ระบบประเมินความเสี่ยงต่อการติดเชื้อ เอช ไอ วี โดยขั้นตอนวิธีการเอ็กซ์มีนส์

Risk Assessment System for HIV Infection by Using X-means Algorithm

จักริน สุขสวัสดิ์ชน (Jakkarin Suksawatchon)* และ อุรีรัฐ สุขสวัสดิ์ชน (Ureerat Suksawatchon)*

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างระบบประเมินความเสี่ยงต่อการติดเชื้อ HIV ที่พิจารณาจากพฤติกรรมของการมีเพศสัมพันธ์เป็นหลักร่วมกับปัจจัยแฝงอื่นๆ โดยใช้เทคนิคการจัดกลุ่มข้อมูลของกลุ่มตัวอย่างจำนวน 639 ตัวอย่างที่มีพฤติกรรมเสี่ยงมาสร้างเป็นตัวแบบในการประเมิน ซึ่งในงานวิจัยนี้ได้เลือกใช้การจัดกลุ่มแบบเอ็กซ์มีนส์ เปรียบเทียบกับการจัดกลุ่มแบบเคมีนส์ โดยพบว่าผลจากการจัดกลุ่มด้วยวิธีการเอ็กซ์มีนจะให้ผลลัพธ์ที่ดีกว่าวิธีการจัดกลุ่มแบบเคมีนส์ โดยได้จำนวนกลุ่มที่เหมาะสมอยู่ที่ 14 กลุ่ม จากการประเมินความถูกต้องโดยพิจารณาผลการจัดกลุ่มกับข้อมูลทดสอบว่าตรงกับที่แพทย์ผู้เชี่ยวชาญด้าน HIV ระบุหรือไม่ พบว่ามีความถูกต้องในการจัดกลุ่มอยู่ที่ 82.14% นอกจากนี้ผู้วิจัยนำผลการจัดกลุ่มไปสร้างเป็นระบบประเมินความเสี่ยงโดยนำเสนอในรูปแบบของแอพพลิเคชั่น ซึ่งทำให้เกิดความสะดวกต่อการใช้งาน และมีความเป็นส่วนตัวในการที่จะประเมินความเสี่ยงด้วย

คำสำคัญ: ระบบประเมินความเสี่ยง เอชไอวี เอ็กซ์มีนส์

Abstract

The objective of this paper is to develop the risk assessment system for HIV infection which is considered by sexual behaviors and other hidden factors. The technique used to develop this system is clustering analysis to cluster the risk persons approximately 639 instances. In this research, we use X-means algorithm and compare with K-means algorithms. The clustering result of X-means algorithm provides the better result than K-means algorithm. There are 14 groups and each group has different risk behaviors and levels. When testing with the testing data, we have found that the accuracy of classification is 82.14%. Besides, we apply the clustering model to develop the assessment system in form of web application. This system is convenience and privacy for risk assessment.

Keyword: Risk Assessment System, HIV, X-means.

1. บทนำ

โรคเอดส์เป็นโรคติดต่อทางเพศสัมพันธ์ที่ส่งผลกระทบทั้งทางด้านร่างกาย จิตใจ สังคม และประเทศชาติ เนื่องจากเป็นโรคที่ยังไม่มียารักษาให้หายขาด และยังไม่มีวัคซีนที่จะสามารถป้องกันได้ผลอันเป็นที่น่าพอใจ ซึ่งการที่คนๆ หนึ่งจะเป็นโรคเอดส์ได้นั้น จะต้องได้รับเชื้อ HIV ที่อยู่ในสารคัดหลั่งของผู้มีเชื้อ เช่น น้ำเลือด น้ำเหลือง น้ำอสุจิ น้ำในช่องคลอด และน้ำนมแม่ โดยจะต้องผ่านช่องทางที่เข้าสู่กระแสเลือดโดยตรง และสาเหตุหลักที่ทำให้คนไทยได้รับเชื้อ HIV ก็คือการมีเพศสัมพันธ์โดยไม่ป้องกันกับผู้ติดเชื้อ HIV ไม่ว่าจะเป็นชายกับชาย หรือชายกับหญิงก็ล้วนแต่มีโอกาสติดเชื้อได้ทั้งนั้น

ในปัจจุบันมีหลายๆ หน่วยงานที่พยายามเข้ามาช่วยแก้ปัญหาเกี่ยวกับการระบาดของเชื้อ HIV โดยมีจุดมุ่งหมาย

* คณะวิทยาการสารสนเทศ มหาวิทยาลัยบูรพา

ให้จำนวนของผู้ติดเชื้อรายใหม่ลดลง แต่อย่างไรก็ตาม จากรายงานเกี่ยวกับผู้ป่วยเอดส์ของสำนักระบาดวิทยา กรมควบคุมโรค กระทรวงสาธารณสุขของไทย พบว่ายังมียอดผู้ติดเชื้อรายใหม่ในปี 2554 สูงถึง 10,000 รายซึ่งถือว่าเป็นจำนวนที่มากเลยทีเดียว และผู้ป่วยเอดส์ส่วนใหญ่ก็เป็นผู้ที่อยู่ในช่วงวัยทำงานและวัยเจริญพันธุ์ คือช่วงอายุระหว่าง 15-59 ปีนั่นเอง [1]

ปัญหาหลักที่ยังคงมีผู้ติดเชื้อรายใหม่เพิ่มขึ้นน่าจะมาจากสาเหตุที่ขาดความรู้ความเข้าใจ และจากการสอบถามแพทย์ผู้เชี่ยวชาญทางด้าน HIV ของคลีนิคแห่งหนึ่ง ยังพบว่าในแต่ละวันจะมีคนโทรศัพท์เข้ามาเพื่อปรึกษาเกี่ยวกับ HIV จำนวนมาก โดยส่วนใหญ่แล้วจะเป็นบุคคลที่ได้ไปมีพฤติกรรมเสี่ยงมา แล้วเกิดความวิตกกังวล ไม่สบายใจ อยากรู้ว่าตัวเองมีความเสี่ยงต่อการติดเชื้อ HIV หรือไม่ แต่ก็ไม่กล้าที่จะไปปรึกษาแพทย์ตามสถานพยาบาลทั่วไป เพราะความอาย ความกลัว ความวิตกกังวล กินไม่ได้นอนไม่หลับ กลัวคนอื่นจะรู้

ดังนั้นในงานวิจัยนี้จึงมีแนวคิดที่ว่า ถ้ามี “ระบบประเมิน” ให้บริการข่าวสารทางด้านการแพทย์รวมไปถึงการถามคำถามเพื่อประเมินความเสี่ยงต่อการติดเชื้อ HIV ด้วยตนเอง พร้อมทั้งคำแนะนำในการปฏิบัติตัวและยังทราบได้อีกว่าตนเองเสี่ยงมาจากปัจจัยใด ก็น่าทำให้บุคคลที่มีพฤติกรรมเสี่ยงสามารถวิเคราะห์หรือประเมินความเสี่ยงต่อการติดเชื้อ HIV ได้ในระดับหนึ่ง

ปัจจุบันงานวิจัยทางการแพทย์ที่เกี่ยวกับการประเมินความเสี่ยงต่อการติดเชื้อ HIV มีอยู่หลายงานวิจัย เช่น งานวิจัยของ Koblin BA และคณะ [2] ที่วิเคราะห์ปัจจัยเสี่ยงต่อการติดเชื้อ HIV ในกลุ่มชายรักชาย โดยใช้กลุ่มตัวอย่างจำนวน 4,295 คน โดยเป็นคนที่ยังไม่ติดเชื้อหรือผลเลือดเป็นลบ จาก 6 เมืองในสหรัฐอเมริกา ในการคัดกรองใช้โปรแกรม ACASI และใช้การวิเคราะห์ทางสถิติที่เรียกว่าสมการถดถอยในการวิเคราะห์ความเสี่ยง ซึ่งพบว่าส่วนใหญ่ในกลุ่มตัวอย่างมีความเสี่ยงต่อการติดเชื้อ HIV จากการมีเพศสัมพันธ์กับคู่นอนหลายคนถึง 32.3%ส่วนอีกงานวิจัยที่เกี่ยวข้องกับการหาปัจจัยเสี่ยงเช่นกันคือ งานวิจัยของ Maevel de Mello และคณะ [3] ซึ่งเป็นงานวิจัยที่ทำเพื่อหาปัจจัยเสี่ยงต่อการติดเชื้อ HIV ในกลุ่มชายรักชาย ประเทศบราซิล ในการคัดเลือกกลุ่มตัวอย่างใช้วิธีการที่เรียก Respondent-Driven Sampling หรือ RDS ซึ่งเป็นเทคนิคที่ใช้ในทางระบาดวิทยา เหมาะกับกลุ่มข้อมูลที่เข้าถึงยาก ส่วนการประเมินความเสี่ยงนั้นใช้การตอบแบบสอบถามผ่านโปรแกรม ACASI และวิเคราะห์ความเสี่ยงโดยใช้ซอฟต์แวร์ RDSAT ของมหาวิทยาลัย Cornell ประเทศสหรัฐอเมริกาซึ่งผลการประเมินทำให้ทราบว่ามีกลุ่มตัวอย่างมีความเสี่ยงต่อการติดเชื้อ HIV มาจากปัจจัยหลักใดบ้าง

นอกจากนี้ยังมีการนำงานวิจัยมาสร้างเป็นเครื่องมือในการประเมินความเสี่ยงในรูปแบบเว็บไซต์ ตัวอย่างเว็บ http://www.medindia.net/patients/calculators/hiv-risk-calculator.asp ที่นำงานวิจัยของ Vincent T DeVita Jr. และคณะ [4] มาสร้างเป็นแบบประเมินบนเว็บที่ให้บริการข่าวสารทางด้านการแพทย์รวมไปถึงการถามคำถามเพื่อประเมินความเสี่ยงต่อการติดเชื้อ HIV ซึ่งประกอบไปด้วย 9 คำถามอย่าง ซึ่งผลการประเมินจะบอกได้แค่ว่า อาจจะเสี่ยง หรือ อาจไม่เสี่ยง แต่ไม่สามารถบอกได้ว่าเสี่ยงมาจากปัจจัยหลักปัจจัยใด

1.1 แนวคิดโดยรวมของงานวิจัย

จากงานวิจัยดังกล่าวข้างต้นคณะผู้วิจัยเกิดความสนใจในประเด็นปัญหาสำหรับงานวิจัยนี้คือเว็บไซต์หลายเว็บที่สร้างเครื่องมือในการประเมินความเสี่ยงต่อการติดเชื้อ HIV โดยใช้ผลการวิจัยจากงานวิจัยที่กล่าวถึงข้างต้นจะประเมินความเสี่ยงเพียงว่า “อาจจะเสี่ยง” หรือ “อาจจะไม่เสี่ยง” แล้วแนะนำให้ผู้ตอบแบบสอบถามไปตรวจเลือดเพื่อความแน่ใจ แต่ไม่สามารถบอกได้ว่าผู้ตอบแบบสอบถามอาจมีความเสี่ยงมาจากปัจจัยใด และควรปรับพฤติกรรมอย่างไร ดังนั้นคำถามคือ “จะรู้ได้อย่างไรว่าความเสี่ยงต่อการติดเชื้อ HIV มาจากปัจจัยใดบ้าง”

นอกจากงานวิจัยที่กล่าวไปแล้วข้างต้น ยังมีงานวิจัยทางด้านการแพทย์ที่ศึกษาเกี่ยวกับปัจจัยเสี่ยงต่อการติดเชื้อ HIV ก็มีด้วยกันอยู่หลายงานวิจัย [5 - 11] แต่งานวิจัยดังกล่าวนั้นเน้นกลุ่มตัวอย่างที่เป็นชายรักชาย เนื่องจากมีจำนวนมากขึ้นในปัจจุบัน แต่สำหรับในประเทศไทยกลุ่มที่เสี่ยงต่อการติดเชื้อ HIV มีด้วยกันหลายกลุ่ม ได้แก่ กลุ่มชายรักชาย ผู้หญิงที่ขายบริการทางเพศ สาวประเภทสอง ซึ่งต่างพื้นที่กันก็มีผลทำให้ปัจจัยเสี่ยงต่อการติดเชื้อ HIV ต่างกันได้เช่นกัน นอกจากนี้ในการวิเคราะห์ปัจจัยเสี่ยงของงานวิจัยดังกล่าวข้างต้นใช้เพียงบางปัจจัย ตัวอย่างเช่นพิจารณาจากจำนวนคู่นอนเพียงอย่างเดียว ไม่ได้มีการวิเคราะห์ชนิด

คู่นอนว่าเป็นคู่นอนประจำ หรือคู่นอนชั่วคราว ซึ่งคู่นอนประจำจะมีความเสี่ยงน้อยกว่า และงานวิจัยดังกล่าวไม่ได้ใช้ความถี่ในการมีเพศสัมพันธ์รวมทั้งปัจจัยเรื่องโรคติดต่อทางเพศสัมพันธ์มาพิจารณาความเสี่ยงร่วมด้วย ซึ่งทำให้งานวิจัยดังกล่าวไม่เหมาะที่จะนำมาใช้กับการวิเคราะห์หรือประเมินความเสี่ยงต่อการติดเชื้อ HIV ในประเทศไทย

แนวทางการแก้ไขปัญหาดังกล่าวสามารถนำเทคนิคการ "วิเคราะห์กลุ่มข้อมูล (Clustering Analysis)" มาประยุกต์ใช้ โดยการนำข้อมูลตัวอย่างมาจัดเป็นกลุ่ม ๆ ซึ่งผลการจัดกลุ่มทำให้ทราบว่าข้อมูลตัวอย่างที่อยู่ในกลุ่มเดียวกันมีพฤติกรรมหรือมีปัจจัยเสี่ยงต่อการติดเชื้อ HIV ที่คล้ายกัน จากนั้นผู้วิจัยจะนำกลุ่มข้อมูลเหล่านี้ไปวิเคราะห์ต่อไปเพื่อหาว่าปัจจัยเสี่ยงใดในแต่ละกลุ่มที่ทำให้เกิดความเสี่ยงต่อการติดเชื้อ HIV ซึ่งในการที่จะพิจารณาว่าแต่ละกลุ่มเป็นกลุ่มที่เสี่ยงมาจากปัจจัยใดหรือมีความโดดเด่นในปัจจัยบ้าง มีความเสี่ยงต่อการติดเชื้อ HIV กี่ปัจจัยนั้น ผู้วิจัยได้นำข้อมูลอ้างอิงจากงานวิจัยต่างประเทศที่ได้กล่าวไว้ในข้างต้นมาเป็นแนวทางในการพิจารณา พร้อมทั้งให้แพทย์ผู้เชี่ยวชาญด้าน HIV เป็นผู้ตรวจสอบความน่าเชื่อถืออีกด้วย จากนั้นผู้วิจัยจึงนำผลของการจัดกลุ่มไปจัดทำเป็นเครื่องมือในการประเมินความเสี่ยงต่อการติดเชื้อ HIV ให้กับผู้ตอบแบบสอบถาม ทำให้รู้ว่าตัวเขาจัดอยู่ในกลุ่มใด มีความเสี่ยงจากปัจจัยใดบ้าง และยังเป็นประโยชน์สำหรับผู้ที่สนใจอื่นที่อยากรู้ว่าตัวเองมีพฤติกรรมการใช้ชีวิตในปัจจุบันที่เสี่ยงต่อการติดเชื้อ HIV หรือไม่ พร้อมทั้งมีคำแนะนำวิธีการปฏิบัติตัวหลังจากรู้ผลการประเมินความเสี่ยงได้อีกด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ขั้นตอนการจัดกลุ่มแบบเคมีนส์ [12]

เคมีนส์คือหนึ่งในวิธีการจัดกลุ่มข้อมูลซึ่งเป็นการเรียนรู้โดยไม่มีผู้สอน และเป็นการแก้ปัญหาการจัดกลุ่มที่รู้จักกันทั่วไป โดยขั้นตอนการจัดกลุ่มแบบเคมีนส์จะแบ่งข้อมูลออกเป็น K กลุ่ม โดยแทนแต่ละกลุ่มด้วยค่าเฉลี่ยของสมาชิกภายในกลุ่ม ซึ่งการจัดกลุ่มแบบเคมีนส์มีขั้นตอนดังนี้

1. กำหนดจำนวนกลุ่ม K กลุ่ม และกำหนดจุดศูนย์กลางเริ่มต้นจำนวน K กลุ่มแบบสุ่ม

2. นำข้อมูลทั้งหมดจัดเข้ากลุ่มที่มีจุดศูนย์กลางที่อยู่ใกล้วัตถุนั้นมากที่สุด โดยคำนวณจากการวัดระยะห่างระหว่างข้อมูลที่น้อยที่สุด ในที่นี้จะใช้มาตรวัดยูคลิด (Euclidean Metric)

3. คำนวณจุดศูนย์กลาง K จุดใหม่ โดยหาจากค่าเฉลี่ยข้อมูลทุกตัวที่อยู่ในกลุ่ม

4. ทำซ้ำในข้อ 2. จนกระทั่งจุดศูนย์กลางไม่เปลี่ยนแปลง

2.2 การจัดกลุ่มแบบเอ็กซ์มีนส์

Dan Peeleg และ Addrew Moore [13] ได้ทำการศึกษาวิจัยเกี่ยวกับขั้นตอนการจัดกลุ่มแบบเอ็กซ์มีนส์ (X-means Algorithm) เพื่อที่จะพัฒนาและปรับปรุงประสิทธิภาพของขั้นตอนการจัดกลุ่มแบบเคมีนส์ เนื่องจากเคมีนส์นั้นมีปัญหาในเรื่องของการกำหนดจำนวนกลุ่ม (K) ที่ต้องกำหนดโดยผู้ใช้ ดังนั้นหากผู้ใช้ไม่มีความรู้จึงเป็นการยากที่จะรู้ว่าจำนวนกลุ่มเท่าไหร่จึงจะเหมาะสมกับข้อมูล เนื่องจากข้อมูลแต่ละชุดจะมีคุณลักษณะที่ไม่เหมือนกัน ดังนั้นการจัดกลุ่มแบบเอ็กซ์มีนส์จะให้ผู้ใช้กำหนดช่วงของจำนวนกลุ่มที่ต้องการ เช่น อยู่ในช่วง 2 กลุ่ม (จำนวนกลุ่มที่น้อยที่สุด) ถึง 20 กลุ่ม (จำนวนกลุ่มที่มากที่สุด) จากนั้นการจัดกลุ่มแบบเอ็กซ์มีนส์จะแบ่งกลุ่มจนกระทั่งได้จำนวนกลุ่มที่เหมาะสมภายในช่วงที่กำหนด โดยมีขั้นตอนดังนี้

1. ขั้นตอน Improve-Params เป็นขั้นตอนของการแบ่งกลุ่มด้วยวิธีการเคมีนส์

2. ขั้นตอน Improve-Structure เป็นขั้นตอนในการตรวจสอบว่าในแต่ละกลุ่มที่ได้จากขั้นตอนแรกนั้นสามารถที่จะแบ่งกลุ่มต่อไปอีกได้หรือไม่ โดยใช้ตัววัดที่เรียกว่า BIC (Bayesian Information Criterion) และ AIC (Akaike Information Criterion)

3. ตรวจสอบค่า K ที่ได้จากขั้นตอนที่ 2 นั่นคือถ้าค่า K มากกว่าค่า Max ให้หยุดการทำงาน แต่ถ้า ค่า K ยังไม่เกินค่า Max ให้กลับไปทำขั้นตอนที่ 1

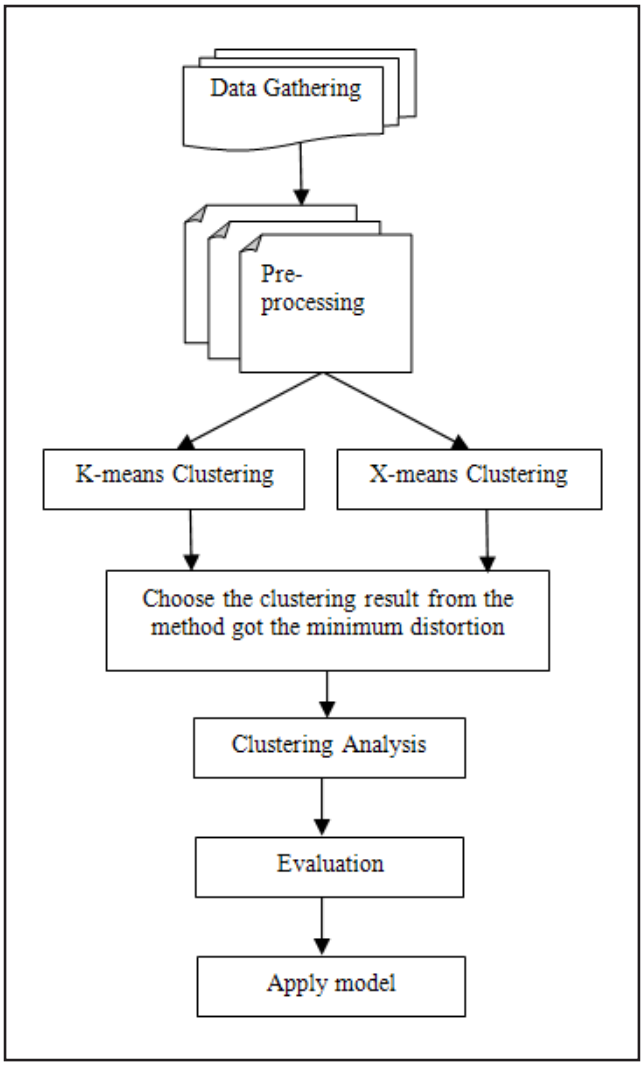
3. วิธีการที่นำเสนอ

ภาพการทำงานโดยรวมเป็นดังภาพที่ 1 โดยมีรายละเอียดดังนี้

3.1 การรวบรวมข้อมูลจากแบบสอบถาม (Data Gathering)

งานวิจัยนี้ได้ทำการรวบรวมข้อมูลจากคลินิคแห่งหนึ่ง โดยการให้อาสาสมัครตอบแบบสอบถาม ซึ่งเป็นคนไทยมีอายุระหว่าง 18-50 ปี เป็นชาย 248 คน หญิง 165 คน และสาวประเภทสอง 226 คนโดยเป็นกลุ่มที่ทำงานในพื้นที่เมือง

พัทยา จ. ชลบุรี ซึ่งคำถามสำคัญๆ จะเกี่ยวข้องกับพฤติกรรมเสี่ยงในการติดเชื้อ HIV เช่น ข้อมูลด้านการมีเพศสัมพันธ์ การใช้ยาเสพติด เป็นต้น



ภาพที่ 1 *ขั้นตอนการดำเนินงานวิจัย*

3.2 การจัดเตรียมข้อมูล (Pre-processing)

ในขั้นตอนของการจัดเตรียมข้อมูลจะเป็นการสกัดเอาคุณลักษณะเด่นของข้อมูล โดยเลือกเอาเฉพาะคุณลักษณะ (Attributes) ที่เป็นปัจจัยเสี่ยงต่อการติดเชื้อ HIV เท่านั้นโดยพิจารณาจากงานวิจัยทางการแพทย์ จากนั้นผู้วิจัยได้ทำการแปลงข้อมูลจากตัวอักขระให้เป็นตัวเลขเพื่อให้สะดวกต่อการคำนวณและการวิเคราะห์ข้อมูล โดยในการเลือกคุณลักษณะนั้นเนื่องจากเป็นข้อมูลเฉพาะด้าน จึงจำเป็นต้องใช้แพทย์และนักวิจัยที่เชี่ยวชาญทางด้าน HIV ซึ่งสามารถแสดงคุณลักษณะ ที่เกี่ยวข้องกับความเสี่ยงต่อการติดเชื้อ HIV ได้ทั้งหมด 72 คุณลักษณะ โดยผู้วิจัยขอยกตัวอย่างการแทนค่าระดับความเสี่ยงเพียงจำนวน 3 คุณลักษณะดังตารางที่ 1

ตารางที่ 1 *ตารางการประเมินประสิทธิภาพการค้นคืนข้อมูล*

Attribute Name	Description	Original Value	Convert to
Gender	เพศ	Male	1
		Female	2
		Transgender	3
STCAT2F	ในช่วง 3 เดือนที่ผ่านมา ท่านมีเพศสัมพันธ์กับคู่นอนประจำเพศหญิงหรือไม่	No	0
		Yes	1
		Don't know	10
		Refuse to Answer	11
NBRSTF	ในช่วง 3 เดือนที่ผ่านมา ท่านมีเพศสัมพันธ์กับคู่นอนประจำเพศหญิงกี่คน	1-2 partners	1
		3-10 partners	2
		More than 10 partners	3
		Don't know	10
		Refuse to Answer	11
		Skipped	0

3.3 ขั้นตอนการจัดกลุ่ม (Clustering)

เป็นการนำข้อมูลตัวอย่างที่ผ่านในขั้นตอนที่ 3.2 มาจัดกลุ่มซึ่งในงานวิจัยนี้ได้เลือกใช้เทคนิคการจัดกลุ่มข้อมูล 2 เทคนิค คือ ขั้นตอนการจัดกลุ่มแบบเคมีนส์และขั้นตอนวิธีการแบบเอ็กซ์มีนส์ โดยทำการทดลองตั้งแต่จำนวน 2 กลุ่มถึง 16 กลุ่ม

ตารางที่ 2 และ 3 เป็นการกำหนดพารามิเตอร์ที่สำคัญของขั้นตอนการจัดกลุ่มแบบเคมีนส์และเอ็กซ์มีนส์ตามลำดับ ซึ่งแสดงให้เห็นว่าได้ผลลัพธ์ที่ดีกว่าขั้นตอนการจัดกลุ่มแบบ

ตารางที่ 2 *ผลการทดลองด้วยวิธีการจัดกลุ่มแบบเคมีนส์*

k	10	11	12	13	14	15	16
Distortion	702.64	708.45	690.72	675.01	655.02	674.10	653.78

ตารางที่ 3 *ผลการทดลองด้วยวิธีการจัดกลุ่มแบบเอ็กซ์มีนส์*

Seed	Min Clusters	Max Clusters	จำนวนกลุ่มที่ได้	Distortion
2	2	16	14	543.554654
5	2	16	13	547.082603
10	2	16	10	560.839225
15	2	16	14	541.680128
16	2	16	13	549.507960
20	2	16	12	548.953354
25	2	16	13	545.110609

เคมีนส์ เนื่องจากขั้นตอนการจัดกลุ่มแบบเอ็กซ์มีนส์ ให้ค่าความผิดพลาด (Distortion) ที่น้อยกว่า ภายใต้เงื่อนไขจำนวนค่า K ที่เท่ากัน ดังนั้นผลการจัดกลุ่มด้วยวิธีการแบบเอ็กซ์มีนส์ให้จำนวนกลุ่มที่เหมาะสมอยู่ที่ 14 กลุ่ม โดยให้ค่าความผิดพลาดที่ 541.68 ในขณะที่การจัดกลุ่มแบบเคมีนส์จะให้ค่าความผิดพลาดที่ 655.02

3.4 การวิเคราะห์กลุ่มข้อมูล (Clustering Analysis)

ในการวิเคราะห์ถึงคุณลักษณะของแต่ละกลุ่มว่ามีความโดดเด่นในปัจจัยบ้าง มีความเสี่ยงต่อการติดเชื้อ HIV กี่ปัจจัยนั้น ผู้วิจัยได้พิจารณาจากงานวิจัยทางการแพทย์เกี่ยวกับปัจจัยเสี่ยงต่อการติดเชื้อ HIV ร่วมกับการพิจารณาจากข้อมูลที่อยู่ในแต่ละกลุ่มว่าข้อมูลโดยส่วนใหญ่มีแนวโน้มไปทางปัจจัยใด แล้วจึงสรุปพฤติกรรมของแต่ละกลุ่มออกมา ตารางที่ 3 เป็นตารางสรุปอันดับความเสี่ยงและภาพรวมคุณลักษณะของแต่ละกลุ่ม โดย M = ชาย, F = หญิง, TG = สาวประเภทสอง, MSM = ชายรักชาย, CD = Condom (ถุงยางอนามัย), STI = โรคติดต่อทางเพศสัมพันธ์ จากนั้นจึงนำผลสรุปจากตารางที่ 4 นี้มาวิเคราะห์คุณลักษณะและพฤติกรรมอย่างละเอียดของแต่ละกลุ่ม แต่เนื่องจากแต่ละกลุ่มมีรายละเอียดที่ค่อนข้างมาก ดังนั้นผู้วิจัยจึงขอยกตัวอย่างเพียงหนึ่งกลุ่มเท่านั้น ดังนี้

ตัวอย่างลักษณะความเสี่ยงของกลุ่มที่ 1 ประกอบด้วยสมาชิกจำนวน 31 ตัวอย่าง โดยกลุ่มนี้ส่วนมากเป็นเพศชายมีเพศสัมพันธ์ทั้งกับหญิง และกับชายด้วยกันเอง และกับสาวประเภทสอง ในทางการแพทย์จะเรียกว่ากลุ่มรักสองเพศ (Bi-Sexual) นอกจากนี้กลุ่มนี้มีจำนวนคู่นอนหลายคน โดยมีคู่นอนประจำหญิง คู่นอนชั่วคราวหญิง ลูกค้าหญิง คู่นอนประจำชาย คู่นอนชั่วคราวชาย และลูกค้าชาย ซึ่งมากกว่า 10 คน และยังมีคู่นอนที่เป็นสาวประเภทสองแต่จำนวนไม่มาก 1-2 คน มีการใช้ถุงยางอนามัยเป็นบางครั้ง กลุ่มนี้จึงมีความเสี่ยงสูงเนื่องจากไม่ได้ป้องกันทุกครั้งที่มีเพศสัมพันธ์ และยังเป็นกลุ่มที่ขายบริการให้กับหญิง ขายบริการให้กับชาย ขายบริการให้กับสาวประเภทสอง แต่กลุ่มนี้จะซื้อบริการเฉพาะกับหญิงเท่านั้น โดยสามารถสรุปปัจจัยที่มีความเสี่ยงได้ 20 ข้อดังนี้

1. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับคู่นอนชั่วคราวหญิง มากกว่า 10 ครั้ง

2. ใช้ถุงยางอนามัยเป็นบางครั้งกับคู่นอนชั่วคราวหญิง

ตารางที่ 4 *อันดับความเสี่ยงและภาพรวมคุณลักษณะของแต่ละกลุ่ม*

ลำดับที่ของความเสี่ยง	กลุ่มที่	คุณลักษณะเด่นของกลุ่ม		
		เพศ	มีคู่นอนเป็นเพศ	จำนวนปัจจัยเสี่ยง
1	8	M	M, F, TG	24
2	1	M	M, F, TG	20
3	12	M	M (MSM)	15
4	7	M	M, F, TG	16
5	11	F	M	12
6	10	TG	M	9
7	13	F	M	9
8	9	TG	M, TG	9
9	14	F	M	6
10	4	TG	M	5
11	6	TG	M	2
12	5	TG	M	2
13	3	M	M (MSM)	2
14	2	M	F	2

3. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับลูกค้าหญิงจำนวนมากกว่า 10 คน

4. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับลูกค้าหญิงมากกว่า 10 ครั้ง

5. ใช้ถุงยางอนามัยเป็นบางครั้งกับลูกค้าหญิง

6. เคยได้รับเงินเพื่อแลกกับการมีเพศสัมพันธ์กับหญิง (ขายบริการให้หญิง)

7. เคยจ่ายเงินเพื่อแลกกับการมีเพศสัมพันธ์กับหญิง (ซื้อบริการจากหญิง)

8. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับเพศหญิงหลังจากดื่มเหล้า เบียร์ หรือแอลกอฮอล์เป็นบางครั้ง

9. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับคู่นอนประจำชาย จำนวนมากกว่า 10 คน

10. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับคู่นอนประจำชาย มากกว่า 10 ครั้ง

11. ใช้ถุงยางอนามัยเป็นบางครั้งกับคู่นอนประจำชาย

12. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับคู่นอนชั่วคราวชาย จำนวนมากกว่า 10 คน

13. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับคู่นอนชั่วคราวชาย มากกว่า 10 ครั้ง

14. ใช้ถุงยางอนามัยเป็นบางครั้งกับคู่นอนชั่วคราวชาย

15. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับลูกค้าชาย จำนวนมากกว่า 10 คน

16. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับลูกค้าชาย มากกว่า 10 ครั้ง

17. ใช้ถุงยางอนามัยเป็นบางครั้งกับลูกค้าชาย

18. ในช่วง 3 เดือนที่ผ่านมามีเพศสัมพันธ์กับเพศชายหลังจากดื่มเหล้า เบียร์ หรือแอลกอฮอล์เป็นบางครั้ง

19. เคยได้รับเงินเพื่อแลกกับการมีเพศสัมพันธ์กับชาย (ขายบริการให้ชาย)

20. เคยได้รับเงินเพื่อแลกกับการมีเพศสัมพันธ์กับสาวประเภทสอง (ขายบริการให้สาวประเภทสอง)

<u>ตัวอย่างคำแนะนำสำหรับกลุ่มที่ 1</u>

เนื่องจากกลุ่มที่ 1 มีจำนวนคู่นอนหลายคน (Multiple Partners Sex) แต่มีการป้องกันไม่ดีนักเพราะใช้ถุงยางอนามัยเป็นบางครั้ง และกลุ่มนี้ยังมีความเสี่ยงจากพฤติกรรมของการขายบริการด้วย ดังนั้นข้อแนะนำเพื่อลดความเสี่ยงสำหรับกลุ่มนี้มีดังต่อไปนี้

1. ปรับพฤติกรรมการใช้ถุงยางอนามัยจากการใช้เป็นบางครั้งให้เป็นใช้ทุกครั้งที่มีเพศสัมพันธ์ ไม่ว่าจะเป็นคู่นอนประจำ คู่นอนชั่วคราว หรือลูกค้าก็ตาม เพราะกลุ่มที่ 1 นี้มีคู่นอนหลายคนโอกาสติดเชื้อมีสูงมาก

2. แนะนำให้ไปปรึกษาแพทย์เกี่ยวกับการขลิบอวัยวะเพศเพื่อลดความเสี่ยง เนื่องจากการขลิบจะทำให้ลดการสะสมของเชื้อโรคลงได้

3. การมีเพศสัมพันธ์หลังจากดื่มแอลกอฮอล์ แม้ว่าการดื่มแอลกอฮอล์จะไม่ได้เป็นปัจจัยเสี่ยงต่อการติดเชื้อ HIV โดยตรง แต่ก็เป็นพฤติกรรมที่จะนำไปสู่การติดเชื้อได้ เนื่องจากคนที่เมาแล้วจะควบคุมสติไม่ได้ ขาดการยับยั้งชั่งใจ เกิดอารมณ์ทางเพศได้สูงกว่าคนปกติ และยังไม่ตระหนักในการใช้ถุงยางอนามัยด้วยเพราะขาดสติ ดังนั้นจึงควรค่อยๆ ลดการดื่มแอลกอฮอล์ลง

4. พฤติกรรมการขายบริการนั้นเนื่องจากด้วยอาชีพที่ต้องขายบริการ โอกาสในการมีเพศสัมพันธ์จึงมีมากกว่าอาชีพอื่นๆ ดังนั้นจึงควรเน้นการมีเพศสัมพันธ์อย่างปลอดภัยโดยการใช้ถุงยางอนามัยทุกครั้งที่มีเพศสัมพันธ์ และการใช้สารหล่อลื่นซึ่งจะต้องเป็นสารหล่อลื่นชนิดที่ละลายกับน้ำได้ ห้ามใช้โลชั่นหรือวาสลิน หรือสารที่มีส่วนประกอบของน้ำมัน เนื่องจากน้ำมันจะทำปฏิกิริยากับถุงยางอนามัยจะมีผลให้ถุงยางอนามัยแตกหรือฉีกขาดได้ และไม่ควรสวมถุงยางอนามัย 2 ชั้น ไม่ควรใช้ถุงยางอนามัยชายคู่กับถุงยางอนามัยหญิง (Female condom) เพราะจะทำให้ถุงยางอนามัยเกิดการเสียดสีและฉีกขาดได้

5. เพื่อประสิทธิภาพของการป้องกันควรเลือกใช้ถุงยางอนามัยที่มีขนาดพอดี ดูวันเดือนปีที่หมดอายุ ซึ่งโดยปกติตามท้องตลาดในเมืองไทยจะมีขนาดให้เลือก เช่น 49 มิลลิเมตร, 52 มิลลิเมตร, 54 มิลลิเมตร, 56 มิลลิเมตร และในกรณีที่ถุงยางอนามัยแตกหรือฉีกขาดระหว่างที่มีเพศสัมพันธ์ จะต้องรีบหยุดการมีเพศสัมพันธ์ทันทีโดยรีบถอนสมอแล้วทำความสะอาดร่างกายทันที จากนั้นให้รีบไปปรึกษาแพทย์เพื่อประเมินความเสี่ยง เพราะไม่ใช่แค่ความเสี่ยงต่อการติดเชื้อ HIV เท่านั้นแต่ยังรวมไปถึงความเสี่ยงต่อการตั้งครรภ์ด้วย (ในกรณีที่คู่นอนเป็นเพศหญิง)

6. เพื่อเป็นการเพิ่มความระมัดระวังในระหว่างมีเพศสัมพันธ์ถ้าเป็นไปได้ควรเปิดไฟบ้างเล็กน้อยเพื่อให้สามารถมองเห็น หรือรับรู้ได้ว่าถุงยางอนามัยยังอยู่ในสภาพปกติไม่แตก ไม่ฉีกขาด และถ้าหากเป็นกรณีที่เสร็จช้าก็ควรจะมีการตรวจสอบถุงยางอนามัยเป็นระยะ หรืออาจจะมีการเปลี่ยนถุงยางอนามัยชิ้นใหม่ระหว่างทางได้

7. ในกลุ่มที่ขายบริการนั้น กรณีที่ลูกค้าไม่ยอมให้ใช้ถุงยางอนามัยเราสามารถที่จะปฏิเสธการมีเพศสัมพันธ์ได้เพื่อปกป้องตัวเอง

8. สำหรับผู้ที่แพ้ถุงยางอนามัย ก็ควรที่จะลองเปลี่ยนยี่ห้อ เปลี่ยนรุ่น เปลี่ยนสี เปลี่ยนกลิ่น หรือถ้าเปลี่ยนแล้วยังไม่หายก็ควรปรึกษาแพทย์ต่อไป

3.5 การประเมินผล (Evaluation)

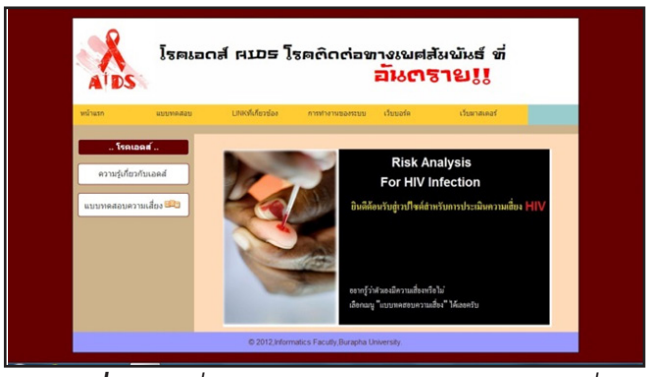
ในการประเมินผลการจัดกลุ่ม งานวิจัยนี้ได้สร้างข้อมูลสำหรับทดสอบ (Testing Data) โดยผู้เชี่ยวชาญ จำนวน 28 ตัวอย่าง แล้วนำข้อมูลเหล่านี้ไปผ่านการจัดกลุ่ม จากนั้นการประเมินความถูกต้องจะพิจารณาผลการจัดกลุ่มว่าตรงกับที่แพทย์ผู้เชี่ยวชาญด้าน HIV ระบุหรือไม่ ซึ่งผลของการจัดกลุ่มพบว่ามีความถูกต้องอยู่ที่ 82.14% (ให้คำตอบที่ตรงกับ

ผู้เชี่ยวชาญจำนวน 23 ตัวอย่าง)

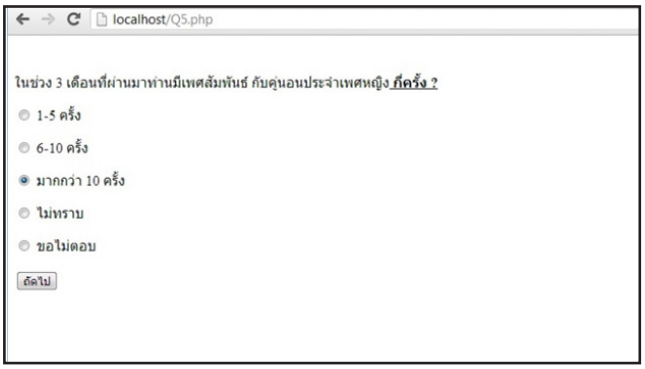
และเมื่อผู้วิจัยได้พิจารณาถึงตัวข้อมูลที่จัดกลุ่มผิดพบว่าเป็นกลุ่มที่มีพฤติกรรมใกล้เคียงกันและมีความเสี่ยงต่อการติดเชื้อ HIV ที่สูงเหมือนกัน เช่น กลุ่มที่ 1 กับกลุ่มที่ 8 ซึ่งเป็นกลุ่มเพศชายที่มีเพศสัมพันธ์กับทั้งชายและหญิงและสาวประเภทสอง มีคู่นอนที่หลายคนและมีการป้องกันไม่ดี โดยทั้งสองกลุ่มนี้จัดว่ามีความเสี่ยงที่สูงเหมือนกันเพียงแต่รายละเอียดปลีกย่อยในพฤติกรรมจะต่างกันบ้าง เช่น กลุ่มที่ 8 เป็นโรคติดต่อทางเพศสัมพันธ์ ส่วนกลุ่มที่ 1 ไม่เป็นโรคติดต่อทางเพศสัมพันธ์ ซึ่งจากความผิดพลาดในการจัดกลุ่มในกรณีนี้ถือว่าไม่ได้มีความรุนแรงมากนัก เพราะผลลัพธ์คือมีความเสี่ยงสูงเหมือนกัน

4. ระบบประเมินความเสี่ยง

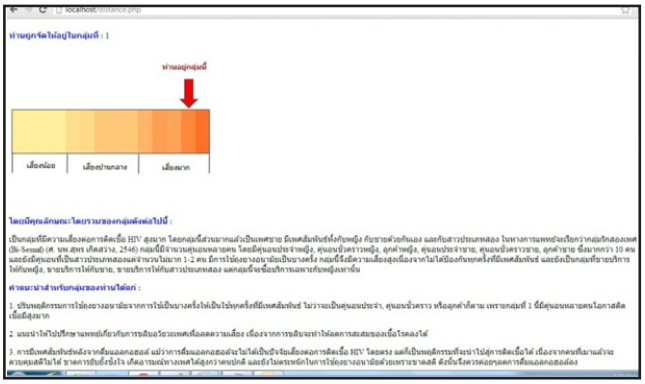
สิ่งที่ได้จากงานวิจัยนี้คือ ปัจจัยที่เสี่ยงต่อการติดเชื้อ HIV และลักษณะพฤติกรรมเฉพาะของแต่ละกลุ่ม ผลการจัดกลุ่มถูกนำมาสร้างเป็นตัวแบบที่สามารถใช้ในการวิเคราะห์ความเสี่ยงต่อการติดเชื้อ HIV พร้อมทั้งคำแนะนำเพื่อหลีกเลี่ยงความเสี่ยง ในรูปแบบของเว็บแอพพลิเคชั่น ดังตัวอย่างภาพที่ 2 โดยเมื่อผู้ใช้ป้อนข้อมูลพฤติกรรมเข้าไปแล้วระบบสามารถที่จะบ่งชี้ได้ว่ามีความเสี่ยงหรือไม่ โดยใช้แทบสีในการระบุความเสี่ยง และถ้ามีความเสี่ยงแล้วเสี่ยงมาจากปัจจัยใด ซึ่งสามารถที่จะนำผลการวิเคราะห์ความเสี่ยงไปใช้เป็นแนวทางในการให้ความรู้ โดยระบบการวิเคราะห์ความเสี่ยงนี้มีส่วนที่เป็นระบบแนะนำวิธีการปฏิบัติตัวหลังจากรู้ผลการประเมินความเสี่ยงอีกด้วย และยังสะดวกต่อการใช้งานสำหรับผู้ที่สนใจแต่ไม่สะดวกใจที่จะมาสถานพยาบาล โดยสามารถประเมินตนเองเบื้องต้นด้วยตัวแบบที่นำเสนอผ่านเครือข่ายอินเทอร์เน็ตได้ด้วย ดังตัวอย่างในภาพที่ 3 และ 4



ภาพที่ 2 ตัวอย่างเว็บเพจของระบบประเมินความเสี่ยง



ภาพที่ 3 ตัวอย่างเว็บเพจส่วนที่เป็นการตอบคำถาม



ภาพที่ 4 ตัวอย่างเว็บเพจในส่วนผลการประเมิน

จากตารางที่ 5 จะเห็นว่าค่าระดับคะแนนโดยเฉลี่ยจากผู้เชี่ยวชาญทั้งสองท่านในส่วนของรูปแบบของการนำเสนอจะอยู่ที่ 3.12 คะแนน จากคะแนนเต็ม 4 และค่าระดับคะแนนในส่วนของการยอมรับในตัวผลลัพธ์จะอยู่ที่ 3.36 คะแนนจากคะแนนเต็ม 4 ซึ่งถือว่าอยู่ในเกณฑ์ดี

5. สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างระบบประเมินสำหรับวิเคราะห์ความเสี่ยงต่อการติดเชื้อ HIV โดยพิจารณาจากพฤติกรรมของการมีเพศสัมพันธ์เป็นหลักรวมถึงปัจจัยแฝงอื่นๆ ซึ่งการที่จะทราบได้ว่าแต่ละคนมีพฤติกรรมเสี่ยงด้านใด จำเป็นจะต้องทำการแบ่งกลุ่มข้อมูลตัวอย่างเสียก่อน ซึ่งงานวิจัยนี้ใช้ขั้นตอนการจัดกลุ่มแบบเอ็กซ์มีนส์มาทำการจัดกลุ่มข้อมูลพฤติกรรมจำนวน 639 ตัวอย่าง ซึ่งผลที่ได้จากการจัดกลุ่มด้วยวิธีการเอ็กซ์มีนให้ผลการจัดกลุ่มที่เหมาะสมกว่าขั้นตอนการจัดกลุ่มแบบเคมีนส์ ด้วยจำนวน 14 กลุ่ม จากนั้นทำการวิเคราะห์คุณลักษณะของแต่ละกลุ่มออกมาว่าแต่ละกลุ่มมีพฤติกรรมที่มีความเสี่ยงต่อการติดเชื้อ HIV ในปัจจัยใดบ้าง มีระดับความเสี่ยงมากน้อยแค่ไหน นอกจากนี้

ตารางที่ 5 *แบบประเมินความพึงพอใจในการใช้ระบบ*

หัวข้อประเมิน		คะแนนจากผู้เชี่ยวชาญคนที่ 1	คะแนนจากผู้เชี่ยวชาญคนที่ 2	คะแนนเฉลี่ย
รูปแบบของการนำเสนอ				
1	ความดึงดูดใจของระบบ	2	3	2.5
2	การใช้งานง่าย เข้าใจง่าย	4	4	4
3	เนื้อหาและข้อมูล	4	4	4
4	ขนาด สี ของตัวอักษร	2	2	2
	ค่าคะแนนเฉลี่ย	3	3.25	3.12
การยอมรับในตัวผลลัพธ์				
1	ความพึงพอใจในผลลัพธ์	3	4	3.5
2	ความถูกต้องของระบบ	3	3	3
3	ความรวดเร็วในการประมวลผล	3	4	3.5
4	ความสะดวกต่อการใช้งาน	3	4	3.5
	ค่าคะแนนเฉลี่ย	**3**	**3.75**	**3.36**
ค่าระดับคะแนน : 4 = ดีมาก 3 = ดี 2 = พอใช้ 1 = ไม่ดีต้องปรับปรุง				

ผู้วิจัยยังนำผลการจัดกลุ่มไปประยุกต์ใช้เพื่อสร้างระบบประเมินความเสี่ยงต่อการติดเชื้อ HIV โดยนำเสนอในรูปแบบของแอพพลิเคชั่น ซึ่งทำให้เกิดความสะดวกต่อการใช้งาน และมีความเป็นส่วนตัวในการที่จะประเมินความเสี่ยงด้วย

6. เอกสารอ้างอิง

[1] สำนักระบาดวิทยา กรมควบคุมโรค. สถานการณ์เอดส์ของประเทศไทยในปัจจุบัน. กรุงเทพฯ : สำนักระบาดวิทยา, กรมควบคุมโรค, 2554.

[2] B. A. Koblin, M. J. Husnik, G. Colfax, Y. Huang, M. Madison, K. Mayer, P. J. Barresi, T. J. Coates. M. A. Chesney, and S. Buchbinder, "Risk Factors for HIV Infection Among Men Who Have Sex with Men." AIDS, Vol. 20, No. 5, pp. 731-9, 2006.

[3] M. Mello, et al. Assessment of Risk Factors for HIV Infection Among Men Who Have Sex With Men in the Metropolitan Area Of Campinas City, Brazil, Using Respondent-Driven Sampling, The Population Council Inc., 2008.

[4] V. T. Vita , S. Hellman, and S. A. Rosenberg. Biology, Diagnosis, Treatment, and Prevention, Lippincott-Raven Publishers, 1997.

[5] Y.W.Paul. AIDS PREVENTION PRO. Available online at http://www. stylite.net/prevaids/000-prevaids-welcome-details-th.htm.

[6] L. George. Sex Frequency Survey. Available online at http://www.intellasia.net/indonesian-men-top-sex-frequency-survey-179671.

[7] U.S. Centers for Disease Control and Prevention. 2009, Oral sex and HIV risk: CDC HIV/AIDS facts. Available online at http://www.cdc.gov/hiv/resources/factsheets/oralsex.htm.

[8] M. C. Boily, et al. "Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies." Lancet Infect Dis, Vol. 9, No. 2, pp. 118-129, 2009.

[9] D. T. Halperin. "Heterosexual anal intercourse: prevalence, cultural factors, and HIV infection and other health risks Part I." AIDS Patient Care STDS, Vol. 13, No. 12, pp. 717-30, 1999.

[10] S. Kalichman, et al. "Heterosexual anal intercourse among community and clinical settings in Cape Town, South Africa." Sex Trans. Infect, Vol. 85, No. 6, pp. 408-10, 2009.

[11] E. Vittinghoff, et al. "Per-contact risk of human immunodeficiency virus transmission between male sexual partners." *American Journal of Epidemiology*, 150: pp. 306-311, 1999.

[12] J. Han, M. Kamber, and J. Pei. Data Mining Concepts and Techniques, Morgan Kaufmann, pp. 383-403, 2006.

[13] D. Pelleg and A. Moore. "X-means: Extending K-means with efficient estimation of the number of clusters." *Proceding of 17th International Conferrence on Machine Learning* (*ICML'00*), pp. 727–734, 2000.