของพนักงานในบริษัทว่า ดีมาก ดี ปานกลาง ไม่ดี โดยพิจารณาจากข้อมูลที่มีอยู่ โดยมีอัลกอริทึมที่ใช้ในการจำแนกประเภท (Classifier algorithm) มีดังนี้

2.4.1 แผนภาพต้นไม้เพื่อการตัดสินใจ (Decision tree) นิยมใช้กันมาก สามารถทำความเข้าใจได้ง่าย มีโครงสร้างเป็นรูปต้นไม้ที่มีโหนด (node) แต่ละโหนดแสดงถึงคุณลักษณะ (Attribute) ของข้อมูลโดยเริ่มจากโหนดของราก (root node) ซึ่งเป็นจุดเริ่มต้นของเหตุการณ์ เชื่อมต่อไปยังโหนดอื่นๆ ด้วยกิ่ง (branches) ซึ่งแสดงผลของการทดสอบ และโหนดของใบ (leaf node) ที่แสดงคำตอบหรือผลลัพธ์ที่ได้ ซึ่งเป็นกลุ่ม (class) ที่กำหนดไว้ แผนภาพต้นไม้เพื่อการตัดสินใจเป็นการเรียนรู้แบบมีผู้สอน (Supervised Learning) โดยในขั้นตอนการฝึก (Training) ต้องเตรียมชุดข้อมูลและคำตอบ (Class) ที่ถูกต้องสำหรับให้เครื่องเรียนรู้ ตัวอย่างแผนภาพต้นไม้เพื่อการตัดสินใจ (สายชล สินสมบูรณ์ทอง 2558) แสดงดังภาพที่ 2.1 ที่แสดงรูปแบบของการตัดสินใจ (if/then/else) ว่าจะเล่นกีฬาหรือไม่ในสภาพอากาศต่างๆ เช่น

if สภาพอากาศ = แดดออก และความชื้น = ปกติ / then เล่นกีฬา = ใช่



ภาพที่ 2.1 ตัวอย่างแผนภาพเพื่อการตัดสินใจ

การเลือก Attribute มาสร้างต้นไม้ในแต่ละโหนดจะใช้ค่า Information Gain (IG) ในการเลือก โดย Attribute ที่มีค่า IG สูงสุด จะถูกเลือกมาเป็นโหนดในแต่ละขั้นตอนในการสร้างต้นไม้ สูตรคำนวณหา Information Gain สามารถแสดงได้ดังนี้ (Margaret H. Dunham 2002)

$$\text{Gain}(A\ \ D) = H(D) - \sum_{i=1}^{S} P(D_i)H(D_i) \quad (2.1)$$

A = ค่าทั้งหมดของชุดตัวอย่าง ที่ประกอบด้วยตัวแปรต้นแต่ละตัวในหลายๆ กรณี

D = ค่าของตัวแปรที่ต้องการนำมาพิจารณา

H(D) = ค่าเอนโทรปีของเป้าหมาย A

$$\text{Entropy}(p_1\ \ p_2\ \ \ldots\ \ p_n) = -\,p_1\,log_2\,p_1\ -\ p_2\,log_2\,p_2\,\ldots\,-\ p_n\,log_2\,p_n \quad (2.2)$$

P = ค่าที่เป็นไปได้ / ค่าทั้งหมด

S = เซตย่อยที่เป็นไปได้ทั้งหมดของแอททริบิวต์ A D

$P(D_i)$ = ค่าความน่าจะเป็นของเซตย่อย i ของแอททริบิวต์ D

$H(D_i)$ = ค่าเอนโทรปีของแอททริบิวต์ D

2.4.2 Naïve Bayes

Naïve Bayes เป็นเทคนิคที่ใช้แก้ปัญหาการจำแนกประเภทที่ใช้ทฤษฎีทางด้านสถิติคำนวณความน่าจะเป็นในการทำนายผล สามารถคาดการณ์ผลลัพธ์และอธิบายได้ โดยจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้สร้างเงื่อนไขความน่าจะเป็นในแต่ละความสัมพันธ์ การทำงานไม่ซับซ้อนเหมาะกับชุดข้อมูลจำนวนมากและคุณสมบัติของตัวอย่างไม่ขึ้นต่อกันสมการของ Bayes แสดงตามสมการด้านล่าง

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} \tag{2.3}$$

จากสมการของ Bayes สามารถอธิบายได้ดังนี้

P(B|A) คือ Posterior probability หมายถึงค่าความน่าจะเป็นที่ข้อมูลที่มีแอทริบิว A จะมีคลาส B

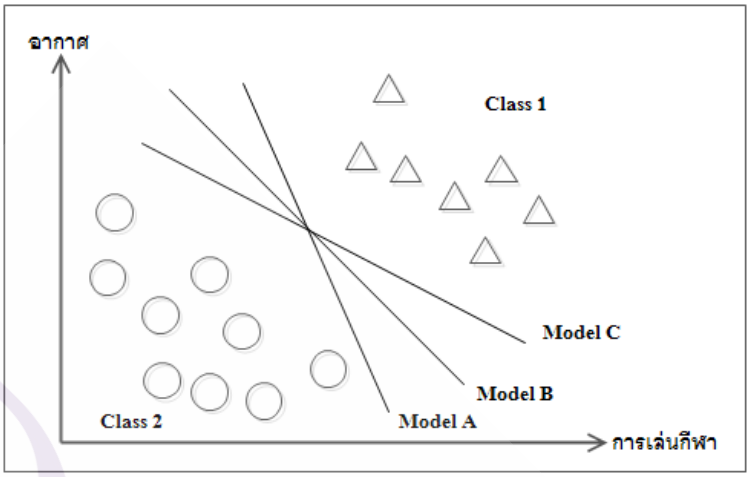
P(A|B) คือ Likelihood หมายถึงชุดข้อมูลที่ใช้ฝึก (Training data set) มีคลาส B และแอทริบิว A โดยที่ $A = A = a_1 \bigcap a_2 \ldots \bigcap a_M$ ซึ่ง M คือจำนวนแอทริบิวทั้งหมดใน Training data set และใช้สมมติฐานที่ว่าแอททริบิวต์ต่างๆ ไม่ขึ้นต่อกัน (คุณลักษณะของตัวอย่างไม่ขึ้นต่อกัน) สามารถเขียนเป็นสมการใหม่ได้ดังนี้ $P(A|B) = P(a_1 | B) \times P(a_2 | B) \times \ldots\ldots\ldots\ldots P(a_M | B)$

P(B) คือ Prior probability หมายถึงค่าความน่าจะเป็นในการเกิดคลาส B

2.4.3 Support Vector Machine (SVM)

SVM เป็นการเรียนรู้แบบใหม่ที่ใช้ทฤษฎีทางสถิติในการวิเคราะห์และจำแนกข้อมูล เป็นเทคนิคที่ใช้ในการแก้ปัญหาการรู้จำข้อมูลโดยจะช่วยลดการเกิดปัญหา Overfitting คือการมีโมเดลของชุดข้อมูลมากเกินไปจนไม่สามารถทำนายข้อมูลที่ไม่ทราบ (unseen data) ได้ SVM จะหาระนาบการตัดสินใจในการแบ่งข้อมูลเป็น 2 ส่วน โดยจะนำค่าของกลุ่มข้อมูลมาวางลงใน Feature space และพยายามสร้างเส้นกึ่งกลาง (Hyperplane) เพื่อแบ่งข้อมูลทั้ง 2 กลุ่ม โดยใช้สมการเส้นตรง (linear model) ในการแบ่งข้อมูลออกเป็น 2 คลาส ซึ่ง linear model สามารถแบ่งเส้นได้มากกว่า 1 เส้น และจะเลือกเส้นที่มีระยะห่างระหว่างกลุ่ม 2 กลุ่มเยอะที่สุด เส้นที่เลือกจะต้องไม่เข้าใกล้ข้อมูลกลุ่มหนึ่งกลุ่มใดมากเกินไปเพื่อป้องกันการเกิด overfitting ของข้อมูล ภาพที่ 2.3 ได้แสดง ตัวอย่าง Linear model ในการแบ่งข้อมูลออกเป็น 2 กลุ่ม ตามหลักการของ SVM การเลือกเส้นระนาบเพื่อนำมาใช้ทำนายผลความสัมพันธ์ของอากาศกับการเล่นกีฬาโดยแบ่งเป็น 2 คลาส คือ Class 1 กับ Class 2 ซึ่งมีการกระจายตัวของข้อมูลออกเป็นกลุ่มอย่างชัดเจน และสามารถแบ่ง Linear model ได้มากกว่า 1 เส้น คือ Model A Model B Model C จากรูป Model B เป็นเส้นที่ถูกเลือก

เนื่องจาก Model A กับ Model C มีจุดที่อยู่ใกล้ข้อมูลมากเกินไป จะทำให้การทำนายผิดพลาดได้ แต่ Model B มีระยะห่างระหว่าง 2 กลุ่มมากที่สุด



ภาพที่ 2.2 ตัวอย่าง Linear model ของ SVM ในการแบ่งข้อมูลออกเป็น 2 กลุ่ม

2.5 เหมืองข้อความ (Text Mining)

เหมืองข้อความคือการค้นหาข้อมูลสารสนเทศที่มีประโยชน์จากภาษาธรรมชาติ (Natural language เป็นสกัดความรู้ เอาข้อมูลที่ไม่เคยได้รู้มาก่อน และสามารถนำไปใช้ประโยชน์ได้ตามวัตถุประสงค์จำเพาะเจาะจง ข้อมูลที่มีอยู่ในโลกส่วนใหญ่เป็นข้อมูลแบบไม่มีโครงสร้าง (Unstructured) หรือเป็นข้อมูลกึ่งโครงสร้าง(Semi-structured) ซึ่งเป็นข้อมูลที่มีหลากหลายมิติ มีความซับซ้อนและมีความไม่ชัดเจน การทำเหมืองข้อความเป็นการใช้เทคนิคทั้งในส่วนของคอมพิวเตอร์ คือเหมืองข้อมูล (Data mining) และส่วนของการวิเคราะห์ทางภาษาของมนุษย์ คือการประมวลผลภาษาทางธรรมชาติ (Natural language processing) รวมถึงการใช้หลักการของการค้นคืนสารสนเทศ (Information retrieval) การเรียนรู้ของเครื่อง (Machine learning) และสถิติ (Statistics) เข้าด้วยกัน

2.5.1 การประมวลผลภาษาธรรมชาติ (Natural language processing - NLP)

การประมวลผลภาษาธรรมชาติหรือ NLP เป็นกระบวนการที่ทำให้คอมพิวเตอร์เข้าใจภาษาธรรมชาติของมนุษย์ คำว่าภาษาธรรมชาติ (Natural language) หมายถึงภาษาที่มนุษย์ใช้สื่อสารทั่วไปในชีวิตประจำวัน ต่างจากภาษาคอมพิวเตอร์ที่ถูกประดิษฐ์ขึ้น คอมพิวเตอร์จำเป็นต้องมีการวิเคราะห์เพื่อจะได้เข้าใจความหมาย โดยมีองค์ประกอบในการวิเคราะห์หลักๆ ดังนี้

1. การวิเคราะห์เชิงโครงสร้าง (Syntactic Analysis) เป็นการตรวจสอบทางไวยากรณ์ ประเภท หน้าที่ของคำ กลุ่มคำ และประโยค โดยจะทำการตรวจสอบประโยคว่าถูกต้องตามหลักไวยากรณ์หรือไม่ เช่น I at work cried ประโยคนี้มีโครงสร้างที่ไม่ถูกต้องตามหลักไวยากรณ์

2. การวิเคราะห์เชิงความหมาย (Semantic Analysis) เป็นการตรวจสอบความถูกต้องในเชิงความหมายของประโยคโดยดูจากหน้าที่ของคำในประโยคและโครงสร้างทางไวยากรณ์ ซึ่งคำบางคำมีความหมายไม่แน่นอน ขึ้นอยู่กับบริบทของแต่ละประโยค เช่น I get the bill กับ Do you get it? คำว่า get ในทั้งสองประโยคนี้มีความหมายต่างกัน โดยประโยคแรกหมายถึง “ได้รับ” และประโยคที่สองหมายถึง “เข้าใจ” (understand) หรือประโยค The moon is walking ประโยคนี้ถูกต้องตามหลักไวยากรณ์ แต่ถ้าดูเชิงความหมายจะพบว่าเป็นไปไม่ได้

3. การวิเคราะห์เชิงตีความ (Pragmatic Analysis) เป็นการวิเคราะห์ตีความหมายของคำถามที่เราถามและตีความตามสถานการณ์ที่เกิดขึ้น เช่นเมื่อเราถามว่าวันนี้อากาศเป็นอย่างไร ผู้ที่ตอบคำถามเราจะต้องเข้าใจว่าเราต้องการรู้อะไร และตอบตามสถานการณ์นั้นๆ เช่น วันนี้อากาศแจ่มใส

2.5.2 กระบวนการอย่างง่ายในการประมวลผลภาษาธรรมชาติ มีขั้นตอนดังนี้

ในการประมวลผลภาษาไทยจะเริ่มจากการตัดคำ (word segmentation) ซึ่งเป็นการแบ่งข้อความออกเป็นคำ ถัดไปเป็นขั้นตอนการกำจัดคำที่ไม่มีความสำคัญออก (Stop word) ซึ่งเป็นการกำจัดคำที่ปรากฏบ่อยในเอกสารและไม่มีความสำคัญออกเพื่อลดขนาดของข้อมูลและป้องกันความผิดพลาดในการวิเคราะห์ข้อมูล และขั้นตอนการหาคำพ้องความ หรือการหารูปเดิมของคำ

ซึ่งเป็นการจัดกลุ่มคำที่มีความหมายเดียวกันเพื่อลดความซับซ้อนของคำ และการวิเคราะห์หน้าที่ของคำ (Part of speech : POS) รายละเอียดของขั้นตอนต่างๆ มีดังนี้

1. การตัดคำ (Word segmentation) คือการแยกข้อความในเอกสารออกเป็นคำ โดยใช้วิธีการต่างๆ เช่นตัดคำตามพจนานุกรม (Dictionary base)

ตารางที่ 2.2 ตัวอย่างการตัดคำ

ประโยค	แม่ออกไปซื้อของที่ตลาดและจะกลับมาตอน 6 โมงค่ะ
ผลการตัดคำ	แม่ | ออก | ไป | ซื้อ | ของ | ที่ | ตลาด | และ | จะ | กลับ |มา | ตอน | 6 | โมง | ค่ะ

2. การกำจัดคำที่ไม่มีความสำคัญในประโยคหรือคำหยุด (Stop word) เมื่อแยกข้อความออกเป็นคำแล้วจะทำการตัดคำที่ไม่มีความสำคัญในประโยคออกไปเพื่อลดความผิดพลาดที่จะ

เกิดขึ้นในการวิเคราะห์ข้อมูล และลดขนาดของจำนวนคำให้เหลือเพียงคำที่จำเป็นเท่านั้น ประเภทของคำหยุดในภาษาไทยมีดังนี้ คำบุพบท คำสันธาน คำสรรพนาม คำวิเศษณ์ และคำอุทาน

ตารางที่ 2.3 ตัวอย่างการกำจัดคำหยุด

ผลการตัดคำ	แม่ | ออก | ไป | ซื้อ | ของ | ที่ | ตลาด | และ | จะ | กลับ |มา | ตอน | 6 | โมง | ค่ะ
การกำจัด Stop word	แม่ | ออก | ไป | ซื้อ | ของ | ที่ | ตลาด | ~~และ~~ | จะ | กลับ |มา | ตอน | 6 | โมง | ~~ค่ะ~~
ผลการกำจัด Stop word	แม่ | ออก | ไป | ซื้อ | ของ | ที่ | ตลาด |จะ | กลับ |มา | ตอน | 6 | โมง

จากตัวอย่าง คำว่า “และ” กับ “ค่ะ” จะถูกตัดออกจากประโยค

3. การหาคำพ้องความ ภาษาไทยเป็นภาษาที่มีคำศัพท์ที่หลากหลายในความหมายเดียวกัน ซึ่งมีทั้งคำที่เป็นคำไทยแท้ และคำที่ยืมมาจากภาษาต่างประเทศเช่น ภาษาเขมร ภาษาบาลี เป็นต้น การหาคำพ้องความในภาษาไทยได้ใช้วิธีสร้างคลังคำพ้องความ หรือดิกชันนารี เพื่อเก็บคำที่มีความหมายเดียวกันไว้

ตารางที่ 2.4 ตัวอย่างคำพ้องความในภาษาไทย

คำศัพท์	คำพ้องความ
พ่อ	พ่อ
บิดา	พ่อ
ป๊า	พ่อ
ป๋า	พ่อ
แม่	แม่
มารดา	แม่
ทำโทษ	ทำโทษ
ลงโทษ	ทำโทษ

4. การวิเคราะห์ชนิดของคำ (Parts of speech)

เป็นการวิเคราะห์คุณลักษณะทางไวยากรณ์ โดยติดแท็กให้กับคำที่ถูกแยกออกแต่ละคำ เพื่อแสดงหน้าที่ของคำในประโยค หรือเป็นคลาสของคำ (word class) โดยบ่งบอกถึงคุณลักษณะและหน้าที่ของคำ

เกณฑ์การจำแนกชนิดของคำมีความแตกต่างกันตามลักษณะโครงสร้างทางไวยากรณ์ในแต่ละภาษา (อมรา ประสิทธิ์รัฐสินธุ์ 2543) เช่นคำว่า สวย ในภาษาไทยเป็นคำกริยา (Verb) ในภาษาเกาหลีเป็นคำกริยาพรรณนา (Descriptive verb) และในภาษาอังกฤษเป็นคำวิเศษณ์ (Adjective) เป็นต้น จำนวนชนิดของคำไม่ได้มีการกำหนดตายตัว สำหรับภาษาไทยการนำเสนอตามงานวิจัยได้มีการจำแนกชนิดของคำไว้หลายแบบ สามารถสรุปได้เป็น 7 ชนิด ดังนี้ คำนาม (Nouns) สรรพนาม (Pronouns) คำกริยา (Verbs) คำวิเศษณ์ (Adverbs) คำบุพบท (Preposition) คำสันธาน (Conjunctions) คำอุทาน (Interjections) (พระยาอุปกิตศิลปสาร 2491)

ตารางที่ 2.5 ตัวอย่างการวิเคราะห์ชนิดของคำ (Part of speech)

คำศัพท์	ชนิดของคำ
แม่	คำนาม
ไป	คำกริยา
ซื้อ	คำกริยา
ของ	คำนาม
ที่	คำบุพบท
ตลาด	คำนาม

2.6 งานวิจัยที่เกี่ยวข้อง

การใช้เทคนิคเหมืองข้อความและการประมวลภาษาธรรมชาติในกระบวนพิจารณาการขอเคลมประกันสุขภาพ (Fred Popowich 2005 : 59-66) เป็นการนำเทคนิคเหมืองข้อความและเทคนิคการประมวลผลภาษาธรรมชาติ (NLP) มาช่วยตัดสินใจในการจ่ายค่าสินไหมประกันสุขภาพ เมื่อมีผู้ทำเรื่องขอเคลมประกัน ขั้นตอนการทำงานของระบบเริ่มจากการบันทึกข้อความ(Medical claim) บันทึกข้อมูลจาก Call center บันทึกข้อมูลผู้ป่วยและข้อมูลที่เกี่ยวข้องอื่นๆ เข้าสู่ระบบ จากนั้นทำการประมวลผลโดยใช้การวิเคราะห์คอนเซป (Concept Analyzer) ในกระบวนการประมวลผลภาษาทางธรรมชาติ และทำการประเมินผลการวิเคราะห์ด้วยเทคนิคเหมืองข้อความ

ในขั้นตอนของการวิเคราะห์คอนเซปผู้วิจัยได้ประยุกต์ใช้หลักการของ Concept taxonomies ที่ใช้ในการจัดกลุ่มคำที่มีความหมายเหมือนกัน โดยนำมาประยุกต์ใช้ในการจัดกลุ่มของลักษณะการเกิดอุบัติเหตุ และการรักษาพยาบาลที่เป็นอาการบาดเจ็บในลักษณะเดียวกัน โดยวิเคราะห์จากบันทึกข้อมูลการเกิดอุบัติเหตุและสถานที่เกิดเหตุ ข้อมูลการตรวจรักษาพยาบาล และข้อมูลอื่นๆ ที่เกี่ยวข้อง แล้วทำการวิเคราะห์ว่าเป็นอุบัติเหตุแบบใด ผลการวิเคราะห์คอนเซปที่ได้จะถูกนำไปใช้เป็นตัวชี้วัด (Indicator) เพื่อนำไปประเมินผลการจ่ายเงินชดเชย ซึ่งเป็นส่วนของการประเมินตัวชี้วัด (evaluating indicators) โดยให้ผู้เชี่ยวชาญกำหนดคะแนนของบันทึกข้อมูลการขอเคลมประกันและแสดงผลคะแนนให้กับผู้ใช้งานซึ่งมีค่าอยู่ระหว่าง 1-1000 เพื่อทำการตัดสินใจว่าควรจะจ่ายเงินทดแทนหรือไม่

ระบบที่ทำประสบความสำเร็จเป็นอย่างมาก เนื่องเป็นข้อมูลเกี่ยวกับการแพทย์ โรคติดต่อ และการรักษา ซึ่งเป็นข้อมูลเฉพาะทางในขอบเขตของงาน ทำให้สามารถสร้างแบบจำลองที่สมบูรณ์ได้ นอกจากนี้ยังสามารถนำไปประยุกต์ใช้กับระบบต่างๆ ที่มีการวิเคราะห์ทางภาษาศาสตร์และงานที่เกี่ยวข้องกับความรู้เชิงความหมายได้

การประเมินโรคเครียดโดยอัตโนมัติจากเหตุการณ์สะเทือนขวัญจากคำบอกเล่าของผู้ป่วยด้วยการประมวลภาษาทางธรรมชาติและเหมืองข้อความ (Qiwei He et al. 2015) งานวิจัยนี้นำเสนอวิธีการประมวลภาษาทางธรรมชาติกับการใช้เทคนิคเหมืองข้อความสำหรับการประเมินอาการของโรคเครียดจากเหตุการณ์สะเทือนขวัญ (Posttraumatic Stress Disorder : PTSD) จากข้อความที่บันทึกด้วยของตัวผู้ป่วยเอง (Self-narratives) ซึ่งเป็นผู้ที่ประสบกับเหตุการณ์สะเทือนใจ (Traumatic experience) ทำให้เกิดความเครียดความหวาดกลัว และความวิตกกังวลที่ส่งผลต่อการใช้ชีวิตประจำวัน

ผู้วิจัยได้เก็บตัวอย่างจากผู้ที่มีประสบการณ์สะเทือนใจ 300 ตัวอย่าง จากการให้ผู้ป่วยเข้าไปบันทึกข้อมูลประสบการณ์ร้ายแรงที่เกิดขึ้นกับตนเองและส่งผลต่อชีวิตประจำวันผ่านทางเว็บไซต์ แบ่งเป็นกลุ่มที่มีแนวโน้มของอาการ PTSD 150 ตัวอย่าง และไม่มีอาการ PTSD (Non-PTSD) 150 ตัวอย่าง เพื่อใช้ในการเรียนรู้ เปรียบเทียบความเหมือนและความต่างโดยใช้ผู้เชี่ยวชาญจำนวน 2 คนทำการประเมิน และหากผลการประเมินต่างกันจะใช้ผู้เชี่ยวชาญเพิ่มอีกหนึ่งคนเป็นคนตัดสิน ในขั้นตอนของการเรียนรู้ ผู้วิจัยได้นำโมเดล n-grams มาประยุกต์ใช้ในส่วนของการเตรียมข้อมูล เพื่อ ลดปัญหาในส่วนของคำหลายพยางค์ รวมถึงการทำให้ความหมายของบางคำมีความชัดเจนขึ้น โดยคำที่ได้จากการนำโมเดล n-grams เข้ามาใช้จะถูกนำไปสร้างเป็นถุงคำ (Bag of words : BOW) ที่เก็บรวบรวมคำกับคุณลักษณะของคำไว้ ในส่วนของการทดสอบ

ในงานวิจัยนี้ได้ใช้เทคนิค 4 เทคนิค ในการสร้างตัวแบบได้แก่ Decision tree Naïve Bayes Support Vector Machine และ Product Score Model และประเมินความแม่นยำ (Accuracy) ของตัวแบบด้วย 10-fold cross-validation จากผลการทดสอบพบว่า Product Score Model ให้ความแม่นยำสูงสุด โดยมีค่าเฉลี่ยอยู่ที่ .82 (SD =.05)

การระบุอาการได้รับบาดเจ็บในถุงลมปอดของผู้ป่วยจากบันทึกรายงานผลการตรวจเอ็กซเรย์ทรวงอก (Meliha Yetisgen-Yildiz et al. 2013 : 10-17) เป็นการใช้เทคนิคการประมวลผลภาษาทางธรรมชาติในการวิเคราะห์ข้อมูลการตรวจเอ็กซเรย์ทรวงอกของผู้ป่วยว่ามีผลเป็นบวก (Positive for Acute Lung Injury) หรือไม่จากข้อมูลที่ตรวจจากการเอ็กซเรย์ ผู้วิจัยมีวัตถุประสงค์ในการพัฒนาวิธีการอัตโนมัติในการตรวจสอบลักษณะอาการบาดเจ็บของถุงลมซึ่งเป็นการบาดเจ็บที่ร้ายแรงโดยใช้เทคนิคการประมวลภาษาทางธรรมชาติ (NLP) กับเทคนิคการจำแนกประเภทด้วยการเรียนรู้แบบมีผู้สอน (Supervised text classification) ใช้ unigram bigram trigram ในการเลือกคุณลักษณะ (feature selection) โดยทดสอบการเลือกคุณลักษณะในการตัดสินใจด้วยการทดสอบสมมติฐานทางสถิติ (statistical hypothesis testing) และใช้อัลกอริทึม Maximum Entropy (MaxEnt)ในการจำแนกประเภท

ผู้วิจัยได้ทำการทดสอบโดยแบ่งตามค่าสูงสุดของค่าพื้นฐาน (threshold value : *N*) ที่ต่างกัน และแบ่งลักษณะ feature เป็น 3 แบบ คือ unigram unigram + bigram และ unigram + bigram + trigram ผลการทดสอบพบว่าค่า f-score ที่ดีที่สุดคือ 78.53% ซึ่งเป็นการทดสอบด้วยค่า *N* = 800 และใช้ ลักษณะของ feature แบบ unigram + bigram + trigram โดยมีค่าความแม่นยำ (precision) = 81.70 และค่าระลึกถึง (recall) = 75.59

การทำเหมืองข้อความสำหรับระบบรายงานอาการหลังได้รับการเสริมสร้างภูมิคุ้มกันด้วยเทคนิคการจำแนกประเภทข้อความทางการแพทย์โดยใช้การเลือกคุณลักษณะ (Taxiachris Botsis Michael D Nguyen Emily Jane Woo Marianthi Markatou Robert Ball 2011 : 631-638) ระบบรายงานอาการหลังได้รับการเสริมสร้างภูมิคุ้มกัน (Vaccine Adverse Event Reporting System : VAERS) เป็นระบบที่เก็บรวบรวมรายงานต่างๆ เกี่ยวกับอาการข้างเคียงจากการฉีดวัคซีนและมีเจ้าหน้าที่ทำการตรวจสอบรายงานที่บันทึกเพื่อดูผลของการฉีดวัคซีน ในงานวิจัยนี้ผู้วิจัยต้องการสร้างระบบอัตโนมัติเพื่อจำแนกอาการข้างเคียงที่เกิดจากการฉีดวัคซีนจากรายงานที่บันทึกในระบบ VAERS ผู้วิจัยเลือกศึกษาข้อมูลรายงานของการฉีดวัคซีนป้องกันไข้หวัดใหญ่ 2009 (H1N1) จำนวน 6034 รายงานและทำการจำแนกการเกิดอาการแพ้เฉียบพลัน (anaphylaxis) เริ่มจากการนำบันทึกรายงานเข้าสู่กระบวนการประมวลผลภาษาธรรมชาติ (Natural language processing) และนำคำที่ทำการตัดคำ (Tokenization) และถูกเปลี่ยนให้อยู่ในรากศัพท์ (Stemming) แล้วมาสกัด

คุณลักษณะโดยทำการเลือกคำที่สำคัญ (Important keywords) และนำ keywords ที่ได้มาวิเคราะห์หาคำที่แสดงถึงอาการแพ้ทั้งอาการหลักและอาการรอง (major / minor criteria) จากนั้นจึงนำมาคัดกรองรูปแบบ (pattern filter) และทำการจำแนกจากลักษณะของรูปแบบว่าใช่อาการแพ้ที่เกิดจากการได้รับวัคซีนหรือไม่

ผู้วิจัยได้ใช้เทคนิคการจำแนกประเภทแบบมีผู้สอนหลากหลายเทคนิค แต่เทคนิคที่ดีที่สุดคือ Weighted support vector machine กับ Boosted trees โดยผลที่ได้มีความเป็นไปได้มากที่จะนำเทคนิคของเหมืองข้อความเข้ามาใช้ในการช่วยจำแนกอาการหลังได้รับการฉีดวัคซีนและมีโอกาสที่จะพัฒนาระบบเพิ่มเติมในงานอื่นๆ ที่มีลักษณะเดียวกัน

การกำหนดศัพท์แพทย์ที่เกี่ยวกับโรคเฉพาะทาง (Mihir Shekhar Veera Ragahvendra Chikka Lini Thomas Sunil Mandhan Kamalakar Karlapalem 2015 : 171-177) งานวิจัยนี้มีวัตถุประสงค์ในการจำแนกประเภทของศัพท์แพทย์ที่บันทึกอยู่ในเวชระเบียนอิเล็กทรอนิกส์เพื่อระบุกลุ่มของโรค ผู้วิจัยได้ศึกษาและจำแนกโรคเฉพาะทางทั้งหมด 5 โรค ได้แก่ โรคมะเร็ง โรคเอดส์ โรคข้ออักเสบ โรคเบาหวาน และโรคที่เกี่ยวกับ ในส่วนของการเตรียมข้อมูล (Preprocessing) ใช้เทคนิคของการประมวลผลภาษาทางธรรมชาติโดยทำการสกัดคำ และเลือกเฉพาะคำศัพท์แพทย์ โดยผู้วิจัยได้สร้างดิกชันนารี (medical dictionary) และใช้ข้อมูลจาก Unified Medical Language System (UMLS) คำศัพท์แพทย์ที่ได้จะนำไปใช้ในขั้นตอนของ Post processing ซึ่งผู้วิจัยต้องการให้คำศัพท์แพทย์ที่เลือกออกมามีความถูกต้องและชัดเจนมากขึ้นจึงใช้ 4 -grams เข้ามาใช้ในการช่วยสกัดคำ โดยจะเลือก n-grams ที่ยาวที่สุด (longest n-grams จากนั้นจึงนำข้อมูลที่ได้มาทำการจำแนกประเภทของโรคว่าเป็นโรคประเภทใด

ผู้วิจัยใช้ Support Vector Machines (SVM) ในการจำแนกประเภทเนื่องจากเป็นโมเดลที่นิยมใช้ในการประมวลผลภาษาธรรมชาติและมีประสิทธิภาพ ให้ผลได้ดีที่สุด โดยจากค่าเฉลี่ยจากผลการจำแนกประเภทโรคทั้ง 5 ประเภท มีค่าเฉลี่ยของ F-score = 0.966

การจำแนกบทความทางวิทยาศาสตร์ที่เกี่ยวกับการแพร่เชื้อของโรคมะเร็ง (Simon Baker Llona Silins Yufan Guo Imran Ali Johan Hogberg Ulla Stenius Ana Korhonen 2015 : 1-9) การแพร่เชื้อของโรคมะเร็ง (Hallmarks of cancer) ที่กำหนดโดย Hanahan กับ Weinberg ได้ส่งผลต่อการวิจัยด้านมะเร็งเป็นอย่างมาก เนื่องจากได้ลดรูปแบบการแพร่เชื้อที่ซับซ้อนของมะเร็งลงโดยเหลือเพียง 10 รูปแบบ ผู้วิจัยเห็นว่าหากนำเทคโนโลยีการจำแนกความหมาย (Semantic Classification) มาใช้จะสามารถช่วยในการค้นพบความรู้ใหม่ๆ จากบทความทางวิทยาศาสตร์ที่ใช้ Hallmarks of cancer ของ Hanahan กับ Weinberg ได้

งานวิจัยนี้เลือกใช้เทคนิคเหมืองข้อความทำการจำแนกประเภทบทความโดยจำแนกจากคำที่แสดงถึงความหมายของโรคมะเร็งที่แบ่งตามการแพร่เชื้อของมะเร็ง 10 ขั้นตอน ในขั้นตอนการประมวลผลภาษาทางธรรมชาติ ผู้วิจัยใช้ pipeline เป็นเครื่องมือในการตัดคำ ติดแท็กชนิดคำ และการหารูปเดิมของคำ (Lemmatization) ในส่วนของการคัดเลือกคุณลักษณะได้ใช้ bigrams ในการช่วยเลือกคำหลังจากการหารูปเดิมของคำแล้ว ซึ่งจะทำให้ได้คำที่ชัดเจนและมีความหมายที่ถูกต้องขึ้น ในการจำแนกประเภทได้เลือกใช้เทคนิค Support vector machine (SVM)

จากผลการทดสอบพบว่าเทคนิคเหมืองข้อความสามารถนำไปใช้ในการจำแนกความหมายในบทความทางวิทยาศาสตร์ได้

จากงานวิจัยที่เกี่ยวข้องได้ใช้เทคนิคที่แตกต่างกันในการประมวลผลภาษาธรรมชาติ มีการวิเคราะห์คอนเซปโดยการใช้ Taxonomies ในการจัดกลุ่มคำ การนำโมเดล n-grams เข้ามาช่วยในส่วนของ pre-processing โดยนำคำที่ได้จากการใช้โมเดล n-grams มาสร้างเป็นถุงคำ (Bag of words) การคัดเลือกคุณลักษณะ (Feature) โดยผู้เชี่ยวชาญ เทคนิคการจำแนกประเภทที่ใช้ คือ Naïve Bayes Neural Network Decision Tree Support Vector Machine และ Product Score Model

ผู้วิจัยได้นำแนวคิดจากที่ได้จากงานวิจัยที่ศึกษาไปประยุกต์ใช้กับงานวิจัยนี้ โดยได้นำเทคนิคการประมวลผลภาษาธรรมชาติ (Natural language processing : NLP) การสร้างถุงคำพ้องความ (Bag of Synonyms) มาประยุกต์ใช้ในงานวิจัย เทคนิคที่นำมาจำแนกในงานวิจัยนี้ คือ Naïve Bayes Decision Tree และ Support Vector Machine

บทที่ 3
ระเบียบวิจัย

ในบทนี้จะกล่าวถึงขั้นตอนการดำเนินงานวิจัย และสถาปัตยกรรมการทำงานของระบบ เริ่มตั้งแต่การเตรียมข้อมูลเพื่อใช้ในการเรียนรู้ (Training) การสกัดคุณลักษณะ การจำแนกประเภท การให้ความช่วยเหลือเด็กที่ถูกทารุณ การสร้างตัวแบบ และการประเมินผล โดยมีรายละเอียดดังนี้

3.1 ขั้นตอนการดำเนินงานวิจัย

วิธีการดำเนินงานวิจัยเป็นการศึกษาเก็บรวบรวมข้อมูลสำหรับการสร้างตัวแบบสำหรับการทำเหมืองข้อความ มีขั้นตอนต่างๆ ดังนี้

3.1.1 การเก็บรวบรวมและการวิเคราะห์ข้อมูล

กระบวนการนี้เป็นการเก็บรวบรวมข้อมูลทั้งแบบปฐมภูมิและทุติยภูมิ เริ่มตั้งแต่การศึกษาความรู้ที่เกี่ยวข้องกับการให้ความช่วยเหลือเด็กที่ถูกทารุณ การเก็บรวบรวมเอกสารบันทึกข้อความที่เกี่ยวข้องกับการทารุณเด็ก การสัมภาษณ์เจ้าหน้าที่ และการเข้าร่วมสังเกตุการณ์การทำงานของทีมสหวิชาชีพ

1. ศึกษาความรู้ที่เกี่ยวข้องกับการให้ความช่วยเหลือเด็กที่ถูกทารุณ

จากการศึกษาประเภทของการทารุณเด็กทั้งหมดมี 4 ประเภท ได้แก่ การทำร้ายร่างกาย การความรุนแรงทางเพศ การทำร้ายจิตใจและการทอดทิ้ง ปล่อยปละละเลย พบว่าการทำร้ายร่างกายกับความรุนแรงทางเพศเป็นการทารุณเด็กที่พบมากที่สุด ในขณะที่การทำร้ายจิตใจและการทอดทิ้งไม่ค่อยได้พบแบบอิสระ แต่มักจะเกิดร่วมกับการทำร้ายร่างกายหรือความรุนแรงทางเพศ ผู้วิจัยจึงเลือกศึกษาเฉพาะกรณีทำร้ายร่างกายและความรุนแรงทางเพศ

2. การเก็บรวบเอกสาร

ผู้วิจัยได้รวบรวมบันทึกเอกสารที่เกี่ยวข้องกับการทารุณเด็กที่เป็นผู้ป่วยของโรงพยาบาลรามาธิบดี ซึ่งบันทึกโดยนักสังคมสงเคราะห์ นักจิตวิทยา และแพทย์ บันทึกที่เกี่ยวข้องกับการทารุณเด็กแบ่งเป็น 7 ส่วนดังนี้

1. บันทึกข้อมูลทั่วไปของตัวเด็ก

2. บันทึกจากการรับแจ้งเหตุ

3. บันทึกจากการสัมภาษณ์เด็กและผู้เกี่ยวข้องเพิ่มเติม

4. บันทึกผลการตรวจร่างกาย

5. บันทึกผลการตรวจไอคิว

6. บันทึกการวางแผนให้ความช่วยเหลือ

7. บันทึกผลความก้าวหน้าในการติดตามให้ความช่วยเหลือ

ข้อมูลทั้งหมดเป็นข้อมูลแบบไม่มีโครงสร้างและเป็นข้อมูลเชิงคุณภาพที่ไม่ได้ถูกจัดเก็บในฐานข้อมูล เนื่องจากยังไม่มีระบบที่ใช้ในการจัดเก็บฐานข้อมูล ทำให้มีข้อมูลไม่ครบถ้วน จำเป็นต้องสอบถามเจ้าหน้าที่เพิ่มเติมเพื่อเติมข้อมูลส่วนที่ขาดหายไปให้สมบูรณ์

3.1.2 การสัมภาษณ์

ใช้วิธีการสัมภาษณ์จากผู้ให้คำตอบโดยตรงเป็นการเก็บข้อมูลแบบปฐมภูมิ เนื่องจากบันทึกที่ได้จากการเก็บรวบรวมข้อมูลยังมีความไม่สมบูรณ์ และระบบการทำงานเป็นลักษณะการใช้ตัวบุคคลเป็นหลักซึ่งข้อมูลจะอยู่ที่ตัวบุคคลที่เป็นผู้ดูแลผู้ป่วยที่ตนรับผิดชอบ โดยผู้วิจัยได้แบ่งการสัมภาษณ์ออกเป็น 2 ส่วนดังนี้

1. สัมภาษณ์เจ้าหน้าที่ดูแลผู้ป่วยเพื่อสอบถามข้อมูลเพิ่มเติม เนื่องจากข้อมูลที่ทำการเก็บรวบรวมมานั้นเป็นบันทึกที่เขียนลงในกระดาษ ข้อมูลไม่ได้ถูกจัดเก็บไว้ที่เดียวกันทั้งหมด และข้อมูลส่วนใหญ่อยู่ที่ตัวบุคคล การเก็บรวบรวมข้อมูลจากการบันทึกงานเพียงอย่างเดียวจึงไม่สามารถเก็บข้อมูลที่สมบูรณ์ได้ทั้งหมด ผู้วิจัยจำเป็นต้องสอบถามเจ้าหน้าที่เพิ่มเติมในแต่ละเคสโดยสัมภาษณ์ผู้ดูแลรับผิดชอบเคสโดยตรงซึ่งเป็นการสอบถามแบบปากเปล่าตามลักษณะข้อมูลของแต่ละเคส

2. สัมภาษณ์เจ้าหน้าที่ถึงปัญหาการทำงานในปัจจุบัน ผู้วิจัยได้ทำการเก็บรวบรวมปัญหาในการทำงานรวมถึงความต้องการในระบบเพื่อนำข้อมูลที่ได้ไปปรับปรุงพัฒนาระบบเว็บแอพพลิเคชั่นสำหรับงานเด็กที่ถูกทารุณโดยใช้คำถามปลายเปิดเพื่อให้ผู้ตอบสามารถตอบคำถามได้ครอบคลุม และเสนอมุมมอง แนวคิด คำแนะนำ ในเรื่องการพัฒนาระบบได้อย่างเต็มที่ คำถามที่ถามเป็นคำถามเกี่ยวกับการทำงานในปัจจุบัน ขั้นตอนการทำงาน หน้าที่ความรับผิดชอบของผู้ตอบคำถาม ปัญหาอุปสรรคในการจัดเก็บและการค้นหาข้อมูล และความคิดเห็นเกี่ยวกับการนำเทคนิคเหมืองข้อความมาใช้เป็นเครื่องมือช่วยในการหาแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณ

ผู้วิจัยได้ทำการสัมภาษณ์บุคลากรในโรงพยาบาลดังนี้ นักสังคมสงเคราะห์กลุ่มเด็กที่ถูกทารุณ กุมารแพทย์และนักจิตวิทยาสำหรับเด็ก เนื่องจากเป็นผู้ที่เกี่ยวข้องกับระบบโดยตรง

3.1.3 การสังเกตการณ์

ผู้วิจัยได้เข้าร่วมสังเกตการทำงานของเจ้าหน้าที่โดยเป็นการสังเกตการณ์แบบไม่มีส่วนร่วม เพื่อดูวิธีการทำงาน การเก็บข้อมูลจากตัวเด็ก การประเมินปัญหาที่เกิดกับเด็ก การสรุปผลเพื่อเตรียมดำเนินการให้ความช่วยเหลือ และการวางแผนให้ความช่วยเหลือ

3.1.4 การเตรียมข้อมูล

หลังจากการรวบรวมข้อมูลที่จะนำมาใช้ในการวิจัยแล้ว จะทำการเตรียมข้อมูลเพื่อใช้ในการทำเหมืองข้อความ โดยเริ่มจากการจัดเก็บข้อมูลที่ได้ลงในฐานข้อมูล และนำข้อมูลที่เก็บไว้มาเข้าสู่การประมวลผลภาษาทางธรรมชาติ (Natural language processing) วิเคราะห์ชนิดของคำ (Part of speech) และจัดสร้างคลังคำ (Words corpus) ทำการสกัดคุณลักษณะ (Feature extraction) เพื่อนำไปใช้สร้างตัวแบบ (Model) สำหรับการจำแนกประเภท (Classification)

3.1.5 การสร้างตัวแบบ

ในขั้นตอนนี้จะเป็นการนำอัลกอริทึมที่ได้ศึกษามาทำงานกับกลุ่มตัวอย่างในการเรียนรู้เพื่อนำมาใช้สร้างตัวแบบ และทดสอบความเหมาะสมของตัวแบบ โดยจะเลือกตัวแบบที่ให้ผลแม่นยำที่สุดมาใช้ในการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณต่อไป ในส่วนนี้จะใช้โปรแกรม RapidMiner Studio Version 8.0 ในการทดสอบตัวแบบ โดยรายละเอียดของผลลัพธ์ที่ได้จากการสร้างตัวแบบจะอธิบายในบทที่ 4

3.1.6 การประเมิน

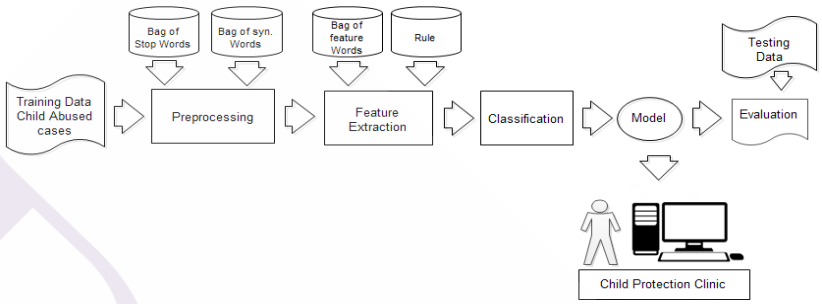
เป็นการประเมินประสิทธิภาพของตัวแบบ โดยนำผลลัพธ์ที่ได้จากเทคนิคที่ใช้ในการสร้างตัวแบบมาวิเคราะห์หาเทคนิคที่ให้ผลได้ถูกต้องแม่นยำที่สุดเพื่อใช้ในการวิเคราะห์หาแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณ

3.1.7 สรุปผลการวิจัยและข้อเสนอแนะ

เป็นการสรุปผลของการดำเนินงานวิจัยทั้งหมด ในการประยุกต์ใช้เทคนิคเหมืองข้อความเพื่อวิเคราะห์หาแนวทางการให้ความช่วยเหลือเด็กที่ถูกทารุณ เพื่อช่วยในการตัดสินใจของเจ้าหน้าที่ และลดระยะเวลาของการตัดสินใจเพื่อสามารถให้ความช่วยเหลือเด็กที่ถูกทารุณได้รวดเร็วขึ้น พร้อมทั้งข้อเสนอแนะสำหรับการนำไปประยุกต์ใช้ในด้านอื่นต่อไป

3.2 สถาปัตยกรรมการทำงานของระบบ

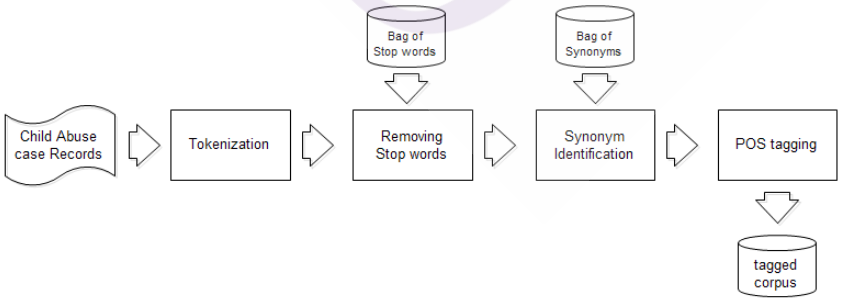
ในส่วนของสถาปัตยกรรมการทำงานของระบบจะเป็นส่วนการเตรียมข้อมูล (Preprocessing) การสกัดคุณลักษณะ (Feature extraction) การสร้างตัวแบบ (Model) โดยใช้เทคนิคการจำแนกประเภท (Classification) และการประเมินผลของข้อมูลที่ใช้เรียนรู้ (Evaluation) ตาม



ภาพที่ 3.1 สถาปัตยกรรมการทำงานของระบบ

3.2.1 การเตรียมข้อมูล (Preprocessing)

ขั้นตอนนี้เริ่มตั้งแต่การนำข้อมูลบันทึกที่เกี่ยวข้องกับการทารุณเด็กที่เก็บอยู่ในฐานข้อมูลโดยข้อมูลแต่ละเคสประกอบด้วย บันทึกเหตุการณ์ ประวัติผู้ป่วย ข้อมูลครอบครัว/ผู้ดูแล ข้อมูลโรงเรียน/ชุมชน และ ข้อมูลผู้กระทำทารุณ เข้าสู่กระบวนประมวลผลภาษาธรรมชาติ (Natural language processing) เริ่มตั้งแต่ การตัดคำ (Word segmentation) การกำจัดคำหยุด (Removing Stop words) และการหาคำพ้องความ (Synonym Identification)



ภาพที่ 3.2 ขั้นตอนการเตรียมข้อมูล

1. การตัดคำ

การตัดคำเริ่มจากการบันทึกข้อมูลเข้าสู่ระบบ โดยข้อมูลจะเป็นเอกสารบันทึกเกี่ยวกับเด็กที่ถูกทารุณทั้งหมด จำนวน 60 เอกสารและมีการจัดประเภทไว้ว่าเอกสารนั้นมีการให้ความช่วยเหลือเป็นประเภทสงเคราะห์หรือคุ้มครองซึ่งได้ถูกจัดแบ่งประเภทโดยนักสังคมสงเคราะห์ หลังจากบันทึกข้อมูลแล้วระบบจะทำการตัดคำ (Tokenized)

ตารางที่ 3.1 ตัวอย่างการตัดคำ

บันทึกข้อความ	ก่อนหน้านี้ เด็กอยู่กับปะป๊า (พ่อเลี้ยง) แม่ พี่บี(13) พี่ซี (นามสมมติ)(11) เรียน ป.3 รร.แห่งหนึ่ง ที่ต้องมาอยู่บ้านพักเด็กเพราะถูกพี่ชายข่มขืน
การตัดคำ	ก่อน | หน้า | นี้ | เด็ก | อยู่ | กับ | ปะป๊า | พ่อเลี้ยง | แม่ | พี่ | บี | 13 | พี่ | ซี | นามสมมติ | 11 | เรียน | ป. 3 | ร.ร. | แห่ง | หนึ่ง | ที่ | ต้อง | มา | อยู่ | บ้านพักเด็ก | เพราะ | ถูก | พี่ชาย | ข่มขืน |

2. การกำจัดคำหยุด (Removing Stop words) เป็นการตัดคำที่ไม่มีความหมายในประโยคออกไป เพื่อไม่ให้กระทบกับจำนวนของคำ จากตารางที่ 3.1 ตัวอย่างการตัดคำได้มีการกำจัด Stop word ออกไปบางส่วน เช่นวงเล็บ คำลงท้ายเช่น “ครับ” “ค่ะ” คำบุพบท เช่น “ที่” “ซึ่ง” “อัน” ซึ่งนอกจากการกำจัด Stop word ในขั้นตอนการตัดคำแล้ว ยังได้เพิ่มรายการคำที่ต้องการตัดออกลงไป (Bag of stop words) เพื่อให้สามารถกำจัดคำที่ต้องการได้มากขึ้น

ตารางที่ 3.2 ตัวอย่างการกำจัดคำที่ไม่จำเป็นในประโยค (Stop word)

ประโยคที่ถูกตัดคำแล้ว	ก่อน | หน้า | นี้ | เด็ก | อยู่ | กับ | ปะป๊า | พ่อเลี้ยง | แม่ | พี่ | เรียน | ป.3 | ร.ร.| แห่ง | หนึ่ง | ที่ | ต้อง | มา | อยู่ | บ้านพักเด็ก | เพราะ | ถูก | พี่ชาย | ข่มขืน |
เพิ่มรายการ Stop word	ค่ะ ครับ ที่ ซึ่ง อัน

ตารางที่ 3.2 (ต่อ)

แสดงผลการกำจัดคำที่ไม่ต้องการ	ก่อน | หน้า | นี้ | เด็ก | อยู่ | กับ | ปะป๊า | พ่อเลี้ยง | แม่ | พี่ | เรียน | ป 3 | รร| แห่ง | หนึ่ง | ต้อง | มา | อยู่ | บ้านพัก เด็ก | เพราะ | ถูก | พี่ชาย | ข่มขืน |

3. การหาคำพ้องความ (Synonym Identification) ข้อมูลเด็กที่ถูกทารุณเป็นบันทึกโดยเจ้าหน้าที่ โดยบันทึกตามคำบอกเล่าของผู้ป่วย และผู้เกี่ยวข้องจำนวนมาก โดยแต่ละเคสมีผู้บันทึกหลายคน ซึ่งแต่ละคนมีการใช้คำที่แตกต่างกันในคำที่มีความหมายเดียวกัน จึงมีการหาคำพ้องความโดยในเป็นการแทนคำที่เขียนต่างกันแต่มีความหมายเหมือนกันด้วยคำในถุงคำพ้องความ (Bag of Synonyms) เพื่อลดความซับซ้อนในการประมวลผล และลดความฟุ่มเฟือยของคำ ได้แสดงตัวอย่างการหาคำพ้องความ ในตารางที่ 3.3 มีรายละเอียดดังนี้

ตารางที่ 3.3 ตัวอย่างการหาคำพ้องความ

คำที่ได้จากการตัดคำและการกำจัดคำหยุด	ก่อน | หน้า | นี้ | เด็ก | อยู่ | กับ | ปะป๊า | พ่อเลี้ยง | แม่ | พี่ชาย | เรียน | ป 3 | รร| แห่ง | หนึ่ง | ที่ | ต้อง | มา | อยู่ | บ้านพักเด็ก | เพราะ | ถูก | พี่ชาย | ข่มขืน |
ตัวอย่างการหาคำพ้องความ	พ่อ / ป๊า / ปะป๊า / ป๋า / บิดา = พ่อ
ผลการหาคำพ้องความ	ก่อน | หน้า | นี้ | เด็ก | อยู่ | กับ | พ่อ | พ่อเลี้ยง | แม่ | เรียน | ป 3 | รร| แห่ง | หนึ่ง | ที่ | ต้อง | มา | อยู่ | บ้านพักเด็ก | เพราะ | ถูก | พี่ชาย | ข่มขืน |

4. การกำกับชนิดของคำ (Part of speech : POS) ในงานวิจัยนี้ได้อ้างอิงการกำกับหน้าที่คำจาก Orchid Corpus ซึ่งเป็นคลังคำภาษาไทยที่รวบรวมและจำแนกชนิดคำโดย NECTEC หน้าที่ของคำสามารถแบ่งได้ตามตารางที่ 3.4

ตารางที่ 3.4 ชนิดของคำ (Part of speech : POS)

ลำดับ	POS Tag	คำอธิบาย	ตัวอย่าง
1	ADVI	คำวิเศษณ์ที่เกิดจากการใช้คำซ้ำ	เสมอๆ ค่อยๆ
2	ADVN	คำวิเศษณ์แท้	เร็ว ช้า สม่ำเสมอ
3	ADVP	คำวิเศษณ์ที่เกิดจากการเติมคำนำหน้า	โดยเร็ว
4	ADVS	คำวิเศษณ์ที่ใช้ขึ้นต้นประโยค	โดยปกติ
5	CFQC	ลักษณะนามบอกความถี่	ครั้ง เที่ยว
6	CLTV	ลักษณะนามบอกหมวดหมู่	คู่ กลุ่ม ฝูง
7	CMTR	ลักษณะนามบอกปริมาณ	กิโลกรัม ชั่วโมง
8	CNIT	ลักษณะนามบอกหน่วยวัด	คน อัน เล่ม
9	CVBL	ลักษณะนามกริยา	ขด รอบ มัด ม้วน
10	DCNM	คำกำกับนามที่แสดงจำนวนนับ	1 คน แมว 2 ตัว
11	DDAC	คำกำกับนามชี้เฉพาะที่ใช้ร่วมกับลักษณะนาม	นี้ นั้น โน้น
12	DDAN	คำกำกับนามชี้เฉพาะที่ตามหลังคำนามแบบไม่มีลักษณะร่วมด้วย	นี่ นั่น โน่น
13	DDAQ	คำกำกับนามชี้เฉพาะที่แสดงจำนวน	พอดี ถ้วน
14	DDBQ	คำกำกับนามชี้เฉพาะที่เชื่อมระหว่างคำนาม คำลักษณะนาม หรือคำที่แสดงจำนวน	ทั้ง อีก เพียง
15	DIAC	คำนำหน้านามที่ไม่ชี้เฉพาะเจาะจง ที่ตามหลังคำนาม มีการให้เลือกระหว่างสิ่งใดสิ่งหนึ่ง	ไหน อื่น ต่างๆ
16	DIAQ	คำนำหน้านามที่ไม่ชี้เฉพาะเจาะจงตามหลังคำแสดงจำนวน	กว่า เศษ
17	DIBQ	คำนำหน้านามที่ไม่ชี้เฉพาะเจาะจง อยู่ระหว่างคำนาม และแสดงถึงจำนวน	บาง ประมาณ เกือบ

ตารางที่ 3.4 (ต่อ)

ลำดับ	POS Tag	คำอธิบาย	ตัวอย่าง
16	DIAQ	คำนำหน้านามที่ไม่ชี้เฉพาะเจาะจง ตามหลังคำแสดงจำนวน	กว่า เศษ
17	DIBQ	คำนำหน้านามที่ไม่ชี้เฉพาะเจาะจง อยู่ระหว่างคำนาม และแสดงถึงจำนวน	บาง ประมาณ เกือบ
18	DONM	จำนวนเชิงอันดับที่	1 2 3
19	EAFF	คำลงท้ายประโยคบอกเล่า	ครับ ค่ะ
20	EITT	คำลงท้ายประโยคคำถาม	ใช่ไหม
21	FIXN	คำอุปสรรค	การ ความ
22	FIXV	คำอุปสรรคนำหน้าคำวิเศษณ์	อย่าง (อย่างเร็ว)
23	INT	คำอุทาน	โอ๊ะ อุ๊ย
24	JCMP	คำเปรียบเทียบ	กว่า เหมือนกับ
25	JCRG	คำสันธานเชื่อมประโยค	แต่ และ หรือ
26	JSBR	คำสันธานเชื่อมประโยคใจความรองเข้ากับประโยคใจความหลัก	เพราะว่า หลังจาก
27	NCMN	คำสามัญนาม	หนังสือ อาหาร
28	NCNM	จำนวนนับ	1 2 3
29	NEG	คำปฏิเสธ	ไม่
30	NLBL	ตัวเลข	1 2 3 a b c
31	NONM	ลำดับที่	ที่ 1 ที่ 2 ที่ 3
32	NPRP	คำนามเฉพาะ	โค้ก โตโยต้า
33	NTTL	คำนำหน้าชื่อ	ดร. พลเอก
34	PDMN	สรรพนามจำเพาะเจาะจง	นี่ นั่น โน่น
35	PNTR	สรรพนามคำถาม	ใคร อะไร อย่างไร
36	PPRS	บุรุษสรรพนาม	ฉัน เธอ เขา
37	PREL	สรรพนามเชื่อมความ	ที่ ซึ่ง โดย
38	PUNC	เครื่องหมายวรรคตอน	“ ;

ตารางที่ 3.4 (ต่อ)

ลำดับ	POS Tag	คำอธิบาย	ตัวอย่าง
39	RPRE	คำบุพบท	บน ล่าง
40	VACT	กริยาที่แสดงการกระทำ	ทำงาน ร้องเพลง
41	VATT	กริยาแสดงคุณลักษณะ	สูง ต่ำ อ้วน ผอม
42	VSTA	กริยาแสดงสภาวะ	เห็น รู้ คือ
43	XVAE	กริยาช่วย	ไป มา ขึ้น
44	XVAM	กริยาช่วย ตามหลังคำปฏิเสธ "ไม่"	ค่อย ได้ น่า
45	XVBB	กริยาช่วยที่ใช้ในการแนะนำหรือออกคำสั่ง	กรุณา เชิญ
46	XVBM	กริยาช่วย นำหน้าคำปฏิเสธ "ไม่"	เกิด เกือบ กำลัง
47	XVMM	กริยาช่วย นำหน้าและตามหลังคำปฏิเสธ "ไม่"	ควร เคย ต้อง

ตารางที่ 3.5 ตัวอย่างการกำกับหน้าที่ของคำ

ก่อนหน้านี้|NCMN เด็ก|NCMN อยู่|VSTA กับ|RPRE พ่อ|NCMN พ่อแม่|NCMN พี่|NCMN เรียน|VACT ป|NCMN 3|NLBL รร|NCMN แห่ง|CNIT หนึ่ง|DCNM ต้อง|XVMM มา|XVAE อยู่|VSTA บ้านพักเด็ก|NCMN เพราะ|JSBR ถูก|XVAM พี่ชาย| NCMN ข่มขืน| VACT

3.2.2 การสกัดคุณลักษณะ

การสกัดคุณลักษณะเพื่อลดขนาดของจำนวนคำ และเพิ่มความรวดเร็ว ในขั้นตอนนี้นิยมใช้การหาความถี่ของคำ หรือวิธีการทางสถิติเพื่อสกัดคุณลักษณะ ในงานวิจัยนี้ ผู้เชี่ยวได้กำหนดคำที่บ่งบอกเหตุการณ์ที่มีความสำคัญต่อการจำแนกประเภทของการทารุณกรรม โดยสามารถแบ่งได้เป็น 6 คุณลักษณะกลุ่มหลัก (Main feature groups) ซึ่งประกอบด้วย 50 คุณลักษณะกลุ่มย่อย (Sub Feature groups) เพื่อให้มีจำนวนคุณลักษณะน้อยลง และเพิ่มความถูกต้องในการทำงานของระบบ โดยคุณลักษณะกลุ่มหลัก สามารถแสดงได้ตามตารางที่ 3.6 คุณลักษณะกลุ่มย่อย สามารถแสดงตาม

ตารางที่ 3.7 โดยคำที่ใช้ในการพิจารณาในการจัดเข้ากลุ่มของแต่ละคุณลักษณะย่อยสามารถแสดงได้ดังภาคผนวก ก

ตารางที่ 3.6 กลุ่มหลัก (Main feature groups)

รหัสกลุ่มหลัก	หัวข้อ	รายละเอียด
1	Abused records	กลุ่ม Feature ที่แสดงการถูกกระทำ
2	Patient	กลุ่ม Feature ที่เกี่ยวข้องกับตัวเด็ก
3	Abuser	กลุ่ม Feature ที่เกี่ยวข้องกับตัวผู้กระทำ
4	Family / Surrounding	กลุ่ม Feature ที่เกี่ยวข้องกับครอบครัว สภาพแวดล้อมที่เด็กอยู่อาศัย
5	Wound	กลุ่ม Feature ที่แสดงถึงลักษณะบาดแผล อากาบาดเจ็บ
6	Child Protect Activity	กลุ่ม Feature ที่เกี่ยวข้องกับการวางแผนป้องกัน ดูแลเด็กหลังเกิดเหตุ

ตารางที่ 3.7 กลุ่มย่อย (Feature groups)

รหัสกลุ่มหลัก	รหัสกลุ่มย่อย	รายละเอียด	พิจารณาจาก
1	G101	ให้เด็กดูสื่อลามก พูดจาแทะโลม	บันทึกเหตุการณ์
1	G102	สัมผัสบริเวณที่ไม่เหมาะสม	บันทึกเหตุการณ์
1	G103	การมีเพศสัมพันธ์ หรือการสอดใส่สิ่งของ / อวัยวะต่างๆ เข้าภายในช่องคลอดเด็ก	บันทึกเหตุการณ์
1	G104	ลงโทษไม่เหมาะสม ตีรุนแรง ใช้อาวุธขนาดเล็ก	บันทึกเหตุการณ์
1	G105	การทำร้ายร่างกายชัดเจนที่อาจก่อให้ เกิดการเสียชีวิต บาดเจ็บหนักหรือพิการ การใช้อาวุธในการทำร้ายร่างกาย	บันทึกเหตุการณ์

ตารางที่ 3.7 (ต่อ)

รหัสกลุ่มหลัก	รหัสกลุ่มย่อย	รายละเอียด	พิจารณาจาก
1	G106	ทอดทิ้ง ปล่อยปละละเลย	บันทึกเหตุการณ์
1	G107	การค้ามนุษย์ ใช้แรงงานเด็ก	บันทึกเหตุการณ์
1	G108	การทำร้ายจิตใจ บังคับ	บันทึกเหตุการณ์
1	G109	การข่มขู่	บันทึกเหตุการณ์
2	G201	ปัญหาสุขภาพ	บันทึกประวัติผู้ป่วย
2	G202	เด็กที่ต้องได้รับการดูแลเป็นพิเศษ	บันทึกประวัติผู้ป่วย
2	G203	มีลักษณะถูกละเลย เช่น เนื้อตัวมอมแมม ไม่ได้เรียนหนังสือ ผอม ไม่มีคนดูแล	บันทึกประวัติผู้ป่วย
2	G204	ปัญหาพฤติกรรม	บันทึกประวัติผู้ป่วย
2	G205	ปัญหาการเรียน	บันทึกข้อมูลครอบครัว/ผู้ดูล
2	G206	โรคจิตเวช	บันทึกประวัติผู้ป่วย
2	G207	ผลกระทบทางจิตใจหลังถูกทารุณ	บันทึกประวัติผู้ป่วย
2	G208	เกี่ยวข้องกับยาเสพติด เครื่องดื่มแอลกอฮอล์	บันทึกประวัติผู้ป่วย
2	G209	ปัญหาด้านพัฒนาการของเด็ก	บันทึกประวัติผู้ป่วย
2	G210	ความซุกซน	บันทึกประวัติผู้ป่วย
2	G211	ภาวะทางอารมณ์ของเด็ก	บันทึกประวัติผู้ป่วย
2	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติต่างๆ	บันทึกประวัติผู้ป่วย
3	G301	โรคจิตเวช ของผู้กระทำ	บันทึกข้อมูลผู้กระทำ
3	G302	ภาวะทางอารมณ์	บันทึกข้อมูลผู้กระทำ
3	G303	ผู้กระทำเกี่ยวข้องกับยาเสพติด	บันทึกข้อมูลผู้กระทำ

ตารางที่ 3.7 (ต่อ)

รหัสกลุ่มหลัก	รหัสกลุ่มย่อย	รายละเอียด	พิจารณาจาก
3	G304	ผู้กระทำเป็นคนในครอบครัว	บันทึกข้อมูลผู้กระทำ
3	G309	ปัญหาความรุนแรงในครอบครัว	บันทึกข้อมูลผู้กระทำ
3	G310	ความเสี่ยงที่จะถูกกระทำทารุณจากความไม่พร้อมของครอบครัว	บันทึกข้อมูลผู้กระทำ
3	G312	การพนัน	บันทึกข้อมูลผู้กระทำ
3	G313	ดื่มเหล้า สูบบุหรี่	บันทึกข้อมูลผู้กระทำ
4	G401	สภาพครอบครัว เช่น แยกกันอยู่ หย่าร้าง	บันทึกข้อมูลครอบครัว
4	G402	ปัญหาความรุนแรงในครอบครัว	บันทึกข้อมูลครอบครัว
4	G403	ปัญหายาเสพติดภายในครอบครัว	บันทึกข้อมูลครอบครัว
4	G404	การพนัน	บันทึกข้อมูลครอบครัว
4	G405	ปัญหาการเลี้ยงดูเด็ก	บันทึกข้อมูลครอบครัว
4	G406	ปัญหาเศษรฐกิจ	บันทึกข้อมูลครอบครัว
4	G407	คนในครอบครัวประกอบอาชีพไม่เหมาะสม	บันทึกข้อมูลครอบครัว

ตารางที่ 3.7 (ต่อ)

รหัสกลุ่มหลัก	รหัสกลุ่มย่อย	รายละเอียด	พิจารณาจาก
4	G408	ภาวะทางอารมณ์ของคนในครอบครัว	บันทึกข้อมูลครอบครัว
4	G409	เหล้า	บันทึกข้อมูลครอบครัว
4	G410	ไม่สามารถปกป้องคุ้มครองเด็กได้	บันทึกข้อมูลครอบครัว
4	G411	สภาพแวดล้อมในบ้านไม่เหมาะสม	บันทึกข้อมูลครอบครัว
4	G412	ปัญหาสุขภาพ	บันทึกข้อมูลครอบครัว
4	G413	ภาวะทางจิต	บันทึกข้อมูลครอบครัว
4	G414	ผู้ต้องขัง	บันทึกข้อมูลครอบครัว
5	G501	รอยฟกช้ำ อาการเจ็บ บาดแผลทั่วไป	บันทึกเหตุการณ์
5	G502	บาดแผล อาการบาดเจ็บรุนแรง หรือรอยแผลเป็นที่แสดงถึงการได้รับบาดเจ็บรุนแรง	บันทึกเหตุการณ์
5	G503	บาดแผล หรืออาการบาดเจ็บที่เกี่ยวข้อง กับการถูกล่วงละเมิดทางเพศ	บันทึกเหตุการณ์
6	G601	มีการวางแผนดูแลเด็กเพื่อป้องกันการถูกกระทำซ้ำ	บันทึกข้อมูลครอบครัว
6	G602	ผู้ปกครองหรือผู้ดูแลเด็กมีความเข้าใจในตัวเด็ก หรือพร้อมทำความเข้าใจเด็ก	บันทึกข้อมูลครอบครัว
6	G603	มีการดำเนินคดีกับผู้กระทำแล้ว	บันทึกข้อมูลครอบครัว

ตารางที่ 3.7 (ต่อ)

รหัสกลุ่มหลัก	รหัสกลุ่มย่อย	รายละเอียด	พิจารณาจาก
6	G604	ผู้ปกครองหรือผู้ดูแลเด็กมีความสามารถที่จะปกป้องเด็กได้	บันทึกข้อมูลครอบครัว

ตัวอย่างการพิจารณาคุณลักษณะ เช่นถ้าในบันทึกเหตุการณ์ประกอบด้วยประโยค “ได้ประวัติว่าถูกพี่ชายกระทำอนาจาร” เคสนี้จะมีคุณลักษณะกลุ่มย่อย คือ G304 (ผู้กระทำเป็นคนในครอบครัว เนื่องจากมีคำว่า “พี่ชาย”) และ กลุ่มย่อย G102 (สัมผัสบริเวณที่ไม่เหมาะสม เนื่องจากมีคำว่า “กระทำอนาจาร”) ตามรายละเอียดในภาคผนวก ก และมีการใช้กฎ จำนวน 19 กฎ ซึ่งสร้างโดยผู้เชี่ยวชาญ เพื่อพิจารณายุบรวมคุณลักษณะที่เกี่ยวข้อง ตามรายละเอียดในภาคผนวก ข

นอกจากนี้ การพิจารณาคุณลักษณะจะพิจารณาร่วมกับการเกิดคำปฏิเสธด้วย โดยพิจารณาคำที่เกิดก่อนหน้าจำนวน 4 คำว่าประกอบด้วยคำปฏิเสธหรือไม่ ถ้ามีคำปฏิเสธ จะไม่พิจารณาว่าเกิดคุณลักษณะนั้น ตัวอย่างดังตารางที่ 3.8

ตารางที่ 3.8 ตัวอย่างการพิจารณาคุณลักษณะร่วมกับคำปฏิเสธ

ลำดับ	Case No	คำศัพท์	Case Type	คำปฏิเสธ	คุณลักษณะกลุ่มย่อย	ค่าคุณลักษณะที่ได้
1	1	มี	บันทึกเหตุการณ์			
2	1	รอยช้ำ	บันทึกเหตุการณ์			501(1)
3	1	มี	บันทึกเหตุการณ์			
4	1	แผลถลอก	บันทึกเหตุการณ์			501(1)
5	1	ไม่มี	บันทึกเหตุการณ์	1		
6	1	พฤติกรรม	บันทึกเหตุการณ์			
7	1	เล่นรุนแรง	บันทึกเหตุการณ์		204	204(-1)

ตารางที่ 3.9 ตัวอย่างกฎที่ใช้ในการจำแนกประเภท

NO	IF	THEN	ความหมาย
1	(G104 = 0 AND G105 = 1) OR (G104 = 1 AND G105 = 1)	1	มีการทำร้ายร่างกาย - G104 คือ การลงโทษไม่เหมาะสม เช่น การใช้ไม้ตีตามร่างกาย - G105 คือ การทำร้ายร่างกายชัดเจน ที่ส่งผลต่อชีวิต หรือการบาดเจ็บรุนแรง เช่น จับหัวกดน้ำ ใช้เหล็กทุบหัว
2	(G108 = 1) OR (G109 = 1)	1	การทำร้ายจิตใจ - G108 คือ การบังคับจิตใจ เช่น บังคับให้เด็มซ้อมมวย พ่อไม่ให้ลูกพบแม่ - G109 คือ การข่มขู่ เช่น ห้ามบอกใคร ถ้าบอกใครจะตีให้ตาย

จากตารางที่ 3.9 จะเป็นการนำคุณลักษณะที่ได้มาสร้างเป็นกฎ โดยคุณลักษณะที่ได้ยังไม่มีความชัดเจนในการนำไปใช้จำแนกประเภท ต้องมีคุณลักษณะอื่นประกอบด้วย จึงจะแสดงเหตุการณ์สำหรับใช้ในการจำแนกประเภทได้ชัดเจนขึ้น รายละเอียดกฎทั้งหมดแสดงได้ดังภาคผนวก ข

3.2.3 การจำแนกประเภทข้อมูล

ในงานวิจัยนี้ได้เลือกเทคนิคการจำแนกประเภทข้อมูลข้อมูล 3 เทคนิค เพื่อใช้สร้างตัวแบบ และประเมินว่าตัวแบบที่สามารถจำแนกข้อมูลได้แม่นยำที่สุด เทคนิคที่ใช้มีดังนี้

1. Decision tress หรือแผนภาพต้นไม้เพื่อการตัดสินใจ เป็นการสร้างข้อมูลแบบลำดับขั้นที่แสดงถึงการตัดสินใจลำดับต่างๆ โดยลำดับสุดท้าย หรือส่วนของโหนดใบ (leaf node) จะแสดงกลุ่ม (class) ตามที่จัดไว้ ซึ่งเป็นผลลัพธ์ของการทำนายหรือการตัดสินใจแบบลำดับขั้น โดยมีลักษณะการตัดสินใจแบบ if/then/else ซึ่งเป็นลักษณะของกฎ และง่ายต่อการนำมาประยุกต์ใช้ในการเขียนโปรแกรม

2. Naïve Bayes เป็นเทคนิคที่นิยมใช้ในการจำแนกประเภทข้อมูลโดยประยุกต์ใช้หลักการทางสถิติมาช่วยในการตัดสินใจ

3. Support Vector Machine (SVM) เป็นการเรียนรู้แบบใหม่โดยใช้ทฤษฎีสถิติวิเคราะห์และจำแนกข้อมูล โดยจะหาระนาบตัดสินใจที่อยู่กึ่งกลางระหว่างข้อมูลซึ่งมีระยะห่างระหว่างข้อมูลทั้งสองกลุ่มมากที่สุดมีข้อดีคือสามารถลดปัญหาการเกิด Overfitting ของข้อมูลได้ และเทคนิคที่มีประสิทธิภาพในการจำแนกประเภทมากที่สุดเทคนิคหนึ่ง

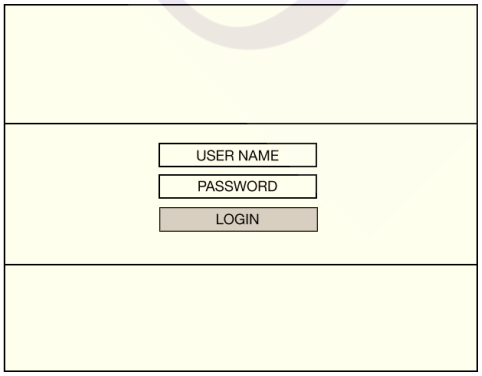
3.2.4 การประเมินผล

ในการประเมินผล เป็นการประเมินประสิทธิภาพของตัวแบบ โดยปกติจะนิยมใช้ Cross-validation test ในการประเมิน ในงานวิจัยนี้เลือกใช้ 10 Fold-Cross-validation ทำการประเมิน เนื่องจากการแบ่งข้อมูลสำหรับการทดสอบออกเป็นกลุ่มเล็กๆ และทำการทดสอบหลายครั้งจะให้ผลการทดสอบที่ถูกต้องกว่า และง่ายต่อการสร้างโมเดลทดสอบ โดยแบ่งข้อมูลเป็นสองชุด คือ Training data กับ Test data ในข้อมูลแต่ละชุดจะแบ่งข้อมูลออกเป็น10 ส่วนเท่าๆ กัน โดยจะแบ่ง 1 ส่วนเป็นข้อมูลที่ใช้ในการเรียนรู้ และอีก 9 ส่วนเป็นข้อมูลที่ใช้สำหรับทดสอบ ทำวนไปจนครบ 10 รอบ

3.3 การพัฒนาระบบ Child Protection Clinic

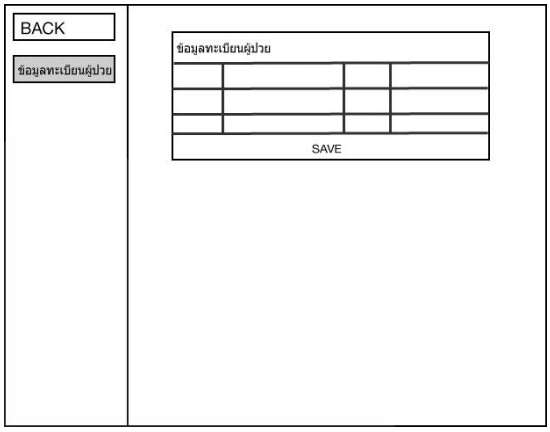
ผลลัพธ์ที่ได้จากการทดลองผู้วิจัยได้นำโมเดลที่ได้มาใช้ในระบบ โดยได้พัฒนาระบบเพื่อช่วยประเมินการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณในเบื้องต้น โดยมีการออกแบบหน้าจอดังนี้

3.3.1 หน้า Login



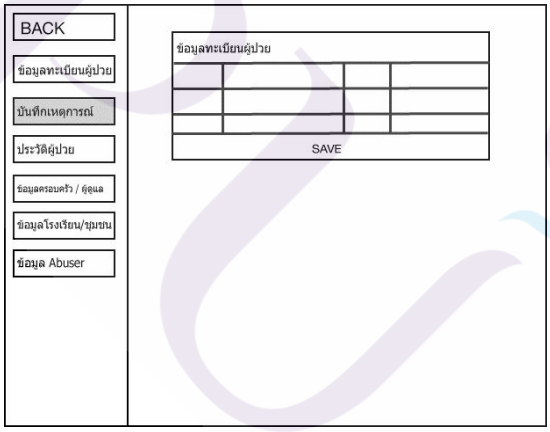
ภาพที่ 3.3 การออกแบบหน้าจอ Login

3.3.2 หน้าบันทึกข้อมูลทะเบียนผู้ป่วย



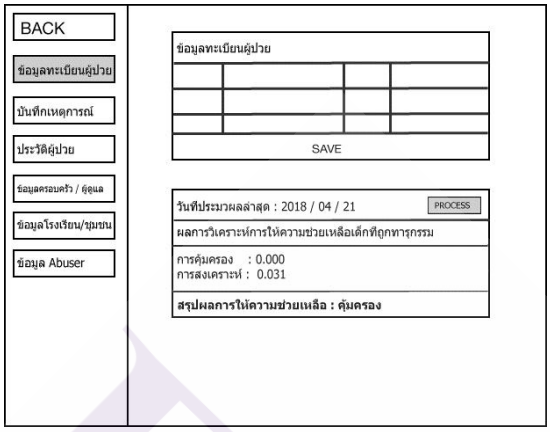
ภาพที่ 3.4 การออกแบบหน้าบันทึกข้อมูลทะเบียนผู้ป่วย

3.3.3 หน้าบันทึกเคส



ภาพที่ 3.5 การออกแบบหน้าบันทึกเคส

3.3.4 หน้าบันสรุปผลการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรม



ภาพที่ 3.6 การออกแบบหน้าบันสรุปผลการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรม

3.4 เครื่องมือที่ใช้ในการดำเนินงานวิจัย

3.4.1 ใช้ฐานข้อมูล SQLserver 2012 สำหรับจัดเก็บข้อมูลบันทึกเด็กที่ถูกทารุณ

3.4.2 ใช้ภาษา Java ในการพัฒนาระบบ

3.4.3 ใช้ Lucene Tokenizer ซึ่งเป็น Java Library ในการตัดคำ

3.4.4 ใช้ Orchid Corpusในการกำกับชนิดคำ

3.4.5 ใช้โปรแกรม RapidMiner Studio 8 ในกาสร้างตัวแบบ

บทที่ 4
ผลการวิจัย

ในบทนี้เป็นการอธิบายการทดสอบและผลการทดสอบการวัดประสิทธิภาพของตัวแบบในการจำแนกประเภทการให้ความเหลือเด็กที่ถูกทารุณกรรม

4.1 ข้อมูลที่นำมาใช้ในการทดสอบ

จำนวนตัวอย่างบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็กที่นำมาใช้ในการทดสอบ มีทั้งหมด 60 ตัวอย่าง แบ่งเป็น 2 ประเภท คือ Class A : การสงเคราะห์เด็ก จำนวน 12 ตัวอย่าง และ Class B : การคุ้มครองสวัสดิภาพ จำนวน 48 ตัวอย่าง แต่ละตัวอย่างประกอบด้วยข้อมูลบันทึกเหตุการณ์ ข้อมูลบันทึกประวัติผู้ป่วย ข้อมูลครอบครัว/ผู้ดูแล ข้อมูลโรงเรียน/ชุมชน และข้อมูลผู้กระทำทารุณ

4.2 ผลการวิจัย

ในงานวิจัยนี้ ได้ใช้เทคนิค 3 ประเภทในการทดสอบ คือ เทคนิค Decision Tree เทคนิค Naïve Bayes และ เทคนิค Support Vector Machine ผลการทดสอบประสิทธิภาพของตัวแบบที่ได้จากเทคนิคทั้ง 3 แสดงดังตารางที่ 4.1

ตารางที่ 4.1 ผลการทดสอบประสิทธิภาพของตัวแบบ

Model	Precision			Recall			F-score
	Class A	Class B	Total	Class A	Class B	Total	
Decision Tree	92.31%	100.00%	98.46%	100.00%	97.92%	98.34%	98.39%
Naïve Bayes	100.00%	97.96%	98.37%	91.67%	100.00%	98.33%	98.35%
SVM	92.31%	100.00%	98.46%	100.00%	97.92%	98.34%	98.39%

ตารางที่ 4.2 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Decision Tree

	True B	True A	Class Precision
Predict B	47	0	100.00%
Predict A	1	12	92.31%
Class Recall	97.92%	100.00%	

ตารางที่ 4.3 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Naïve Bayes

	True B	True A	Class Precision
Predict B	48	1	97.96%
Predict A	0	11	100.00%
Class Recall	100.00%	91.67%	

ตารางที่ 4.4 ตาราง Confusion Matrix ของตัวแบบที่ได้จากเทคนิค Support Vector Machine

	True B	True A	Class Precision
Predict B	47	0	100.00%
Predict A	1	12	92.31%
Class Recall	97.92%	100.00%	

จากผลการทดสอบ ดังตารางที่ 4.1 Decision Tree กับ Support Vector Machine ให้ค่า Precision Recall และ F-score ดีที่สุด คือ 92.31% 98.34% และ 98.39% ตามลำดับ โดยตัวแบบทั้งสองให้ค่าเท่ากัน และผลการประเมินจากตาราง Confusion Matrix ในตารางที่ 4.2 ตารางที่ 4.3 และตารางที่ 4.4 ของตัวแบบ Decision Tree Naïve Bayes และ Support Vector Machine ตามลำดับ จะพบว่าตัวแบบที่ได้จากเทคนิค Naïve Bayes สามารถทำนาย Class B ได้ถูกต้องทั้งหมด 48 ตัวอย่าง

งานวิจัยนี้ให้ความสำคัญกับการทำนาย class B : การคุ้มครองสวัสดิภาพ ให้ถูกต้อง ซึ่งไม่ต้องการให้ทำนายผิดไปเป็น class A เนื่องจากกระบวนการให้ความช่วยเหลือเด็กตาม พรบ. คุ้มครองเด็ก พ.ศ. 2546 มีข้อกำหนดเรื่องอำนาจของเจ้าหน้าที่ในการให้ความช่วยเหลือแตกต่างกัน โดยการคุ้มครองสวัสดิภาพ ได้ให้อำนาจเจ้าหน้าที่ในการแยกเด็กออกจากผู้ดูแล และเป็นการให้

ความช่วยเหลือที่เร่งด่วน หากระบบทำนายประเภทการให้ความช่วยเหลือใน Class B ไม่ถูกต้องจะทำให้เกิดความล่าช้า และตัวเด็กมีความเสี่ยงที่จะถูกกระทำทารุณซ้ำได้สูง จึงเลือก Naïve Bayes เป็นตัวแบบที่นำมาใช้พัฒนาระบบ เนื่องจากตัวแบบนี้ให้ค่า False Positive ของ Class B เป็น 0 หรือสามารถทำนายการเกิด Class B ได้ถูกต้องทั้งหมด

4.3 การวิเคราะห์ผลการทดลอง

จากผลการทดสอบตัวแบบที่ได้จากเทคนิค Naïve Bayes ตัวแบบสามารถจำแนกประเภทใน Class B ได้ถูกต้องทั้งหมด 48 ตัวอย่าง และจำแนก Class A ได้ถูกต้อง 11 ตัวอย่าง จากตัวอย่างทั้งหมด 12 ตัวอย่าง โดยมี 1 ตัวอย่างที่ถูกจำแนกผิดเป็น Class B โดยเมื่อตรวจสอบข้อมูลพบว่า ตัวอย่างที่จำแนกผิดเป็นตัวอย่างเด็กที่ถูกทำร้ายร่างกาย ซึ่งผู้ป่วยเป็นเด็กพิเศษ มีการบันทึกรายละเอียดของบาดแผลและลักษณะการถูกทำร้าย รวมถึงสภาพครอบครัวที่พ่อแม่หย่าร้างกัน แต่ในส่วนบันทึกข้อมูลการแก้ปัญหา การปรับความเข้าใจของผู้ปกครอง และการวางแผนดูแลเด็กยังมีรายละเอียดไม่เพียงพอ จึงทำให้ระบบจะแนกประเภทผิด

จากการศึกษาข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณกรรมเด็ก ผู้วิจัยได้แบ่งคุณลักษณะเป็น 6 กลุ่มใหญ่ ซึ่งประกอบด้วยคุณลักษณะกลุ่มย่อย 50 กลุ่ม โดยลักษณะของการทารุณเด็กจะมีรูปแบบเหตุการณ์ที่มักเกิดขึ้นด้วยกัน เช่น เด็กที่ถูกทารุณกรรมทางเพศ มักจะเกิดจากปัญหาครอบครัว ปัญหายาเสพติด ปัญหาสภาพแวดล้อม และปัญหาพฤติกรรมของเด็ก

Naïve Bayes เป็นเทคนิคที่ใช้ทฤษฎีความน่าจะเป็นที่จะเกิดเหตุการณ์ A และมีเหตุการณ์ B ตามมา โดยจะใช้ความน่าจะเป็นที่มากที่สุดของสมมติฐานในการทำนาย Class เทคนิค Naïve Bayes มีความเหมาะสมกับข้อมูลขนาดใหญ่ สามารถเรียนรู้เพิ่มได้ โดยจะปรับรูปแบบการทำนายความน่าจะเป็นเมื่อมีการเรียนรู้ใหม่เพิ่มเข้ามา แต่ยังผนวกกับความรู้เดิมที่เรียนรู้ก่อนหน้า และมีความเหมาะกับข้อมูลที่มีความหลากหลาย

จากการศึกษาข้อมูลบันทึกเหตุการณ์และเทคนิคการจำแนกข้อมูล Naïve Bayes ผู้วิจัยจึงเลือกตัวแบบ Naïve Bayes มาใช้ในการพัฒนาระบบ

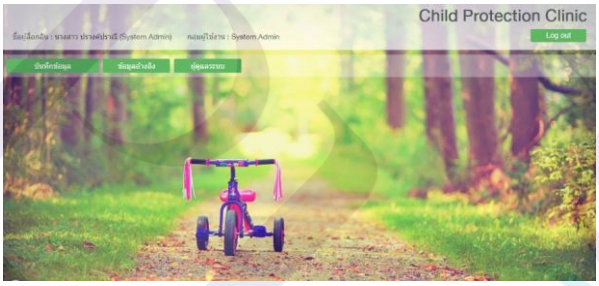
4.4 ระบบจำแนกการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรม

ผู้วิจัยได้เลือกตัวแบบ Naïve Bayes มาใช้ในการพัฒนาระบบเพื่อช่วยประเมินการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณ โดยเป็นระบบเว็บแอปพลิเคชัน พัฒนาด้วยภาษา Java และเก็บข้อมูลในฐานข้อมูล SQL Server version 2012 มีรายละเอียดดังนี้

4.4.1 หน้าแรกของระบบจะเป็นหน้าจอสำหรับลงชื่อเข้าใช้ระบบ (Login)

ภาพที่ 4.1 หน้าจอ Login

เมื่อผู้ใช้ลงชื่อเข้าใช้ระบบแล้วจะพบกับหน้าหลักของระบบ ซึ่งประกอบด้วยเมนูบันทึกข้อมูล เมนูข้อมูลอ้างอิง และเมนูผู้ดูแลระบบ

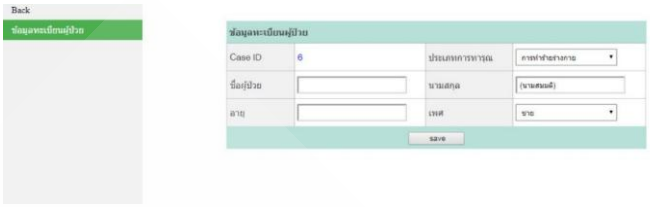


ภาพที่ 4.2 หน้าจอหลัก Login

เมื่อคลิกที่เมนูบันทึกข้อมูล ระบบจะแสดงหน้ารายการผู้ป่วย ประกอบด้วย ส่วนค้นหาข้อมูลและ ส่วนแสดงรายการผู้ป่วย

ภาพที่ 4.3 หน้าจอแสดงรายการผู้ป่วย

เมื่อคลิกปุ่ม เพิ่มข้อมูลใหม่ ระบบจะแสดงหน้าจอการเพิ่มข้อมูลใหม่ ซึ่งผู้ใช้ต้องกรอกข้อมูลพื้นฐานของผู้ป่วย ซึ่งประกอบด้วย Case ID ประเภทการทารุณ ชื่อผู้ป่วย นามสกุล อายุ และเพศ

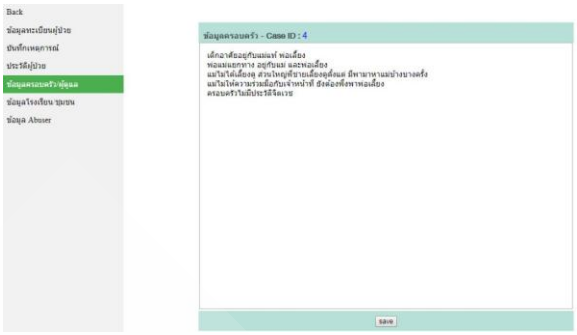


ภาพที่ 4.4 หน้าจอการเพิ่มข้อมูลใหม่

เมื่อผู้ใช้กรอกข้อมูลพื้นฐานของผู้ป่วยเสร็จ แล้วกดปุ่ม save ระบบจะแสดงเมนูให้กรอกข้อมูลที่เกี่ยวข้อง ซึ่งได้แก่ ข้อมูลบันทึกเหตุการณ์ ข้อมูลประวัติผู้ป่วย ข้อมูลครอบครัว/ผู้ดูแล ข้อมูลโรงเรียน/ชุมชน ข้อมูล Abuser ตามภาพที่ 4.5 และ 4.9 ตามลำดับ

ภาพที่ 4.5 หน้าจอบันทึกเหตุการณ์

ภาพที่ 4.6 หน้าจอประวัติผู้ป่วย



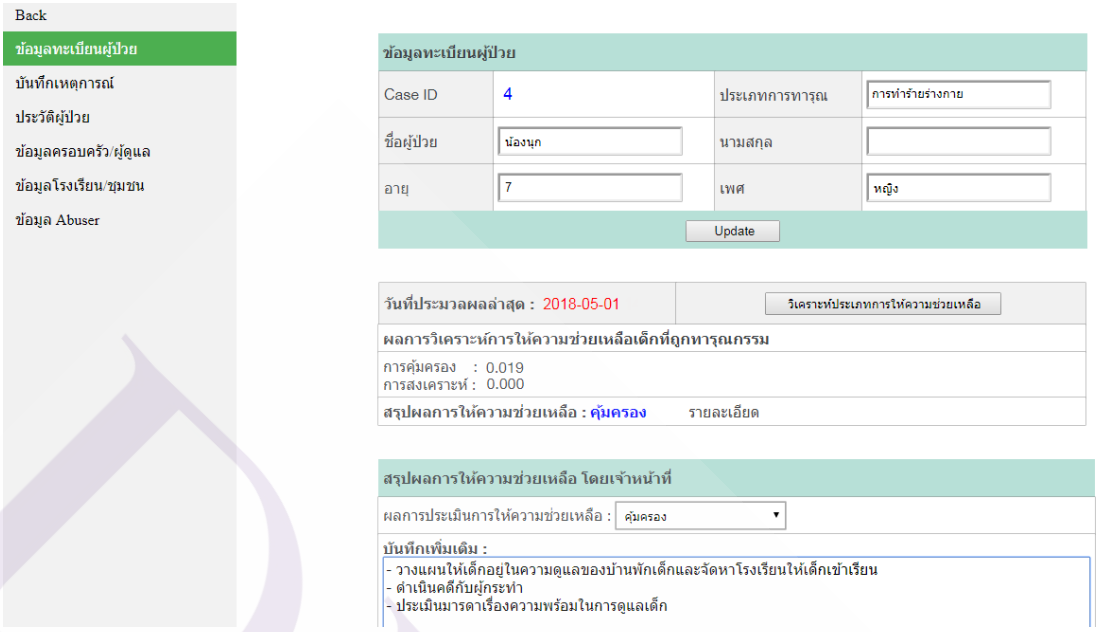
ภาพที่ 4.7 หน้าจอข้อมูลครอบครัว/ผู้ดูแล

ภาพที่ 4.8 หน้าจอข้อมูลโรงเรียน/ชุมชน

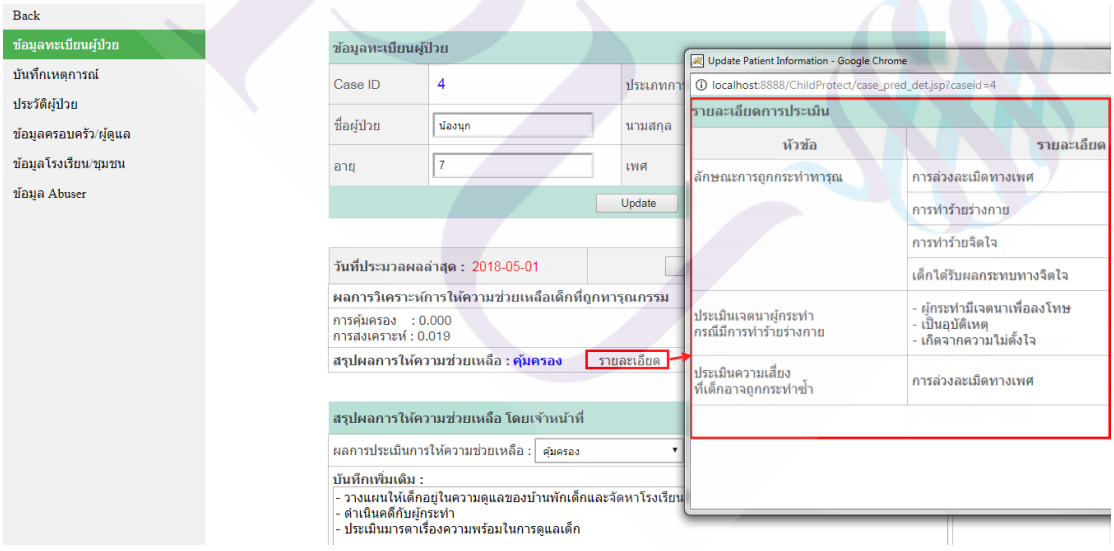
ภาพที่ 4.9 หน้าจอข้อมูล Abuser

เมื่อผู้ใช้กรอกรายละเอียดในหน้าจอต่างๆ ครบถ้วน ให้มาที่หน้าแรก โดยคลิกที่เมนูข้อมูลทะเบียนผู้ป่วย และคลิกที่ปุ่ม วิเคราะห์ประเภทการให้ความช่วยเหลือ ระบบจะทำการวิเคราะห์ข้อมูลจากรายละเอียดต่างๆ ของผู้ป่วย และแสดงผลการวิเคราะห์ที่ได้ ผู้ใช้สามารถคลิกปุ่ม

รายละเอียด เพื่อดูรายละเอียดการประเมิน ดังแสดงในภาพที่ 4.11 และสามารถใส่ข้อมูลสรุปผลการให้ความช่วยเหลือ โดยเจ้าหน้าที่ เพื่อเก็บข้อมูลการให้ความช่วยเหลือที่ได้ปฏิบัติจริง



ภาพที่ 4.10 หน้าจอผลการวิเคราะห์ประเภทการให้ความช่วยเหลือ



ภาพที่ 4.11 หน้าจอแสดงรายละเอียดการประเมิน

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

5.1 สรุปการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเทคนิคการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมด้วยเทคนิคการทำเหมืองข้อความ โดยการวิเคราะห์ข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็ก ซึ่งอยู่ในรูปของการบันทึกข้อความ ประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมสามารถแบ่งได้ 2 ประเภทใหญ่ๆ คือ การสงเคราะห์เด็ก กับการคุ้มครองสวัสดิภาพเด็ก โดยเอกสารบันทึกเหตุการณ์จะถูกนำมาผ่านการประมวลผลภาษาธรรมชาติก่อนเพื่อสกัดคุณลักษณะ โดยแบ่งคุณลักษณะเป็นกลุ่มใหญ่ 6 กลุ่ม และกลุ่มย่อย 50 กลุ่ม มีคำที่ถูกเลือกเป็นคุณลักษณะทั้งหมด 692 คำ และมีการใช้กฎ 19 กฎ เทคนิคการจำแนกประเภทที่นำมาทดสอบมี 3 เทคนิค คือ Decision Tree, Naïve Bayes และ Support Vector Machine แบ่งประเภทการให้ความช่วยเหลือเป็น 2 ประเภทคือการสงเคราะห์เด็กกับการคุ้มครองสวัสดิภาพเด็ก จากผลการทดสอบพบว่า Naïve Bayes ให้ค่าดีที่สุด จากผลการทดลองสามารถสรุป ปัญหาอุปสรรคและข้อจำกัดของงานวิจัย ข้อเสนอแนะ โดยมีรายละเอียดดังต่อไปนี้

5.2 ปัญหาอุปสรรคและข้อจำกัดของงานวิจัย

1. ในการบันทึกข้อมูลยังไม่ได้ถูกบันทึกในระบบคอมพิวเตอร์ โดยข้อมูลที่ถูกบันทึกนั้นเป็นการบันทึกลงในกระดาษ ซึ่งเป็นการเขียนโดยตัวผู้บันทึกเอง ทำให้มีความลำบากในการอ่าน

2. ข้อมูลที่นำมาใช้มีจำนวนน้อย โดยมีเพียง 60 ตัวอย่าง เนื่องจากเวลาที่จำกัด และปัญหาอุปสรรคในการรวบรวมข้อมูล ทำให้มีคำที่จำกัด

3. การตัดคำและการกำกับชนิดคำ ยังต้องนำไฟล์เอกสารไปทำการตัดคำและกำกับชนิดคำเอง

5.3 ข้อเสนอแนะ

ในการพัฒนาระบบการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมด้วยการใช้เทคนิคการทำเหมืองข้อความครั้งต่อไป มีข้อเสนอแนะดังนี้

1. สามารถนำข้อมูลที่ได้ไปวิเคราะห์เพื่อใช้ในการป้องกันก่อนเกิดเหตุ
2. สามารถเชื่อมโยงกับข้อมูลผู้ป่วยของโรงพยาบาลได้

DPU

บรรณานุกรม

บรรณานุกรม

ภาษาไทย

นันทกา สวัสดิพานิช. (2555). *"ความรุนแรงต่อเด็ก : ความท้าทายใหม่สำหรับพยาบาล".* พยาบาลสาธารณสุข ปีที่ 26 ฉบับที่ 1 2555.

นงลักษณ์ เทพสวัสดิ์. (2540). *ทฤษฎีและการปฏิบัติงานสังคมสงเคราะห์.* กรุงเทพฯ: มหาวิทยาลัยธรรมศาสตาร์.

พรรณี ประจวบเหมาะ. (2523). *การสังคมสงเคราะห์*. กรุงเทพฯ: เจริญกิจ .

สายชล สินสมบูรณ์ทอง. (2558). *การทำเหมืองข้อมูล Data Mining*. กรุงเทพฯ: จามจุรีโปรดักส์.

สายชล สินสมบูรณ์ทอง. (2559). *การทำเหมืองข้อมูล เล่ม 2 วิธีการและตัวแบบ Data Mining 2 : Methods and Models*. กรุงเทพฯ: จามจุรีโปรดักส์.

สุนีย์ ยอดเยี่ยม.(2554).*Child abuse or neglect*. สืบค้น 30 พฤษภาคม 2559, จาก http://med.mahidol.ac.th/ramamental/sites/default/files/public/pdf/Child%20abuse.pdf.

อุปกิตศิลปสาร, พระยา. (2491). *หลักภาษาไทย*. กรุงเทพฯ: ไทยวัฒนาพานิช.

อมรา ประสิทธิ์รัฐสินธุ์. (2543). *ชนิดของคำในภาษาไทย : การวิเคราะห์ทางวกยสัมพันธ์โดยอาศัยฐานข้อมูลภาษาไทยปัจจุบันสองล้านคำ*.กรุงเทพมหานคร : สำนักงานคณะกรรมการวิจัย.

ภาษาต่างประเทศ

Fred Popowich. (2005). *Using Text Mining and Natural Language Processing for Health Care Claims Processing*. SIGKDD Explorations Volume 7 Issue 1, 59-66.

Margaret H. Dunham. (2003). *Data Mining Introductory and Advanced topics*. USA: Person Education, Inc..

Mehmed Kantardzie. (2003). *Data Mining Concepts, Models, Methods, And Algorithms*. USA: A John Wiley & Son, Inc..

Meliha Yetisgen-Yildiz, Cosmin Adrian Bejan, Mark M. Wurfel. (2013).*Identification of Patients with Acute Lung Injury from Free-Text Chest X-Ray Reports*.Workshop on Biomedical

Natural Language Processing (BioNLP 2013), 10-17.

Mihir Shekhar, Veera Raghavendra Chikka, Lini Thomas, et al.(2015). *Identifying Medical Terms Related to Specific Diseases*. 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 170-177.

Qiwei He, et al.(2015). *Automated Assessment of Patients' Self-Narratives for Posttraumatic Stress Disorder Screening Using Natural Language Processing and Text Mining*. Assessment, 24, 157-172.

Simon Baker, et al. (2015). *Automatic Semantic Classification of Scientific Literature According to the Hallmarks of Cancer*. Oxford University Press 2015, 1-9 .

Taxiachris Botsis, et al.(2011). *Text Mining for the Vaccine Adverse Event Reporting System: medical classification using informative feature selection*. J Am Med inform Assoc 2011, 631-63.

ภาคผนวก

ภาคผนวก ก
กลุ่มคุณลักษณะและคำที่แสดงคุณลักษณะ

ตารางที่ ก.1 คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	ให้เด็กดูวิดีโอโป๊	G101	ให้เด็กดูสื่อลามก พูดจาแทะโลม
2	ขอมีอะไรกัน	G101	ให้เด็กดูสื่อลามก พูดจาแทะโลม
3	พูดจีบ	G101	ให้เด็กดูสื่อลามก พูดจาแทะโลม
4	กอด	G102	สัมผัสบริเวณที่ไม่เหมาะสม
5	กอดจูบ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
6	การสัมผัสทางเพศไม่เหมาะสม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
7	การสัมผัสอวัยวะเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
8	จับตามตัวหน้าอกของสงวน	G102	สัมผัสบริเวณที่ไม่เหมาะสม
9	จับนม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
10	จับอวัยวะเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
11	จูบ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
12	ถูกสัมผัสไม่เหมาะสม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
13	ถูไถด้านนอก	G102	สัมผัสบริเวณที่ไม่เหมาะสม
14	ทำอะไรไม่เหมาะสม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
15	บีบอวัยวะเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
16	ปล้ำ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
17	ลวนลาม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
18	ลูบคลำ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
19	ลูบคลำจิ๋ม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
20	ลูบคลำช่องคลอด	G102	สัมผัสบริเวณที่ไม่เหมาะสม
21	ลูบคลำอวัยวะเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
22	เล่นไม่เหมาะสม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
23	สัมผัสไม่เหมาะสม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
24	สัมผัสอวัยวะเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
25	หอมแก้ม	G102	สัมผัสบริเวณที่ไม่เหมาะสม
26	หอมคอ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
27	เอากระจู๋มาชน	G102	สัมผัสบริเวณที่ไม่เหมาะสม
28	เอาอวัยวะเพศมาโดน	G102	สัมผัสบริเวณที่ไม่เหมาะสม
29	ล่วงละเมิด	G102	สัมผัสบริเวณที่ไม่เหมาะสม
30	ล่วงละเมิดทางเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
31	สัมผัสหน้าอก	G102	สัมผัสบริเวณที่ไม่เหมาะสม
32	การถูกล่วงละเมิด	G102	สัมผัสบริเวณที่ไม่เหมาะสม
33	จับหน้าอก	G102	สัมผัสบริเวณที่ไม่เหมาะสม
34	โดนละเมิด	G102	สัมผัสบริเวณที่ไม่เหมาะสม
35	ถูกล่วงละเมิดทางเพศ	G102	สัมผัสบริเวณที่ไม่เหมาะสม
36	ทำอนาจาร	G102	สัมผัสบริเวณที่ไม่เหมาะสม
37	กระทำชำเรา	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
38	กระทำอนาจาร	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
39	กลั้นปัสสาวะไม่อยู่	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
40	การสอดใส่ทางทวารหนัก	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
41	การสอดใส่อวัยวะเพศ	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
42	ข่มขืน	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
43	คุมกำเนิด	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
44	ใช้ถุงยางอนามัย	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
45	ใช้นิ้ว	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
46	ใช้ปาก	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
47	ใช้ลิ้น	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
48	ตั้งครรภ์	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
49	แตกใน	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
50	ถูกกระทำอนาจาร	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
51	ถูกข่มขืน	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
52	พยายามข่มขืน	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
53	ไม่ใช้ถุงยาง	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
54	ยาคุม	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
55	ยุติการตั้งครรภ์	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
56	เอาอวัยวะเพศใส่ปาก	G103	มีเพศสัมพันธ์ หรือสอดใส่ของ
57	กัด	G104	ลงโทษไม่เหมาะสม
58	ใช้มือบิดหู	G104	ลงโทษไม่เหมาะสม
59	ใช้ไม้ฟาดตามตัว	G104	ลงโทษไม่เหมาะสม
60	ทำโทษ	G104	ลงโทษไม่เหมาะสม
61	ทำโทษรุนแรง	G104	ลงโทษไม่เหมาะสม
62	ตีด้วยไม้	G104	ลงโทษไม่เหมาะสม
63	หยิก	G104	ลงโทษไม่เหมาะสม

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
64	ไม้	G104	ลงโทษไม่เหมาะสม
65	ถีบ	G104	ลงโทษไม่เหมาะสม
66	ตบ	G104	ลงโทษไม่เหมาะสม
67	ตี	G104	ลงโทษไม่เหมาะสม
68	ใช้กำลัง	G104	ลงโทษไม่เหมาะสม
69	ใช้ความรุนแรง	G104	ลงโทษไม่เหมาะสม
70	เอารองเท้าฟาดหน้า	G104	ลงโทษไม่เหมาะสม
71	ไม้แขวนเสื้อ	G104	ลงโทษไม่เหมาะสม
71	ตีด้วยไม้แขวนเสื้อ	G104	ลงโทษไม่เหมาะสม
72	โขกพื้น	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
73	จับหัวกดน้ำ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
74	เอาไม้แขวนเสื้อรูดปาก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
75	หยิกอวัยวะเพศ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
76	กระชากหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
77	กระทืบหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
78	กระบอง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
79	ขว้างของใส่	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
80	คีมหนีบ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
81	จับศีรษะโขกกำแพง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
82	จิกผม	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
83	ชกต่อย	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
84	ใช้แขนล็อคคอ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
85	ใช้ไม้ตีหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
86	ใช้ลวดตีหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
87	ใช้สายยางตี	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
88	ใช้อาวุธ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
89	โดนบุหรี่จี้	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
90	ตบหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
91	ต่อยที่ท้อง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
92	ตะหลิว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
93	ตีด้วยเข็มขัด	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
94	ตีด้วยไม้บรรทัดเหล็ก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
95	ตีด้วยไม้หน้าสาม	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
96	ตีด้วยสายไฟ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
97	ตีที่หลัง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
98	ตีรุนแรง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
99	ตีหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
100	เตะท้อง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
101	เตะที่ศรีษ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
102	เตะหลัง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
103	เตะหัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
104	เตารีดนาบ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
105	ถีบหน้าอก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
106	ถีบหลัง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
107	ท่อน้ำ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
108	ท่อแป๊บ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
109	ทับบนตัว	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
110	ที่ล็อกล้อรถ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
111	ทุบตี	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
112	แท่งเหล็ก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
113	บีบคอ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
114	ผลักล้ม	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
115	มัดแขน	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
116	มัดด้วยเชือก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
117	มัดไว้กับเตียง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
118	มีด	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
119	ไม้บรรทัดเหล็ก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
120	รุมทำร้าย	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
121	ลวดสลิง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
122	ลากลงบันได	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
123	สลิง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
124	สาก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
125	สายไฟ	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
126	เหยียบขา	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
127	เหล็ก	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
128	เอาขวดปา	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
129	เอาเชือกรัด	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
130	เอามือมัดไขว้หลัง	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
131	เอามุ้งมามัด	G105	การทำร้ายร่างกายรุนแรง และอาวุธที่ใช้ทำร้าย
132	ขังในห้องน้ำ	G106	ทอดทิ้ง ปล่อยปละละเลย
133	บังคับให้ต่อยมวย	G106	ทอดทิ้ง ปล่อยปละละเลย
134	ขัง	G106	ทอดทิ้ง ปล่อยปละละเลย

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
135	ให้นอนนอกบ้าน	G106	ทอดทิ้ง ปล่อยปละละเลย
136	ไม่เคยได้รับการตรวจประเมิน	G106	ทอดทิ้ง ปล่อยปละละเลย
137	ไม่ได้ฉัดวัคซีน	G106	ทอดทิ้ง ปล่อยปละละเลย
138	พ่อแม่เด็กไม่สนใจ	G106	ทอดทิ้ง ปล่อยปละละเลย
139	ไม่ให้กินข้าว	G106	ทอดทิ้ง ปล่อยปละละเลย
140	ไม่ให้เรียนหนังสือ	G106	ทอดทิ้ง ปล่อยปละละเลย
141	เลี้ยงดูเด็กไม่เหมาะสม	G106	ทอดทิ้ง ปล่อยปละละเลย
142	เร่ร่อน	G106	ทอดทิ้ง ปล่อยปละละเลย
143	ดูแลเด็กไม่เหมาะสม	G106	ทอดทิ้ง ปล่อยปละละเลย
144	กินอาหารหมา	G106	ทอดทิ้ง ปล่อยปละละเลย
145	กินข้าวกับหมา	G106	ทอดทิ้ง ปล่อยปละละเลย
146	เด็กเร่ร่อน	G106	ทอดทิ้ง ปล่อยปละละเลย
147	เลี้ยงดูไม่เหมาะสม	G106	ทอดทิ้ง ปล่อยปละละเลย
148	ยินยอมให้เด็กหมั้น	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
149	ให้สินสอด	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
150	ขายบริการ	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
151	ค้าประเวณี	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
152	แม่เลี้ยงต้องการเอาไว้ใช้งาน	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
153	ให้เด็กมาช่วยขายของ	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
154	ขายเด็กให้นายหน้า	G107	การค้ามนุษย์ ใช้แรงงานเด็ก
155	ถูกขาย	G107	การค้ามนุษย์ ใช้แรงงานเด็ก

ตารางที่ ก.1 (ต่อ)

คุณลักษณะกลุ่มหลัก 1- ข้อมูลการกระทำทารุณ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
156	ให้อยู่ในกรอบ	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
157	มีปากเสียง	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
158	ว่าเสียงดัง	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
159	คำพูดแรง	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
160	ด่า	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
161	ด่าหยาบคาย	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
162	ชี้หน้าด่า	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
163	บังคับให้ต่อยมวย	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
164	บังคับให้ขอเงิน	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
165	พูดให้เจ็บใจ	G108	การทำร้ายจิตใจ บังคับ ข่มขู่
166	มีการขู่	G109	การข่มขู่
167	เดี๋ยวจะเอามีดแทง	G109	การข่มขู่
168	จะเอาค้อนทุบหัว	G109	การข่มขู่
169	ขู่	G109	การข่มขู่
170	ขู่ฆ่า	G109	การข่มขู่
171	ข่มขู่	G109	การข่มขู่
172	ห้ามบอกใคร	G109	การข่มขู่
173	ห้ามไปบอกใคร	G109	การข่มขู่
174	อย่าไปบอกใคร	G109	การข่มขู่

ตารางที่ ก.2 คุณลักษณะกลุ่มหลัก 2 – ข้อมูลประวัติผู้ป่วย

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
64	ไม้	G104	ลงโทษไม่เหมาะสม
1	ชัก	G201	ปัญหาสุขภาพ
2	เป็นไทรอยด์	G201	ปัญหาสุขภาพ
3	สมาธิสั้น	G202	เด็กที่ต้องได้รับการดูแลเป็นพิเศษ
4	Difficult child	G202	เด็กที่ต้องได้รับการดูแลเป็นพิเศษ
5	คนพิการ	G202	เด็กที่ต้องได้รับการดูแลเป็นพิเศษ
6	adhd	G202	เด็กที่ต้องได้รับการดูแลเป็นพิเศษ
7	ไม่ได้ไปโรงเรียน	G203	มีลักษณะถูกทอดทิ้ง ปล่อยปละละเลย
8	เนื้อตัวมอมแมม	G203	มีลักษณะถูกทอดทิ้ง ปล่อยปละละเลย
9	ผอม	G203	มีลักษณะถูกทอดทิ้ง ปล่อยปละละเลย
10	เร่ร่อน	G203	มีลักษณะถูกทอดทิ้ง ปล่อยปละละเลย
11	ขโมย	G204	ปัญหาพฤติกรรม
12	หนีเรียน	G204	ปัญหาพฤติกรรม
13	ชู้สาว	G204	ปัญหาพฤติกรรม
14	ฉุยน้ำลายให้คนอื่นกิน	G204	ปัญหาพฤติกรรม
15	เล่นรุนแรง	G204	ปัญหาพฤติกรรม
16	มีเพศสัมพันธ์	G204	ปัญหาพฤติกรรม
17	หนีเที่ยว	G204	ปัญหาพฤติกรรม

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
18	หนีเรียน	G204	ปัญหาพฤติกรรม
19	หนีออกจากบ้าน	G204	ปัญหาพฤติกรรม
20	อยู่กับผู้ชาย	G204	ปัญหาพฤติกรรม
21	จูบ	G204	ปัญหาพฤติกรรม
22	ชอบแย่งของ	G204	ปัญหาพฤติกรรม
23	เถียง	G204	ปัญหาพฤติกรรม
24	ทะเลาะวิวาท	G204	ปัญหาพฤติกรรม
25	เที่ยวเล่น	G204	ปัญหาพฤติกรรม
26	นัดเจอผู้ชาย	G204	ปัญหาพฤติกรรม
27	บุหรี่	G204	ปัญหาพฤติกรรม
28	พูดคำหยาบ	G204	ปัญหาพฤติกรรม
29	พูดไม่ฟัง	G204	ปัญหาพฤติกรรม
30	มีปัญหาพฤติกรรม	G204	ปัญหาพฤติกรรม
31	มีปัญหาพฤติกรรมทางเพศ	G204	ปัญหาพฤติกรรม
32	มีพฤติกรรมก้าวร้าว	G204	ปัญหาพฤติกรรม
33	ไม่กลับบ้าน	G204	ปัญหาพฤติกรรม
34	กลับบ้านดึก	G204	ปัญหาพฤติกรรม
35	หนีออกจากบ้าน	G204	ปัญหาพฤติกรรม
36	แกล้งป่วย	G204	ปัญหาพฤติกรรม
37	ใช้ความรุนแรง	G204	ปัญหาพฤติกรรม
38	สูบบุหรี่	G204	ปัญหาพฤติกรรม

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
39	ใช้ยาแก้ไอ	G204	ปัญหาพฤติกรรม
40	ซิ่งรถมอเตอร์ไซค์	G204	ปัญหาพฤติกรรม
41	โดดเรียน	G204	ปัญหาพฤติกรรม
42	ตบตี	G204	ปัญหาพฤติกรรม
43	ติดเพื่อน	G204	ปัญหาพฤติกรรม
44	แต่งกายเปิดเผย	G204	ปัญหาพฤติกรรม
45	มีเพศสัมพันธ์สมยอม	G204	ปัญหาพฤติกรรม
46	เรียนไม่ดี	G205	ปัญหาการเรียน
47	การเรียนตก	G205	ปัญหาการเรียน
48	สติปัญญาไม่ดี	G205	ปัญหาการเรียน
49	ไม่ค่อยสนใจเรียน	G205	ปัญหาการเรียน
50	อ่านเขียนไม่ค่อยได้	G205	ปัญหาการเรียน
51	อ่านหนังสือไม่ออก	G205	ปัญหาการเรียน
52	เขียนไม่ได้	G205	ปัญหาการเรียน
53	เขียนหนังสือไม่ค่อยได้	G205	ปัญหาการเรียน
54	เรียนไม่เก่ง	G205	ปัญหาการเรียน
55	เรียนไม่ค่อยดี	G205	ปัญหาการเรียน
56	เรียนอ่อน	G205	ปัญหาการเรียน
60	มีปัญหาสุขภาพจิต	G206	โรคจิตเวช
61	รักษารพ.ศรีธัญญา	G206	โรคจิตเวช
62	รู้สึกแย่	G207	ผลกระทบทางจิตใจหลังถูกทารุณ

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
63	ร้องไห้	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
64	วิตกกังวล	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
65	โกรธ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
66	หวาดผวา	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
67	เสียใจ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
68	มีความคิดรุนแรง	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
69	ฝันร้าย	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
70	นอนไม่ค่อยหลับ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
71	อยากฆ่าตัวตาย	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
72	อยากยิงให้ตาย	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
73	อยากให้ผู้กระทำตาย	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
74	เครียด	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
75	แค้น	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
76	ซึมเศร้า	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
77	อึดอัด	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
78	อยากเอามีดแทง	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
79	ไม่มีความสุข	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
80	ไม่อยากรักพ่อ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
81	ไม่อยากอยู่กับพ่อ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
82	แยกตัว	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
83	เกลียดผู้กระทำ	G207	ผลกระทบทางจิตใจหลังถูกทารุณ

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
84	กดดัน	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
85	กินได้น้อย	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
86	กลัว	G207	ผลกระทบทางจิตใจหลังถูกทารุณ
87	กินเหล้า	G208	ยาเสพติด เครื่องดื่มแอลกอฮอล์
88	สารเสพติด	G208	ยาเสพติด เครื่องดื่มแอลกอฮอล์
89	เหล้า	G208	ยาเสพติด เครื่องดื่มแอลกอฮอล์
90	ยาเสพติด	G208	ยาเสพติด เครื่องดื่มแอลกอฮอล์
91	ใช้ยา	G208	ยาเสพติด เครื่องดื่มแอลกอฮอล์
92	เรียกสีไม่ถูก	G209	ปัญหาด้านพัฒนาการของเด็ก
93	บวกเลขยังไม่ได้	G209	ปัญหาด้านพัฒนาการของเด็ก
94	ท่องกฮยังไม่ได้	G209	ปัญหาด้านพัฒนาการของเด็ก
95	ไม่ค่อยมองหน้า	G209	ปัญหาด้านพัฒนาการของเด็ก
96	ไม่ค่อยสบตา	G209	ปัญหาด้านพัฒนาการของเด็ก
97	ติดขวดนม	G209	ปัญหาด้านพัฒนาการของเด็ก
98	ไม่ค่อยมั่นใจในตัวเอง	G209	ปัญหาด้านพัฒนาการของเด็ก
99	ลำบากในการพูดโต้ตอบ	G209	ปัญหาด้านพัฒนาการของเด็ก
100	เขียนกขคไม่ได้	G209	ปัญหาด้านพัฒนาการของเด็ก
101	เขียนหนังสือไม่สวย	G209	ปัญหาด้านพัฒนาการของเด็ก
102	iqต่ำ	G209	ปัญหาด้านพัฒนาการของเด็ก
103	ดื้อ	G210	มีความซุกซน
104	ดื้อมาก	G210	มีความซุกซน

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
105	ซน	G210	มีความซุกซน
106	ก้าวร้าวมากขึ้นเมื่อไม่ได้ดังใจ	G210	มีความซุกซน
107	แกล้งน้อง	G210	มีความซุกซน
108	แกะพัดลม	G210	มีความซุกซน
109	ทำลายข้าวของ	G210	มีความซุกซน
110	เลี้ยงยาก	G210	มีความซุกซน
111	อยู่ไม่นิ่ง	G210	มีความซุกซน
112	อารมณ์แปรปรวน	G211	ภาวะทางอารมณ์ของเด็ก
113	หงุดหงิด	G211	ภาวะทางอารมณ์ของเด็ก
114	หมกมุ่นเรื่องเพศ	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
115	ติดเกม	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
116	ทำร้ายตัวเอง	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
117	แววตาเอาจริงเมื่อขัดใจ	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
118	อาละวาด	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
119	เอาแต่ใจ	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ

ตารางที่ ก.2 (ต่อ)

คุณลักษณะกลุ่มหลัก 2– ข้อมูลประวัติผู้ป่วย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
120	พฤติกรรมที่โตเกินวัย	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
121	เคลื่อนไหวช้ากว่าปกติ	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ
122	เฉื่อย	G212	พฤติกรรมที่บ่งบอกถึงความผิดปกติ

ตารางที่ ก.3 คุณลักษณะกลุ่มหลัก 3 – ข้อมูลผู้กระทำ

คุณลักษณะกลุ่มหลัก 3 – ข้อมูลผู้กระทำ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	มีอาการทางจิต	G301	โรคจิตเวช
2	อารมณ์ไม่คงที่	G302	ภาวะทางอารมณ์
3	อารมณ์ร้อน	G302	ภาวะทางอารมณ์
4	อารมณ์รุนแรง	G302	ภาวะทางอารมณ์
5	โมโหง่าย	G302	ภาวะทางอารมณ์
6	โมโหร้าย	G302	ภาวะทางอารมณ์
7	ไม่มีวุฒิภาวะทางอารมณ์	G302	ภาวะทางอารมณ์
8	ควบคุมอารมณ์ตนเองไม่ได้	G302	ภาวะทางอารมณ์
9	ควบคุมอารมณ์ลำบาก	G302	ภาวะทางอารมณ์
10	ภาวะอารมณ์รุนแรง	G302	ภาวะทางอารมณ์
11	ติดยา	G303	เกี่ยวข้องกับยาเสพติด

ตารางที่ ก.3 (ต่อ)

คุณลักษณะกลุ่มหลัก 3 – ข้อมูลผู้กระทำ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	มีอาการทางจิต	G301	โรคจิตเวช
12	ใช้ยา	G303	เกี่ยวข้องกับยาเสพติด
13	ใช้ยาเสพติด	G303	เกี่ยวข้องกับยาเสพติด
14	ค้ายา	G303	เกี่ยวข้องกับยาเสพติด
15	ยาเสพติด	G303	เกี่ยวข้องกับยาเสพติด
16	สารเสพติด	G303	เกี่ยวข้องกับยาเสพติด
17	สูบกัญชา	G303	เกี่ยวข้องกับยาเสพติด
18	ดมกาว	G303	เกี่ยวข้องกับยาเสพติด
19	เสพยา	G303	เกี่ยวข้องกับยาเสพติด
20	น้า	G304	ผู้กระทำเป็นคนในครอบครัว
21	ผู้ดูแลเด็ก	G304	ผู้กระทำเป็นคนในครอบครัว
22	ลุง	G304	ผู้กระทำเป็นคนในครอบครัว
23	แม่เลี้ยง	G304	ผู้กระทำเป็นคนในครอบครัว
24	ย่า	G304	ผู้กระทำเป็นคนในครอบครัว
25	แม่	G304	ผู้กระทำเป็นคนในครอบครัว
26	ตาเลี้ยง	G304	ผู้กระทำเป็นคนในครอบครัว
27	ปู่	G304	ผู้กระทำเป็นคนในครอบครัว
28	พ่อ	G304	ผู้กระทำเป็นคนในครอบครัว
29	พ่อเลี้ยง	G304	ผู้กระทำเป็นคนในครอบครัว
30	พี่ชาย	G304	ผู้กระทำเป็นคนในครอบครัว
31	ทะเลาะวิวาท	G305	ปัญหาความรุนแรงในครอบครัว

ตารางที่ ก.3 (ต่อ)

คุณลักษณะกลุ่มหลัก 3 – ข้อมูลผู้กระทำ			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
32	ทำร้ายร่างกาย	G305	ปัญหาความรุนแรงในครอบครัว
33	ทำอนาจาร	G305	ปัญหาความรุนแรงในครอบครัว
34	ด่าทอ	G305	ปัญหาความรุนแรงในครอบครัว
35	คุกคาม	G305	ปัญหาความรุนแรงในครอบครัว
36	เคยฆ่าคนตาย	G305	ปัญหาความรุนแรงในครอบครัว
37	พกปืน	G305	ปัญหาความรุนแรงในครอบครัว
38	ผู้ต้องหายังไม่ถูกจับ	G306	ความเสี่ยงที่จะถูกกระทำทารุณซ้ำ
39	ไม่อยากให้มีการดำเนินคดี	G306	ความเสี่ยงที่จะถูกกระทำทารุณซ้ำ
40	เคยติดคุก	G307	มีประวัติการทำผิด
41	ติดคุก	G307	ประวัติการทำผิด
42	ติดหวย	G308	การพนัน
43	เล่นไพ่	G308	การพนัน
44	เล่นไฮโล	G308	การพนัน
45	ดื่มเหล้า	G309	ดื่มเหล้า สูบบุหรี่
46	สูบบุหรี่	G309	ดื่มเหล้า สูบบุหรี่
47	อาการมึนเมา	G309	ดื่มเหล้า สูบบุหรี่

ตารางที่ ก.4 คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	พ่อเลี้ยง	G401	สถานภาพครอบครัว
2	แม่เลี้ยง	G401	สถานภาพครอบครัว
3	แยกทาง	G401	สถานภาพครอบครัว
4	เลิกกัน	G401	สถานภาพครอบครัว
5	แต่งงานใหม่	G401	สถานภาพครอบครัว
6	ปัญหาครอบครัว	G401	สถานภาพครอบครัว
7	พ่อแม่แยกกันอยู่	G401	สถานภาพครอบครัว
8	พ่อแม่เลิกกัน	G401	สถานภาพครอบครัว
9	พ่อแม่เสียชีวิต	G401	สถานภาพครอบครัว
10	มีครอบครัวใหม่	G401	สถานภาพครอบครัว
11	ใช้ความรุนแรง	G402	ปัญหาความรุนแรงในครอบครัว
12	พ่อแม่ทะเลาะกัน	G402	ปัญหาความรุนแรงในครอบครัว
13	พ่อตีแม่	G402	ปัญหาความรุนแรงในครอบครัว
14	เตะ	G402	ปัญหาความรุนแรงในครอบครัว
15	ถีบ	G402	ปัญหาความรุนแรงในครอบครัว
16	ถืออาวุธ	G402	ปัญหาความรุนแรงในครอบครัว
17	ถูกตบตี	G402	ปัญหาความรุนแรงในครอบครัว
18	ทะเลาะ	G402	ปัญหาความรุนแรงในครอบครัว
20	ทำลายข้าวของ	G402	ปัญหาความรุนแรงในครอบครัว
21	ตบตี	G402	ปัญหาความรุนแรงในครอบครัว
22	ตบหน้า	G402	ปัญหาความรุนแรงในครอบครัว

ตารางที่ ก.4 (ต่อ)

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
23	ต่อย	G402	ปัญหาความรุนแรงในครอบครัว
24	ขู่	G402	ปัญหาความรุนแรงในครอบครัว
25	ใช้มีด	G402	ปัญหาความรุนแรงในครอบครัว
26	ใช้ไม้ตีแม่	G402	ปัญหาความรุนแรงในครอบครัว
27	ด่า	G402	ปัญหาความรุนแรงในครอบครัว
29	ใช้ความรุนแรง	G402	ปัญหาความรุนแรงในครอบครัว
30	ใช้คีมหนีบนิ้วมือ	G402	ปัญหาความรุนแรงในครอบครัว
31	ระแวงเรื่องผู้หญิง	G402	ปัญหาความรุนแรงในครอบครัว
32	มีปากเสียง	G402	ปัญหาความรุนแรงในครอบครัว
33	แม่ถูกพ่อตี	G402	ปัญหาความรุนแรงในครอบครัว
34	เอามีดปังตอสับหลัง	G402	ปัญหาความรุนแรงในครอบครัว
35	กระทำชำเรา	G402	ปัญหาความรุนแรงในครอบครัว
40	ใช้ยา	G403	ปัญหายาเสพติด
41	ใช้ยาเสพติด	G403	ปัญหายาเสพติด
36	ทำร้ายร่างกาย	G402	ปัญหาความรุนแรงในครอบครัว
37	ตีรุนแรง	G402	ปัญหาความรุนแรงในครอบครัว
38	คุกคาม	G402	ปัญหาความรุนแรงในครอบครัว
39	บังคับ	G402	ปัญหาความรุนแรงในครอบครัว
40	ใช้ยา	G403	ปัญหายาเสพติด
41	ใช้ยาเสพติด	G403	ปัญหายาเสพติด
42	เฮโรอีน	G403	ปัญหายาเสพติด

ตารางที่ ก.4 (ต่อ)

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
43	คดียาเสพติด	G403	ปัญหายาเสพติด
44	ยาบ้า	G403	ปัญหายาเสพติด
45	ยาเสพติด	G403	ปัญหายาเสพติด
46	ดมกาว	G403	ปัญหายาเสพติด
47	คดียา	G403	ปัญหายาเสพติด
48	กัญชา	G403	ปัญหายาเสพติด
49	ติดยา	G403	ปัญหายาเสพติด
50	ติดหวย	G403	ปัญหายาเสพติด
51	หวยใต้ดิน	G404	การพนัน
52	เล่นการพนัน	G404	การพนัน
53	เล่นปอกเด้ง	G404	การพนัน
54	เล่นไพ่	G404	การพนัน
55	เล่นไฮโล	G404	การพนัน
56	เลี้ยงดูไม่เหมาะสม	G405	ปัญหาการเลี้ยงดูเด็ก
57	ไม่เคยชื่นชมเด็ก	G405	ปัญหาการเลี้ยงดูเด็ก
58	ไม่ได้ดูแลเด็ก	G405	ปัญหาการเลี้ยงดูเด็ก
59	ไม่มีแผนดูแล	G405	ปัญหาการเลี้ยงดูเด็ก
60	ไม่ได้เลี้ยงดู	G405	ปัญหาการเลี้ยงดูเด็ก
61	ไม่ได้เรียนหนังสือ	G405	ปัญหาการเลี้ยงดูเด็ก
62	ไม่พร้อมดูแล	G405	ปัญหาการเลี้ยงดูเด็ก
63	ไม่มีคนดูแลเด็ก	G405	ปัญหาการเลี้ยงดูเด็ก

ตารางที่ ก.4 (ต่อ)

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
64	ไม่มีความผูกพันธ์	G405	ปัญหาการเลี้ยงดูเด็ก
65	Interpersonal problem	G405	ปัญหาการเลี้ยงดูเด็ก
66	แม่ไม่เชื่อเด็ก	G405	ปัญหาการเลี้ยงดูเด็ก
67	แม่ไม่รักเด็ก	G405	ปัญหาการเลี้ยงดูเด็ก
68	ให้เด็กดูวีดีโอโป๊	G405	ปัญหาการเลี้ยงดูเด็ก
69	พาลูกไปดื่มเหล้า	G405	ปัญหาการเลี้ยงดูเด็ก
70	ยินยอมให้เด็กหมั้น	G405	ปัญหาการเลี้ยงดูเด็ก
71	บกพร่องทางสติปัญญา	G405	ปัญหาการเลี้ยงดูเด็ก
72	มองว่าเด็กเป็นภาระ	G405	ปัญหาการเลี้ยงดูเด็ก
73	ไม่ให้ไปโรงเรียน	G405	ปัญหาการเลี้ยงดูเด็ก
74	ดูแลไม่เหมาะสม	G405	ปัญหาการเลี้ยงดูเด็ก
75	สนับสนุนให้เด็กมีครอบครัว	G405	ปัญหาการเลี้ยงดูเด็ก
76	ใช้เงินเก่ง	G406	ปัญหาเศษรฐกิจ
77	มีปัญหาเรื่องการเงิน	G406	ปัญหาเศษรฐกิจ
78	เศรษฐกิจไม่ดี	G406	ปัญหาเศษรฐกิจ
79	เครียดเรื่องเงิน	G406	ปัญหาเศษรฐกิจ
80	สร้างหนี้	G406	ปัญหาเศษรฐกิจ
81	รายได้ไม่ดี	G406	ปัญหาเศษรฐกิจ
82	แม่เป็นแม่เล้า	G407	ประกอบอาชีพไม่เหมาะสม
83	ขายcdเถื่อน	G407	ประกอบอาชีพไม่เหมาะสม
84	ขายบริการ	G407	ประกอบอาชีพไม่เหมาะสม

ตารางที่ ก.4 (ต่อ)

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
85	ขายยาบ้า	G407	ประกอบอาชีพไม่เหมาะสม
86	กลัว	G408	ภาวะทางอารมณ์
87	ควบคุมอารมณ์ตนเองไม่ได้	G408	ภาวะทางอารมณ์
88	เครียด	G408	ภาวะทางอารมณ์
89	เครียดเรื่องงาน	G408	ภาวะทางอารมณ์
90	โกรธแรง	G408	ภาวะทางอารมณ์
91	โมโหง่าย	G408	ภาวะทางอารมณ์
92	หงุดหงิด	G408	ภาวะทางอารมณ์
93	อารมณ์ร้าย	G408	ภาวะทางอารมณ์
94	อารมณ์รุนแรง	G408	ภาวะทางอารมณ์
95	เอาแต่ใจตัวเอง	G408	ภาวะทางอารมณ์
96	เอาแน่นอนไม่ได้	G408	ภาวะทางอารมณ์
97	ภาวะอารมณ์รุนแรง	G408	ภาวะทางอารมณ์
98	ไม่มีวุฒิภาวะทางอารมณ์	G408	ภาวะทางอารมณ์
99	ติดสุรา	G409	ปัญหาสุรา
100	ติดเหล้า	G409	ปัญหาสุรา
101	เมาเหล้า	G409	ปัญหาสุรา
102	สูบบุหรี่	G409	ปัญหาสุรา
103	ดื่มเหล้า	G409	ปัญหาสุรา
104	พึ่งพาพ่อเลี้ยง	G410	ไม่สามารถปกป้อเด็กได้
105	แม่ไม่เชื่อเด็ก	G410	ไม่สามารถปกป้อเด็กได้

ตารางที่ ก.4 (ต่อ)

คุณลักษณะกลุ่มหลัก 4 – ข้อมูลสภาพครอบครัว ที่อยู่อาศัย และสภาพแวดล้อม			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
106	ไม่ได้ปกป้องเด็ก	G410	ไม่สามารถปกป้อเด็กได้
107	ไม่มีแนวทางป้องกัน	G410	ไม่สามารถปกป้อเด็กได้
108	เด็กอยู่บ้านผู้กระทำ	G410	ไม่สามารถปกป้อเด็กได้
109	ผู้กระทำเป็นผู้ดูแลเด็ก	G410	ไม่สามารถปกป้อเด็กได้
110	ไม่แจ้งความ	G410	ไม่สามารถปกป้อเด็กได้
111	ยังไม่มีการแจ้งความ	G410	ไม่สามารถปกป้อเด็กได้
112	ไม่สามารถปกป้องลูกได้	G410	ไม่สามารถปกป้อเด็กได้
113	นอนรวมกัน	G411	สภาพแวดล้อมในบ้านไม่เหมาะสม
114	เห็นพ่อแม่มีเพศสัมพันธ์	G411	สภาพแวดล้อมในบ้านไม่เหมาะสม
115	ดูสื่อลามกในบ้าน	G411	สภาพแวดล้อมในบ้านไม่เหมาะสม
116	hiv	G412	ปัญหาสุขภาพ
117	พิการ	G412	ปัญหาสุขภาพ
118	อาเจียนเป็นเลือด	G412	ปัญหาสุขภาพ
119	มีอาการทางจิต	G413	ภาวะทางจิต
120	แม่เด็กติดคุก	G414	มีประวัติการทำผิด
121	เรือนจำ	G414	มีประวัติการทำผิด
122	ถูกจับ	G414	มีประวัติการทำผิด
123	พ่อถูกจับ	G414	มีประวัติการทำผิด

ตารางที่ ก.5 คุณลักษณะกลุ่มหลัก 5 – บาดแผล ร่องรอยการถูกทำร้าย

คุณลักษณะกลุ่มหลัก 5 – บาดแผล ร่องรอยการถูกทำร้าย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	แผลเป็น	G501	อาการเจ็บ บาดแผลทั่วไป
2	แผลฟกช้ำ	G501	อาการเจ็บ บาดแผลทั่วไป
3	เป็นแผล	G501	อาการเจ็บ บาดแผลทั่วไป
4	แผลเก่า	G501	อาการเจ็บ บาดแผลทั่วไป
5	แผลถลอก	G501	อาการเจ็บ บาดแผลทั่วไป
6	ได้รับบาดเจ็บ	G501	อาการเจ็บ บาดแผลทั่วไป
7	เจ็บ	G501	อาการเจ็บ บาดแผลทั่วไป
8	รอยกัด	G501	อาการเจ็บ บาดแผลทั่วไป
9	รอยช้ำ	G501	อาการเจ็บ บาดแผลทั่วไป
10	รอยถลอก	G501	อาการเจ็บ บาดแผลทั่วไป
11	รอยแผล	G501	อาการเจ็บ บาดแผลทั่วไป
12	รอยหยิก	G501	อาการเจ็บ บาดแผลทั่วไป
13	แสบ	G501	อาการเจ็บ บาดแผลทั่วไป
14	สลบ	G502	อาการเจ็บ บาดแผลทั่วไป
15	หัวโน	G502	บาดแผล อาการบาดเจ็บรุนแรง
16	สำลัก	G502	บาดแผล อาการบาดเจ็บรุนแรง
17	เลือด	G502	บาดแผล อาการบาดเจ็บรุนแรง
18	เลือดออก	G502	บาดแผล อาการบาดเจ็บรุนแรง
19	ปากแตก	G502	บาดแผล อาการบาดเจ็บรุนแรง
20	แผลที่ศีรษ	G502	บาดแผล อาการบาดเจ็บรุนแรง
21	แผลที่หู	G502	บาดแผล อาการบาดเจ็บรุนแรง

ตารางที่ ก.5 (ต่อ)

คุณลักษณะกลุ่มหลัก 5 – บาดแผล ร่องรอยการถูกทำร้าย			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
22	แผลที่อวัยวะเพศ	G503	อาการบาดเจ็บเกี่ยวกับการถูกล่วงละเมิดทางเพศ
23	เจ็บอวัยวะเพศ	G503	อาการบาดเจ็บเกี่ยวกับการถูกล่วงละเมิดทางเพศ
24	เลือดออกอวัยวะเพศ	G503	อาการบาดเจ็บเกี่ยวกับการถูกล่วงละเมิดทางเพศ

ตารางที่ ก.6 คุณลักษณะกลุ่มหลัก 6 – มีการดำเนินการปกป้องเด็ก

คุณลักษณะกลุ่มหลัก 6 – มีการดำเนินการปกป้องเด็ก			
ลำดับ	**คำแสดงคุณลักษณะ**	**รหัสกลุ่มย่อย**	**คุณลักษณะกลุ่มย่อย**
1	วางแผนดูแลเด็ก	G601	การวางแผนดูแลเด็ก
2	ปรับปรุงการดูแลเด็ก	G601	การวางแผนดูแลเด็ก
3	เปลี่ยนผู้ดูแล	G601	การวางแผนดูแลเด็ก
4	วิเคราะห์จุดอ่อนของตัวเอง	G602	มีความเข้าใจในตัวเด็ก
5	พยายามทำความเข้าใจเด็ก	G602	มีความเข้าใจในตัวเด็ก
6	ถูกจับดำเนินคดีแล้ว	G603	มีการดำเนินคดีกับผู้กระทำ
7	คัดค้านการประกันตัว	G603	มีการดำเนินคดีกับผู้กระทำ
8	ผู้ต้องหาถูกจับตัวแล้ว	G603	มีการดำเนินคดีกับผู้กระทำ
9	สามารถดูแลลูกได้	G604	มีความสามารถที่จะปกป้องเด็กได้
10	ควบคุมพฤติกรรมได้	G604	มีความสามารถที่จะปกป้องเด็กได้
11	เด็กอยู่กับแม่ตลอด	G604	มีความสามารถที่จะปกป้องเด็กได้

ภาคผนวก ข
กฎที่ใช้ในการวิเคราะห์ประเภท
การให้ความช่วยเหลือเด็กที่ถูกทารุณ

ตารางที่ ข.1 กฎที่ใช้ในการจำแนกประเภท

ลำดับ	รหัสกฎ	คำอธิบาย
1	Lv101	เด็กถูกล่วงละเมิดทางเพศ
2	Lv102	เด็กถูกทำร้ายร่างกาย
3	Lv103	เด็กมีลักษณะถูกทอดทิ้ง
4	Lv104	มีความเกี่ยวข้องกับการค้ามนุษย์ หรือการใช้แรงงานเด็ก
5	Lv105	เด็กได้รับความบอบช้ำทางจิตใจ
6	Lv106	เด็กมีความกระทบกระเทือนทางจิตใจจากการถูกกระทำ
7	Lv107	ปัจจัยการทารุณเกิดจากตัวผู้กระทำ
8	Lv108	ปัจจัยการทารุณเกิดจากตัวเด็ก ด้วยสาเหตุปัญหาด้านสุขภาพ
9	Lv109	ปัจจัยการทารุณเกิดจากตัวเด็ก ด้วยสาเหตุปัญหาด้านพฤติกรรม
10	Lv110	การทารุณที่เกิดสภาพครอบครัวแตกแยก
11	Lv111	การทารุณที่เกิดจากความรุนแรงภายในครอบครัว
12	Lv112	พ่อแม่ ผู้ดูแลเด็กขาดความสามารถในการดูแลเด็ก กรณีผู้กระทำเป็นบุคคลภายนอก
13	Lv113	พ่อแม่ ผู้ดูแลเด็กขาดความสามารถในการดูแลเด็ก กรณีผู้กระทำเป็นบุคคลภายในครอบครัว
14	Lv114	ข้อมูลครอบครัวมีความเสี่ยงต่อการทารุณเด็ก
15	Lv115	การทำร้ายที่เกิดจากอุบัติเหตุ ความไม่ตั้งใจ และรู้เท่าไม่ถึงการณ์
16	Lv116	พ่อแม่ ผู้ดูแลเด็กมีแผนการในการดูแลเด็ก เพื่อป้องกันเหตุ
17	PA_Risk	ประเมินความเสี่ยงในการถูกระทำซ้ำ ด้านการถูกทำร้ายร่างกาย
18	SA_Risk	ประเมินความเสี่ยงในการถูกระทำซ้ำ ด้านการถูกล่งละเมิดทางเพศ
19	Fam_Risk	ประเมินความเสี่ยงในการถูกระทำซ้ำ จากครอบครัว

ตารางที่ ข.2 การสร้างกฎจากคุณลักษณะ

รหัสกฎ	คำอธิบาย	IF	THEN
Lv101	เด็กถูกล่วงละเมิดทางเพศ	(G101 = AND G102 = 1 AND G103 = 0) OR (G101 = 0 AND G102 = 0 AND G103 = 1) OR (G101 = 0 AND G102 = 1 AND G103 = 1) OR (G101 = 1 AND G102 = 1 AND G103 = 1) OR (G101 = 1 AND G102 = 0 AND G103 = 1)	1
Lv102	เด็กถูกทำร้ายร่างกาย	(G104 = 0 AND G105 = 1) OR (G104 = 1 AND G105 = 1)	1
Lv103	เด็กมีลักษณะถูกทอดทิ้ง	(G106 = 1)	1
Lv104	มีความเกี่ยวข้องกับการค้ามนุษย์ หรือการใช้แรงงานเด็ก	(G107 = 1)	1
Lv105	เด็กได้รับความบอบช้ำทางจิตใจ	(G108 = 1 AND G109 = 0) OR (G108 = 1 AND G109 = 1) OR (G108 = 0 AND G109 = 1)	1
Lv106	เด็กมีความกระทบกระเทือนทางจิตใจจากการถูกกระทำ	(G207 = 1)	1
Lv107	ปัจจัยการทารุณเกิดจากตัวผู้กระทำ	(G301 = 1) OR (G303 = 1) OR (G309 = 1) OR (G312 = 1) OR (G313 = 1)	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
Lv108	ปัจจัยการทารุณเกิดจากตัวเด็ก ด้วยสาเหตุปัญหาด้านสุขภาพ	(G205 = 1 AND G202 = 1 AND G210 = 1 AND G211 = 1 AND G212 = 1) OR (G205 = 1 AND G202 = 0 AND G210 = 1 AND G211 = 1 AND G212 = 1) OR (G205 = 1 AND G202 = 1 AND G210 = 1 AND G211 = 0 AND G212 = 1) OR (G205 = 1 AND G202 = 0 AND G210 = 1 AND G211 = 0 AND G212 = 1) OR (G204 = 1 AND G205 = 0 AND G209 = 0 AND G212 = 1) OR (G204 = 1 AND G205 = 1 AND G209 = 1 AND G212 = 1) OR (G204 = 1 AND G205 = 1 AND G209 = 0 AND G212 = 1) OR (G204 = 1 AND G206 = 1) OR (G206 = 1)	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
Lv109	ปัจจัยการทารุณเกิดจากตัวเด็ก ด้วยสาเหตุปัญหาด้านพฤติกรรม	(G204 = 1 AND G205 = 1 AND G209 = 0 AND G210 = 1 AND G212 = 0) OR (G208 = 1 AND G205 = 0 AND G209 = 0 AND G210 = 0 AND G212 = 0) OR (G208 = 1 AND G205 = 1 AND G209 = 0 AND G210 = 0 AND G212 = 0) OR (G208 = 1 AND G205 = 1 AND G209 = 0 AND G210 = 1 AND G212 = 0)	1
Lv110	การทารุณที่เกิดสภาพครอบครัวแตกแยก	(G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G401 = 1 AND G402 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G401 = 1 AND G405 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G401 = 1 AND G410 = 1)	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
Lv111	การทารุณที่เกิดจากความรุนแรงภายในครอบครัว	(G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G402 = 1 AND G403 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G402 = 1 AND G409 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G402 = 1 AND G414 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G402 = 1 AND G404 = 1 AND G406 = 1) OR (G601 = 0 AND G602 = 0 AND G603 = 0 AND G604 = 0 AND G402 = 1 AND G408 = 1 AND G406 = 1)	1
Lv112	พ่อแม่ ผู้ดูแลเด็กขาดความสามารถในการดูแลเด็ก กรณีผู้กระทำเป็นบุคคลภายนอก	(G304 = 0 AND G402 = 1) OR (G304 = 0 AND G403 = 1) OR (G304 = 0 AND G405 = 1) OR (G304 = 0 AND G409 = 1) OR (G304 = 0 AND G410 = 1) OR (G304 = 0 AND G411 = 1) OR (G304 = 0 AND G107 = 1 AND G407 = 1)	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
Lv113	พ่อแม่ ผู้ดูแลเด็กขาดความสามารถในการดูแลเด็ก กรณีผู้กระทำเป็นบุคคลภายในครอบครัว	(G304 = 1 AND G403 = 1) OR (G304 = 1 AND G404 = 1) OR (G304 = 1 AND G405 = 1) OR (G304 = 1 AND G408 = 1) OR (G304 = 1 AND G409 = 1) OR (G304 = 1 AND G410 = 1) OR (G304 = 1 AND G412 = 1) OR (G304 = 1 AND G413 = 1)	1
Lv114	ข้อมูลครอบครัวมีความเสี่ยงต่อการทารุณเด็ก	G403 = 1 OR G413 = 1 OR G410 = 1 OR G405 = 1 OR G401 = 1 AND G414 = 1 OR G401 = 1 AND G402 = 1 AND G409 = 1 OR G401 = 1 AND G402 = 1 AND G404 = 1 OR G406 = 1 AND G106 = 1 OR G406 = 1 AND G107 = 1 OR G407 = 1 AND G107 = 1 OR G409 = 1 AND G402 = 1 OR G412 = 1 AND G408 = 1 OR G412 = 1 AND G403 = 1 OR G412 = 1 AND G413 = 1 OR G412 = 1 AND G410 = 1 OR G412 = 1 AND G405 = 1	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
Lv115	การทำร้ายที่เกิดจากอุบัติเหตุ ความไม่ตั้งใจและรู้เท่าไม่ถึงการณ์	G101 = 1 AND G501 = 0 AND G502 = 0 AND G503 = 0 OR G101 = 1 AND G501 = 1 AND G502 = 0 AND G503 = 0 OR G102 = 1 AND G501 = 0 AND G502 = 0 AND G503 = 0 OR G102 = 1 AND G501 = 1 AND G502 = 0 AND G503 = 0 OR G103 = 1 AND G501 = 0 AND G502 = 0 AND G503 = 0 OR G103 = 1 AND G501 = 1 AND G502 = 0 AND G503 = 0 OR G104 = 1 AND G501 = 0 AND G502 = 0 AND G503 = 0 OR G104 = 1 AND G501 = 1 AND G502 = 0 AND G503 = 0 OR G105 = 1 AND G501 = 0 AND G502 = 0 AND G503 = 0 OR G105 = 1 AND G501 = 1 AND G502 = 0 AND G503 = 0	1
Lv116	พ่อแม่ ผู้ดูแลเด็กมีแผนการในการดูแลเด็ก เพื่อป้องกันเหตุ	(G601= 1 OR G602 = 1 OR G603 = 1) OR (G604 = 1)	1

ตารางที่ ข.2 (ต่อ)

รหัสกฎ	คำอธิบาย	IF	THEN
PA	ประเมินความเสี่ยงในการถูกระทำซ้ำ ด้านการถูกทำร้ายร่างกาย	Lv1 = 113 and Lv 1 = 114	1
SA	ประเมินความเสี่ยงในการถูกระทำซ้ำ ด้านการถูกล่วงละเมิดทางเพศ	Lv1 = 101 and Lv0 = 116	1
Fam	ประเมินความเสี่ยงในการถูกระทำซ้ำ จากครอบครัว	(PA = 1 and Lv114 = 1) OR (SA = 1 and Lv114 = 1)	1

ภาคผนวก ค

บทความประชุมวิชาการ

การประชุมวิชาการมหาวิทยาลัยเทคโนโลยีราชมงคล ครั้งที่ 9

เรื่อง"ราชมงคลสร้างสรรค์กับนวัตกรรมที่ยั่งยืนสู่ประเทศไทย 4.0"

ศูนย์แสดงสินค้า และการประชุม อิมแพ็คเมืองทองธานี

การจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมโดยเทคนิคการทำเหมืองข้อความ
Classification of Helping Abused Children by using Text Mining Techniques

ปรางค์ปราณี ล้ำเลิศพระคุณ* **และ อรวรรณ อิ่มสมบัติ**
สาขาวิศวกรรมเว็บ วิทยาลัยครีเอทีฟดีไซน์ แอนด์ เอนเตอร์เทนเมนต์เทคโนโลยี
มหาวิทยาลัยธุรกิจบัณฑิตย์ จังหวัดนนทบุรี

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาเทคนิคการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมด้วยเทคนิคการทำเหมืองข้อความ โดยการวิเคราะห์ข้อมูลบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็ก ซึ่งอยู่ในรูปของการบันทึกข้อความ ประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมสามารถแบ่งได้ 2 ประเภทใหญ่ๆ คือ การสงเคราะห์เด็ก กับการคุ้มครองสวัสดิภาพเด็ก โดยเอกสารบันทึกเหตุการณ์จะถูกนำมาผ่านการประมวลผลภาษาธรรมชาติก่อนเพื่อสกัดคุณลักษณะซึ่งเป็นคำที่เกี่ยวกับการทำร้ายร่างกาย การล่วงละเมิดทางเพศ การทำร้ายจิตใจ อาวุธบาดแผลและพฤติกรรมที่เกี่ยวข้อง เทคนิคในการจำแนกประเภทที่นำมาทดสอบ มี 4 เทคนิค คือ Decision Tree, Neural Network, Naïve Bayes และ Support Vector Machine จากผลการประเมินพบว่า Naïve Bayes มีประสิทธิภาพในการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณมากที่สุด โดยให้ค่า Precision 88.53 % ค่า Recall83.02 % และ F-score 85.68 %

คำสำคัญ : การให้ความช่วยเหลือเด็กที่ถูกทารุณกรรม, เทคนิคการจำแนกประเภท, เหมืองข้อความ

Abstract

The purpose of this research was to study the classification techniques of helping abused children by using text mining techniques. The data of this research were the records of children abused cases which were in memorandum format. The helping abused children was classified into 2 types: child assistance and child welfare protection. Natural Language Process (NLP) was used for features extraction from case records. The features are the words concerning to physical abuse, sexual abuse, emotional abuse, weapon, wound and related behavior. Four classification algorithms used in this study were Decision Tree, Neural Network, Naïve Bayes and Support Vector Machine. Naïve Bayes showed the best performance with achieved 88.53% precision 83.02% recall and 85.68% F-score.

Keywords : Helping Abused Children, Classification Techniques, Text Mining

* ผู้นิพนธ์ประสานงาน 575159090025@dpu.ac.th โทร. 08 3669 9819

1. บทนำ

การทารุณกรรมเด็กเป็นปัญหาที่เกิดขึ้นทั่วโลกรวมถึงในประเทศไทยเด็กที่ถูกทารุณจะได้รับผลกระทบในทุกด้าน ทั้งทางร่างกาย จิตใจ สังคม และยังส่งผลต่อพัฒนาการในการเติบโตของเด็ก โดยได้แบ่งรูปแบบของการทารุณกรรมเป็น 4 ประเภท ได้แก่การทำร้ายร่างกาย ความรุนแรงทางเพศ การทารุณจิตใจ และการทอดทิ้งปล่อยปละละเลย งานวิจัยนี้ได้ศึกษาการทารุณกรรมประเภทการทำร้ายร่างกาย กับความรุนแรงทางเพศ เนื่องจากเป็นรูปแบบการทารุณกรรมที่เกิดขึ้นมากที่สุด (นันทนา สวัสดิพานิช, 2555 : 25) การให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมแบ่งเป็น 2 ประเภทคือ การสงเคราะห์เด็ก หมายถึงเด็กพิการ เด็กที่อยู่ในสภาพยากลำบาก และเด็กที่อยู่ในสภาพที่จำเป็นต้องได้รับการสงเคราะห์ การส่งเด็กเข้ารับการอุปการะในสถานแรกรับ ครอบครัวอุปถัมภ์ สถานรับเลี้ยง

เด็ก สถานสงเคราะห์ สถานพัฒนาและฟื้นฟูหรือสถานที่ทางศาสนา ต้องได้รับความยินยอมจากผู้ปกครอง กับการคุ้มครอง สวัสดิภาพเด็ก หมายถึงเด็กที่ถูกทารุณกรรม เสี่ยงต่อการกระทำผิด และเด็กที่อยู่ในสภาพที่จำต้องได้รับการคุ้มครองสวัสดิภาพ โดยมีการให้อำนาจเจ้าหน้าที่เข้าตรวจค้นและแยกตัวเด็กออกจากครอบครัว (พรบ. คุ้มครองเด็ก พ.ศ. 2546, 2546) โดยพิจารณาจากบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณกรรม ซึ่งเป็นบันทึกข้อมูลความรุนแรงที่เกิดกับเด็ก ความเสี่ยง ลักษณะของเด็ก อารมณ์ สภาพจิตของผู้ก่อเหตุ บาดแผล ความสัมพันธ์ของเด็กกับผู้กระทำทารุณ สภาพครอบครัว ฐานะความเป็นอยู่ และสภาพแวดล้อม ข้อมูลเหล่านี้อยู่ในรูปของบันทึกข้อความซึ่งบันทึกโดยแพทย์ นักสังคมสงเคราะห์ และนักจิตวิทยา โดยบันทึกข้อความเหล่านี้มีส่วนในการพิจารณาแนวทางการให้ความช่วยเหลือ และดำเนินคดีกับผู้กระทำผิด ในการดำเนินการจะอาศัยความคิดเห็นของบุคลากรที่เกี่ยวข้องจำนวนมาก บางครั้งมีความเห็นที่แตกต่างกัน และอาจทำให้การตัดสินใจเกิดความล่าช้า โดยเฉพาะกรณีที่ต้องเร่งให้ความช่วยเหลือ รวมถึงการทำงานในปัจจุบันยังไม่มีเครื่องมือที่จะช่วยในการตัดสินใจ และการตัดสินใจให้ความช่วยเหลือเป็นดุลพินิจส่วนบุคคล

เนื่องจากบันทึกข้อความที่ใช้ในงานวิจัยนี้อยู่ในรูปแบบของภาษาธรรมชาติ ดังนั้นจึงมีการประยุกต์ใช้เทคนิคเหมืองข้อความ โดยเหมืองข้อความเป็นการค้นหาข้อมูลสารสนเทศที่มีประโยชน์จากภาษาธรรมชาติ เป็นการสกัดความรู้ เอาข้อมูลที่ไม่เคยรู้มาก่อนสามารถนำไปใช้ประโยชน์ได้ตามวัตถุประสงค์จำเพาะเจาะจง ข้อมูลสารสนเทศที่นำมาทำเหมืองข้อความเป็นข้อมูลที่ไม่มีโครงสร้างหรือกึ่งโครงสร้าง ซึ่งเป็นข้อมูลที่มีหลากหลายมิติ มีความซับซ้อนและไม่ชัดเจน เช่น บันทึกเอกสาร การทำเหมืองข้อความต้องมีการประมวลผลภาษาทางธรรมชาติ คือการวิเคราะห์ภาษาของมนุษย์ เป็นกระบวนการที่ทำให้คอมพิวเตอร์เข้าใจภาษาธรรมชาติของมนุษย์ งานวิจัยนี้นำเสนอการประยุกต์ใช้เทคนิคการทำเหมืองข้อความในการวิเคราะห์เอกสารที่เกี่ยวข้องกับการทารุณกรรมเด็กที่อยู่ในรูปแบบบันทึกข้อความ และจำแนกประเภทการให้ความช่วยเหลือ ซึ่งจะเกิดประโยชน์ต่อการดำเนินการตัดสินใจให้ความช่วยเหลือเพื่อให้ตัดสินใจได้รวดเร็ว และมีประสิทธิภาพมากขึ้น โดยในการทดลองได้ใช้เทคนิคการจำแนก 4 เทคนิค ได้แก่ Decision Tree, Neural Network, Naïve Bayes และ Support Vector Machine จากผลการประเมินพบว่า Naïve Bayes มีประสิทธิภาพในการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมมากที่สุด

2. วัตถุประสงค์การวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อวิเคราะห์เอกสารบันทึกข้อความที่เกี่ยวข้องกับการทารุณกรรมเด็ก และศึกษาเทคนิคการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณ

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

Popowich [6] ได้นำเสนอการใช้เทคนิคเหมืองข้อความและการประมวลผลภาษาธรรมชาติในกระบวนการพิจารณาการขอเคลมประกันสุขภาพ โดยประยุกต์ใช้เทคนิคการประมวลผลภาษาธรรมชาติมาช่วยวิเคราะห์ข้อมูลในการตัดสินใจในการพิจารณาการจ่ายค่าสินไหมทดแทน งานวิจัยนี้ได้ใช้หลักการของ Concept Taxonomies ในการวิเคราะห์ประโยค ซึ่งเป็นบันทึกผลการตรวจบันทึกข้อความจากการรับแจ้งโดย Call center และข้อมูลอื่นๆที่เกี่ยวข้อง ผลการวิจัยพบว่าระบบที่พัฒนาประสบความสำเร็จเป็นอย่างมาก เนื่องจากเป็นข้อมูลเกี่ยวกับการแพทย์ โรคติดต่อ และการรักษาเท่านั้น ซึ่งเป็นข้อมูลเฉพาะทางในขอบเขตของงาน ทำให้สามารถสร้างแบบจำลองที่สมบูรณ์ได้ นอกจากนี้ยังสามารถนำไปประยุกต์ใช้กับระบบต่าง ๆ ที่มีการวิเคราะห์ทางภาษาศาสตร์และงานที่เกี่ยวข้องกับความรู้เชิงความหมายได้

He, et al. [7] ได้นำเสนองานวิจัยในการประเมินโรคเครียดจากเหตุการณ์สะเทือนขวัญ (Posttraumatic Stress Disorder : PTSD) จากคำบอกเล่าของผู้ป่วยด้วยการประมวลผลภาษาทางธรรมชาติและเหมืองข้อความ โดยการวิเคราะห์ข้อความจากข้อมูลที่บันทึกด้วยตัวของผู้ป่วยเอง (Self-narratives) โดยให้ผู้ป่วยเข้าไปบันทึกข้อมูลประสบการณ์ร้ายแรงที่เกิดขึ้นกับตนเองและผลกระทบที่เกิดขึ้นในชีวิตประจำวันผ่านทางเว็บไซต์ แบ่งเป็นกลุ่มที่มีแนวโน้มของอาการ PTSD และไม่มีอาการ PTSD (Non-PTSD) เพื่อใช้ในการเรียนรู้ ในส่วนของการทดสอบได้ทดลองใช้อัลกอริทึม 4 อัลกอริทึมในการสร้างตัวแบบได้แก่ Decision tree, Naïve

Bayes, Support Vector Machine และ Product Score Model และประเมินความแม่นยำ (Accuracy) ของตัวแบบด้วย 10-fold cross-validation ผลการทดสอบพบว่า เทคนิค Product Score Model ให้ความแม่นยำสูงสุด โดยมีค่าเฉลี่ยอยู่ที่ .82 (SD =.05)

Yetisgen-Yildiz, et al. [8] ได้นำเสนองานวิจัยในการระบุอาการได้รับบาดเจ็บในถุงลมปอดของผู้ป่วยจากบันทึกรายงานผลการตรวจเอ็กซเรย์ทรวงอก เป็นการใช้เทคนิคการประมวลผลภาษาทางธรรมชาติ กับเทคนิคการจำแนกประเภทด้วยการเรียนรู้แบบมีผู้สอน (Supervised text classification) ในการวิเคราะห์ข้อมูลการตรวจเอ็กซเรย์ทรวงอกของผู้ป่วยว่ามีผลเป็นบวก (Positive for Acute Lung Injury) หรือไม่ ผลการทดสอบพบว่าค่าf-score ที่ดีที่สุดคือ 78.53%

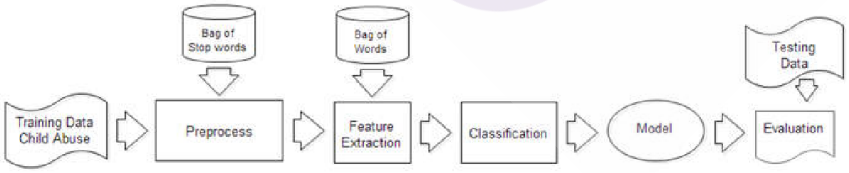
Botsis, et al. [9] ได้นำเสนองานวิจัยในการสร้างระบบอัตโนมัติเพื่อจำแนกอาการข้างเคียงที่เกิดจากการฉีดวัคซีนโดยใช้รายงานที่บันทึกในระบบรายงานอาการหลังได้รับการเสริมสร้างภูมิคุ้มกัน (Vaccine Adverse Event Reporting System : VAERS) โดยใช้เทคนิคการประมวลผลภาษาทางธรรมชาติ ในการวิเคราะห์อาการแพ้ว่าเป็นอาการแพ้ที่เกิดจากการได้รับวัคซีนหรือไม่ จากผลการทดสอบพบว่ามีความเป็นไปได้มากที่จะนำเทคนิคของเหมืองข้อความเข้ามาใช้ในการช่วยจำแนกอาการหลังได้รับการฉีดวัคซีนและมีโอกาสที่จะพัฒนาระบบเพิ่มเติมในงานอื่น ๆที่มีลักษณะเดียวกันได้

Shekhar , et al. [10] ได้นำเสนองานวิจัยในการจำแนกประเภทของศัพท์แพทย์ที่บันทึกอยู่ในเวชระเบียนอิเล็กทรอนิกส์เพื่อระบุกลุ่มของโรค ผู้วิจัยได้ศึกษาและจำแนกโรคเฉพาะทางทั้งหมด 5 โรค ได้แก่ โรคมะเร็ง, โรคเอดส์, โรคข้ออักเสบ, โรคเบาหวาน และโรคที่เกี่ยวกับหัวใจ ในส่วนของการเตรียมข้อมูลการประมวลผล (Preprocessing) ใช้เทคนิคของการประมวลผลภาษาทางธรรมชาติโดยทำการสกัดคำ และเลือกเฉพาะคำศัพท์แพทย์ โดยใช้ Support Vector Machines (SVM) ในการจำแนกประเภท และมีค่าเฉลี่ยของ F-score จากการจำแนกประเภทโรค 5 ประเภท เป็น 96.6%

Baker, et al. [11] ได้นำเสนองานวิจัยในการจำแนกบทความทางวิทยาศาสตร์ที่เกี่ยวกับการแพร่เชื้อของโรคมะเร็ง โดยจำแนกจากคำที่แสดงถึงความหมายของโรคมะเร็งที่แบ่งตามการแพร่เชื้อของมะเร็ง 10 ขั้นตอน ในการจำแนกประเภทได้เลือกใช้โมเดล Support vector machine (SVM) จากผลการวิจัยพบว่าเทคนิคเหมืองข้อความสามารถนำไปใช้ในการจำแนกความหมายในบทความทางวิทยาศาสตร์ได้

4. การดำเนินการวิจัย

งานวิจัยนี้เป็นการศึกษารวบรวมข้อมูลเพื่อสร้างตัวแบบสำหรับการทำเหมืองข้อความ โดยเริ่มจากการศึกษาความรู้ที่เกี่ยวข้องกับเด็กที่ถูกทารุณกรรม วิธีการให้ความช่วยเหลือ เก็บรวบรวมเอกสารบันทึกข้อความที่ได้รับแจ้งเหตุทารุณกรรมเด็ก การสัมภาษณ์เจ้าหน้าที่ การร่วมสังเกตุการณ์การทำงานของเจ้าหน้าที่และทีมสหวิชาชีพในการช่วยเหลือเด็กที่ถูกทารุณ กลุ่มตัวอย่างที่นำมาศึกษาคือผู้ป่วยเด็กที่ถูกทารุณ อายุไม่เกิน 18 ปีที่อยู่ในความดูแลของนักสังคมสงเคราะห์ของโรงพยาบาลแห่งหนึ่ง โดยข้อมูลที่ได้มาจะไม่มีการเปิดเผยชื่อของเด็กและบุคคลที่เกี่ยวข้อง บันทึกเอกสารที่เกี่ยวข้องกับการทารุณเด็กจะถูกนำเข้าสู่การประมวลผลภาษาทางธรรมชาติ สกัดคุณลักษณะที่เกี่ยวข้องกับการจำแนก สร้างตัวแบบการจำแนก และประเมินประสิทธิภาพของตัวแบบ ดังแสดงในรูปที่ 1



รูปที่ 1 สถาปัตยกรรมการทำงานของระบบ

ขั้นตอน Preprocess เป็นขั้นตอนการนำเอกสารบันทึกข้อความมาผ่านขั้นตอนการประมวลผลภาษาทางธรรมชาติ ซึ่งประกอบด้วยการตัดคำ (Word Segmentation) และการกำกับชนิดของคำ (POS tagging) โดยการตัดคำเป็นการแบ่งคำแต่ละคำในประโยคออกจากกันเพื่อให้สามารถประมวลผลในขั้นตอนถัดไปได้ ซึ่งงานวิจัยนี้ใช้ Lucene เป็นเครื่องมือในการตัดคำ ในส่วนของการกำกับชนิดคำ (POS tagging) ในงานวิจัยนี้ใช้คลังชนิดคำของ Orchid Corpus ในการเรียนรู้

การสกัดคุณลักษณะ (Feature extraction) เป็นการค้นหาคำหรือกลุ่มคำที่แสดงคุณลักษณะต่างๆ ที่เกี่ยวข้อง กับการทำทารุณกรรมเด็ก ซึ่งกำหนดโดยผู้เชี่ยวชาญ ซึ่งจะเป็นการนับจำนวนครั้งในการปรากฏของกลุ่มคำเหล่านี้ นอกจากนี้จะมีคุณลักษณะเกี่ยวกับการปฏิเสธการเกิดเหตุการณ์เหล่านี้โดยหากหน้ากลุ่มคำเหล่านี้ปรากฏคำที่เป็นปฏิเสธนำหน้า เช่น ไม่พบ, ไม่เคย จะเป็นการเพิ่มค่าคะแนนในกลุ่มปฏิเสธของกลุ่มลักษณะเหล่านี้ คุณลักษณะที่ใช้ในการจำแนกประเภทมีทั้งหมด 248 คุณลักษณะ แบ่งเป็นกลุ่มใหญ่ๆ ได้ 16 กลุ่ม ดังนี้

กลุ่มการทำร้ายร่างกาย สามารถแบ่งระดับการทำร้ายได้ 3 ระดับตามความรุนแรง ประกอบด้วย 51 กลุ่มคำ เช่น ตีศีรษะ, ตีหลัง, จับกดน้ำ

กลุ่มอาวุธที่ใช้ในการทำร้าย สามารถแบ่งได้ 3 ระดับตามความรุนแรง ประกอบด้วย 8 กลุ่มคำ เช่น ไม้, ไม้บรรทัด, ท่อแป๊บ

กลุ่มการล่วงละเมิดทางเพศ สามารถแบ่งได้ 3 ระดับตามความรุนแรง ประกอบด้วย 46 กลุ่มคำ เช่น จับหน้าอก, สอดนิ้วเข้าทางช่องคลอด, สอดใส่อวัยวะเพศเข้าทางช่องคลอด

กลุ่มบาดแผล สามารถแบ่งได้ 2 ระดับตามความรุนแรง ประกอบด้วย 26 กลุ่มคำ เช่น แผลถลอก, รอยช้ำ, เลือดออก

กลุ่มความเสี่ยงที่อาจถูกกระทำทารุณ ประกอบด้วย 7 กลุ่มคำ เช่น โตเกินวัย

กลุ่มปัญหาครอบครัว ประกอบด้วย 16 กลุ่มคำ เช่น พ่อแม่ทะเลาะกัน, พ่อทำร้ายแม่, พ่อแม่แยกทาง

กลุ่มอาการทางจิตของผู้กระทำทารุณ ประกอบด้วย 5 กลุ่มคำ เช่น ใช้ยา

กลุ่มอารมณ์ของผู้กระทำขณะก่อเหตุ ประกอบด้วย 6 กลุ่มคำ เช่น โกรธ, เครียด

กลุ่มปัญหายาเสพติด สามารถแบ่งได้ 3 ระดับตามความรุนแรง ประกอบด้วย 17 กลุ่มคำ เช่น เสพเฮโรอีน, มีประวัติการใช้ยาเสพติด, ชักชวนเสพยา

กลุ่มการเล่นการพนัน ประกอบด้วย 3 กลุ่มคำ ได้แก่ ปอกเต้ง, หวยใต้ดิน, เล่นการพนัน

กลุ่มการข่มขู่ ประกอบด้วย 5 กลุ่มคำ เช่น ขู่, อย่าบอกใคร, ไม่ให้บอกครู

กลุ่มการทำร้ายทางอารมณ์ ประกอบด้วย 8 กลุ่มคำ เช่น บังคับ, แม่กดดัน

กลุ่มการทอดทิ้ง ปล่อยปละละเลย ประกอบด้วย 13 กลุ่มคำ เช่น เร่ร่อน, แม่ไม่ได้เลี้ยงดูลูก, ไม่ให้เด็กไปโรงเรียน

กลุ่มปัญหาพฤติกรรม ปัญหาการเรียน ประกอบด้วย 18 กลุ่มคำ เช่น หนีเรียน

กลุ่มความผิดปกติของเด็ก ประกอบด้วย 2 กลุ่มคำ ดังนี้ difficult child, สมาธิสั้น

กลุ่มสภาพจิตของเด็กหลังเกิดเหตุ สามารถแบ่งได้ 2 ระดับตามความรุนแรง ประกอบด้วย 17 กลุ่มคำ เช่น ร่าเริง, ซึมเศร้า, แยกตัว

การจำแนกประเภท (Classification) สามารถแบ่งการให้ความช่วยเหลือเด็กที่ถูกทารุณออกเป็น 2 ประเภท คือ การสงเคราะห์เด็ก กับ การคุ้มครองสวัสดิภาพเด็ก และในการสร้างตัวแบบการจำแนกใช้โปรแกรม RapidMiner Studio เพื่อทดสอบเทคนิคการจำแนกทั้งหมด 4 เทคนิค คือ Decision Tree, Neural Network, Naïve Bayes และ Support Vector Machine

การทดสอบประสิทธิภาพของตัวแบบ (Evaluation) ทำการประเมินโดยใช้ 10-fold-cross-validation คือการแบ่งข้อมูลออกเป็น 10 ชุดเท่าๆ กันโดยใช้ข้อมูล 9 ชุดในการเรียนรู้และใช้ข้อมูล 1 ชุด เป็นตัวทดสอบประสิทธิภาพของโมเดล และทำซ้ำไปจนครบจำนวน 10 รอบ

5. ผลการวิจัย

ในจำแนกประเภทการให้ความช่วยเหลือได้มีการกำกับประเภทการให้ความช่วยเหลือเป็น 2 ประเภท คือ Class A : การสงเคราะห์เด็ก จำนวน 9 ตัวอย่าง และ Class B : การคุ้มครองสวัสดิภาพเด็ก จำนวน 21 ตัวอย่าง โดยจำนวนคำในตัวอย่างบันทึกเหตุการณ์ที่เกี่ยวข้องกับการทารุณเด็กเฉลี่ยประมาณ 155 คำ ต่อ 1 ตัวอย่าง

การวัดประสิทธิภาพของตัวแบบ ได้ใช้ค่าความแม่นยำ Precision, ค่าระลึก Recall และค่า F-score ในการวัดประสิทธิภาพโดยมีรายละเอียดดังนี้

$$\text{ความแม่นยำ } (Precision) = \frac{\text{จำนวนตัวอย่างที่โมเดลทำนายถูกต้อง}}{\text{จำนวนตัวอย่างที่โมเดลทำนายทั้งหมด}} \times 100$$

$$\text{ค่าระลึก } (Recall) = \frac{\text{จำนวนตัวอย่างที่โมเดลทำนายถูกต้อง}}{\text{จำนวนตัวอย่างที่มีทั้งหมด}} \times 100$$

$$F - score = \frac{2 \text{ x Preciosion x Recall}}{\text{Precision} + \text{Recall}}$$

จากผลการทดสอบ แสดงดังในตารางที่ 1 พบว่า Naïve Bayes ให้ค่า Precision, Recall และ F-score ดีที่สุด คือ 88.53%, 83.02%, 85.68% ตามลำดับ Decision Tree มีค่าต่ำที่สุด โดยมีค่า Precision, Recall และ F-score เป็น 49.60%, 41.27%, 45.06% ตามลำดับ

ตารางที่ 1 ผลการทดสอบประสิทธิภาพของตัวแบบ

Model	Precision			Recall			F-score
	Class A	Class B	Total	Class A	Class B	Total	
Decision Tree	72.00%	40.00%	49.60%	85.71%	22.22%	41.27%	45.06%
Neural Network	86.96%	85.71%	86.08%	95.24%	66.67%	75.24%	80.03%
Naïve Bayes	**90.91%**	**87.50%**	**88.53%**	**95.24%**	**77.78%**	**83.02%**	**85.68%**
SVM	82.61%	71.43%	74.78%	90.48%	55.56%	70.12%	70.12%

การจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณ ประเภทที่ต้องให้ความช่วยเหลืออย่างเร่งด่วนคือการคุ้มครองสวัสดิภาพเด็ก (Class B) จากตัวอย่างที่นำมาทดสอบ มีจำนวน 21 ตัวอย่างที่ถูกจัดประเภทให้เป็นการคุ้มครองสวัสดิภาพ ผลการทดสอบโดยการใช้เทคนิค Naïve Bayes พบว่า มีการจำแนกประเภทได้ถูกต้อง 20 ตัวอย่าง โดย 1 ตัวอย่างที่ถูกจำแนกผิด เป็นบันทึกข้อความเกี่ยวกับการถูกล่วงละเมิดทางเพศ เมื่อตรวจสอบเอกสารและการสกัดคุณลักษณะแล้วพบว่ามีความไม่ชัดเจนในการถูกกระทำตัวอย่างจากส่วนหนึ่งของบันทึกข้อมูลการถูกล่วงละเมิดทางเพศมีการบันทึกว่า “เด็กถูกข่มขืนโดยคนข้างบ้าน” ซึ่งสามารถจำแนกเป็นคุณลักษณะในกลุ่มของการถูกล่วงละเมิดทางเพศในระดับความรุนแรงที่ 3 เป็น 1 และในบันทึกยังพบข้อความ “เด็กมีแฟนแต่ไม่เคยมีเพศสัมพันธ์กับแฟน” ซึ่งประกอบด้วยคำปฏิเสธ เช่น “ไม่”, “ไม่มี”, “ไม่เคย” จึงมีค่าคุณลักษณะในกลุ่มคำปฏิเสธการถูกล่วงละเมิด

ทางเพศในระดับความรุนแรงที่ 3 เป็น 1 เช่นกัน ทำให้คะแนนของการถูกกระทำและไม่ถูกกระทำมีค่าเท่ากัน จึงเกิดความไม่ชัดเจนในส่วนของการถูกล่วงละเมิดทางเพศและทำให้ระบบจำแนกประเภทผิด

6. บทสรุปและข้อเสนอแนะ

จากการศึกษาบันทึกเอกสารที่เกี่ยวข้องกับการทารุณเด็กและการจำแนกประเภทการให้ความช่วยเหลือเด็กที่ถูกทารุณกรรมด้วยการประยุกต์ใช้เทคนิคการทำเหมืองข้อความ การจำแนกประเภทที่นำมาทดสอบมี 4 เทคนิค คือ Decision Tree, Neural Network, Naïve Bayes และ Support Vector Machine แบ่งประเภทการให้ความช่วยเหลือเป็น 2 ประเภทคือการสงเคราะห์เด็กกับการคุ้มครองสวัสดิภาพเด็ก จากผลการทดสอบพบว่า Naïve Bayes ให้ค่าดีที่สุด

จากผลการทดสอบพบว่าผลลัพธ์ที่ได้สามารถนำไปใช้จำแนกการให้ความช่วยเหลือเด็กที่ถูกทารุณได้ในขอบเขตของข้อมูลที่ทำการศึกษา สามารถพัฒนาการจำแนกประเภทให้มีความถูกต้อง แม่นยำมากขึ้นและสามารถเพิ่มประเภทการจำแนกที่ซับซ้อนขึ้นได้ โดยใช้กลุ่มตัวอย่างที่มีความหลากหลายและมีจำนวนมากขึ้น

เอกสารอ้างอิง

[1] สายชล สินสมบูรณ์ทอง. (2558). การทำเหมืองข้อมูล Data Mining. กรุงเทพฯ: จามจุรีโปรดักส์

[2] นันทกา สวัสดิพานิช. (2555). "ความรุนแรงต่อเด็ก: ความท้าทายใหม่สำหรับพยาบาล". พยาบาลสาธารณสุข ปีที่ 26 ฉบับที่ 1 2555, 101-110

[3] พระราชบัญญัติคุ้มครองเด็ก. (2546). ราชกิจจานุเบกษา เล่ม 120 ตอนที่ 95 ก.วันที่ 2 ตุลาคม 2546 หน้า 1-29.

[4] คลังข้อความไทย-อังกฤษ, https://www.nectec.or.th/corpus/index.php?league=pm [17 พฤศจิกายน 2559]

[5] Mehmed Kantardzie. (2003). Data Mining Concepts, Models, Methods, And Algorithms. USA: A John Wiley & Son, Inc.

[6] Fred Popowich. (2005). Using Text Mining and Natural Language Processing for Health Care Claims Processing. SIGKDD Explorations Volume 7 Issue 1, 59-66

[7] Qiwei He, et al.(2015). Automated Assessment of Patients' Self-Narratives for Posttraumatic Stress Disorder Screening Using Natural Language Processing and Text Mining. **Assessment, 24**, 157-172

[8] Meliha Yetisgen-Yildiz, Cosmin Adrian Bejan, Mark M. Wurfel. (2013). Identification of Patients with Acute Lung Injury from Free-Text Chest X-Ray Reports. Workshop on Biomedical Natural Language Processing (BioNLP 2013), 10-17

[9] Taxiachris Botsis, et al.(2011). Text Mining for the Vaccine Adverse Event Reporting System: medical text classification using informative feature selection. **J Am Med inform Assoc 2011**, 631-63

[10] Mihir Shekhar, Veera Raghavendra Chikka, Lini Thomas, et al.(2015). Identifying Medical Terms Related to Specific Diseases. 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 170-177

[11] Simon Baker, et al. (2015). Automatic Semantic Classification of Scientific Literature According to the Hallmarks of Cancer. **Oxford University Press 2015**, 1-9

[12] Download Rapidminer Studio,https://my.rapidminer.com/nexus/account/index.html #downloads[15 มิถุนายน 2559]

[13] Download Lucene, http://archive.apache.org/dist/lucene/java [19 สิงหาคม 2559]

ประวัติผู้เขียน

ชื่อ-นามสกุล	ปรางค์ปราณี ล้ำเลิศพระคุณ
ประวัติการศึกษา	ปีการศึกษา 2550 สำเร็จการศึกษาระดับปริญญาตรี คณะศิลปะศาสตร์ คณะมนุษยศาสตร์และสังคมศาสตร์ เอกภาษาเกาหลี มหาวิทยาลัย สงขลานครินทร์ (วิทยาเขตปัตตานี)
ตำแหน่งและสถานที่ทำงานปัจจุบัน	Web Application Programmer บริษัท SP COMNET จำกัด

DPU