

การพัฒนาระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วย โดยใช้เทคนิคเหมืองข้อความ

Development of A Decision Support System for the Patient's Symptoms and Herbal Respond by Text Mining Techniques.

Received:	February	25, 2019
Revised:	May	23, 2019
Accepted:	May	27, 2019

พิศาล สุขชี (Phisan Sookhee)*

เจษฎา ชาตรี (Jesada Chatree)**

สิริกัลยา สุขชี (Sirikanlaya Sookkhee)***

บทคัดย่อ

ความเชื่อมั่นต่อสมุนไพรไทยที่สามารถรักษาอาการเจ็บป่วยตามที่ใช้งานต้องการ เป็นปัจจัยสำคัญในการตัดสินใจเลือกใช้สมุนไพรไทย เพื่อเป็นการส่งเสริมและสร้างความมั่นใจในการใช้สมุนไพรไทยให้ถูกต้องตามสรรพคุณและเหมาะสมกับอาการเจ็บป่วย งานวิจัยนี้จึงได้นำเสนอถึงกระบวนการวิเคราะห์ข้อมูลที่เกี่ยวข้องกับสมุนไพรที่เผยแพร่อยู่บนเว็บไซต์ภาษาไทย ซึ่งเป็นข้อมูลที่มีลักษณะแบบไม่มีโครงสร้าง โดยการประยุกต์ใช้เทคนิคการทำเหมืองข้อความเพื่อประมวลผลข้อมูลให้มีลักษณะแบบมีโครงสร้าง และวิเคราะห์ข้อมูลโดยการประยุกต์ใช้เทคนิคการค้นหากฎความสัมพันธ์ เพื่อค้นหารูปแบบความสัมพันธ์ระหว่างอาการเจ็บป่วยและสมุนไพรที่ตอบสนองการรักษา โดยผู้วิจัยใช้ข้อมูลจากเว็บไซต์ภาษาไทย 40 เว็บไซต์ และใช้คำสำคัญอาการ ความเจ็บป่วยจำนวน 379 อาการ และชื่อสมุนไพรจำนวน 994 ชนิด ผู้วิจัยคัดกรองเอกสารจากแหล่งข้อมูลเว็บไซต์ภาษาไทยทำให้ได้เอกสารที่เกี่ยวข้องทั้งสิ้น 31,452 เอกสาร โดยผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลด้วยเทคนิคดังกล่าวถูกนำเสนอในสองส่วนคือ 1) ผลของการวิเคราะห์ข้อมูลเชิงสารสนเทศซึ่งถูกนำเสนอออกมาในรูปแบบความถี่ที่ปรากฏร่วมกันของคำสำคัญ 2) ผลของการค้นหากฎความสัมพันธ์ระหว่างคำสำคัญอาการเจ็บป่วย และชื่อสมุนไพร โดยในกระบวนการสร้างกฎความสัมพันธ์ผู้วิจัยกำหนดให้ค่า min-support และค่า min-confidence มีค่าเท่ากับ 50% และ 90% ตามลำดับ โดยผลของการศึกษานี้จะถูกใช้เป็นข้อมูลป้อนเข้าสำหรับการพัฒนาระบบสนับสนุนการตัดสินใจ

* อาจารย์ ประจำสาขาวิชาวิทยาการคอมพิวเตอร์ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ
Teacher Department of Computer Science, Faculty of Liberal Arts and Science, Rajabathat Sisaket University. , Email phisan.s@sskru.ac.th, Tel 089-9697483

** อาจารย์ ดร. ประจำสาขาวิชาวิทยาการคอมพิวเตอร์ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ
Teacher Dr. Department of Computer Science, Faculty of Liberal Arts and Science, Rajabathat Sisaket University. , jedsada.s@sskru.ac.th, 098-0970059

*** อาจารย์ ประจำสาขาวิชาคณิตศาสตร์ คณะครุศาสตร์ มหาวิทยาลัยราชภัฏศรีสะเกษ
Teacher Department of Mathematics Education, Faculty of Education, Rajabathat Sisaket University. , sirikalaya.s@sskru.ac.th, 084-0994144

โดยระบบสนับสนุนการตัดสินใจ ที่ได้ทำการพัฒนาขึ้นมีผลของการประเมินประสิทธิภาพของระบบ โดยผู้ใช้ในกลุ่มตัวอย่างเป็นผู้ประเมิน ผลการประเมินประสิทธิภาพของระบบในภาพรวม มีค่าเฉลี่ยเท่ากับ 4.17 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.49 โดยด้านที่มีค่าเฉลี่ยมากที่สุด คือ ด้านการใช้งานระบบ โดยมีค่าเฉลี่ยเท่ากับ 4.23 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.57 รองลงมา คือด้านการออกแบบระบบ โดยมีค่าเฉลี่ยเท่ากับ 4.16 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.53 และค่าเฉลี่ยด้านที่น้อยที่สุดคือ ด้านเนื้อหาที่น่าสนใจ โดยมีค่าเฉลี่ยเท่ากับ 4.13 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.49 ตามลำดับ

คำสำคัญ : การทำเหมืองข้อความ, กฎความสัมพันธ์, สมุนไพรรักษา, อาการเจ็บป่วย, ระบบสนับสนุนการตัดสินใจ

Abstract

The confidence of users on herbal medicines that can treat symptoms is a major factor to decide to use herbal medicine. To promote and increase confidence in using herbal medicine that corresponding its effect to symptoms, This research proposed a process to analyze data about herbs that published on Thai language web sites, and these original data sources are unstructured data e.g. texts and articles. Text mining is used to process unstructured data to become structured data. The association rules technique is used to analyze and to find association patterns between symptoms and herbs from the data set. Original data sources are collected from 40 Thai language web sites. The major keywords are 379 Thai words referring to symptoms and 994 words referring to herbs and collected documents from the original source from the Thai language website, 31,452 totals. The results of the study are presented as two parts: (1) The results of the analysis of the information that is presented in the pattern of frequency that appears together. (2) The results of finding the association rules between symptoms and herbs. In the generation of association rules, we set min-support value at 50% and min-confidence value at 90% and filtered only rules that we are interested in. The results of this study will be used as inputs for the development of decision support systems further.

The developed decision support system was evaluated for effectiveness by the sample users. The overall system performance evaluation results show that the mean was 4.17. The standard deviation was 0.49. The most significant aspect was the system utilization. The mean was 4.23, standard deviation was 0.57, followed by system design. The average was 4.16, the standard deviation was 0.53, and the smallest mean was 4.13, standard deviation was 0.49, respectively.

Keywords: Text Mining, Association Rules, Herbs, Symptoms, Decision Support System

บทนำ

ข้อมูลต่าง ๆ ที่ถูกเผยแพร่สู่สาธารณะผ่านระบบอินเทอร์เน็ตในรูปแบบของเว็บเพจและเว็บไซต์กลายเป็นแหล่งข้อมูลขนาดใหญ่ที่บุคคลทั่วไปสามารถเข้าถึงข้อมูลได้โดยง่าย และมีทรัพยากรสารสนเทศที่มีคุณค่ามหาศาล และพร้อมที่จะแปรสภาพให้มีรรถประโยชน์ต่อการสร้างองค์ความรู้ในด้านต่าง ๆ อย่างหลากหลาย โดยการเติบโตของสารสนเทศในรูปแบบข้อความ (Textual Information) มีอัตราที่สูงขึ้น ทั้งจากเครือข่ายอินเทอร์เน็ตและเอกสารที่องค์กรจัดเก็บไว้ และแม้ว่าเทคโนโลยีการสืบค้น (Search Engine) จะมีขีดความสามารถสูงขึ้น โดยสามารถค้นหาคำสำคัญ (Keywords) และให้ผลลัพธ์อย่างรวดเร็ว แต่วิธีดังกล่าวก็ยังไม่สามารถสรุปความ ประมวลความหมาย หรือความสัมพันธ์ของคำได้อย่างตรงประเด็น ทำให้การสืบค้นเอกสารไม่ตรงกับความต้องการของผู้ใช้ (วรรณวิภา วงศ์วิไลสกุล, 2556, น. 158) ดังนั้นจะเห็นได้ว่าปัญหาที่สำคัญอย่างยิ่งในยุคที่ข้อมูล มีความหลากหลายและเติบโตขึ้นอย่างต่อเนื่อง คือการไม่สามารถเชื่อมโยงข้อมูลที่หลากหลายเข้าด้วยกัน และไม่สามารถสกัดองค์ความรู้จากข้อมูลที่มีจำนวนมหาศาลเพื่อให้ได้องค์ความรู้ที่แท้จริงและตอบสนองต่อความต้องการในการใช้งานได้อย่างตรงประเด็น

ความรู้ด้านสมุนไพร อันเกิดจากการคัดสรร สังเคราะห์ความรู้จากหมอพื้นบ้าน การแพทย์แผนไทย และการแพทย์ทางเลือก เป็นข้อมูลความรู้อีกประเภทหนึ่งที่ถูกเรียบเรียงและเผยแพร่ผ่านระบบอินเทอร์เน็ตเป็นจำนวนมากและมีแนวโน้มที่เพิ่มมากขึ้นเรื่อย ๆ ซึ่งเป็นผลสืบเนื่องมาจากการผลักดันด้านนโยบายที่ถูกระบุไว้ในแผนยุทธศาสตร์ชาติ การพัฒนาภูมิปัญญาไทย สุขภาพวิถีไทย ฉบับที่ 2 พ.ศ. 2555 – 2559 ระบุให้มีการเร่งจัดทำข้อมูลดิจิทัลภูมิปัญญาการแพทย์พื้นบ้าน และการแพทย์แผนไทยของประเทศไทย โดยมีการเชื่อมโยงระบบข้อมูลส่วนกลางกับชุมชนท้องถิ่น (สำนักข้อมูลและประเมินผล กรมพัฒนาการแพทย์แผนไทยและการแพทย์ทางเลือก กระทรวงสาธารณสุข, 2556, น. 27) โดยจากการส่งเสริม ผลักดันจากนโยบายดังกล่าว จึงเห็นได้ว่าทั้งหน่วยงานภาครัฐ และเอกชนได้มีการศึกษาวิจัย โดยการสำรวจ รวบรวมข้อมูลสมุนไพรจากตำรายาการแพทย์แผนไทย หรือข้อมูลจากคำบอกเล่าของหมอพื้นบ้าน และนำมาเผยแพร่สู่สาธารณะในรูปแบบของเว็บเพจจำนวนมาก ตลอดจนกระแสนของการดูแลตนเองด้วยวิถีธรรมชาติบำบัดเริ่มได้รับความนิยมเพิ่มมากขึ้นในสังคมไทยปัจจุบัน ผู้คนให้ความสนใจและมีความตื่นตัวในการหาความรู้จากการใช้ยาสมุนไพรเพื่อนำมาใช้ในการดูแลสุขภาพด้วยตนเองแทนการพึ่งพาจากระบบการแพทย์สมัยใหม่เพียงอย่างเดียว ซึ่งการสืบค้นข้อมูลผ่านระบบเครือข่ายอินเทอร์เน็ตนั้น ถือเป็นช่องทางในการค้นหาความรู้ด้านการดูแลสุขภาพที่ได้รับความนิยมสูงมากที่สุดสำหรับคนไทยในยุคปัจจุบัน เนื่องจากเป็นช่องทางที่ใช้งานได้ง่าย รวดเร็ว และเป็นแหล่งที่มีข้อมูลหลากหลาย และจำนวนมหาศาล

อย่างไรก็ตามการสืบค้นข้อมูลด้านการดูแลสุขภาพด้วยสมุนไพรผ่านระบบอินเทอร์เน็ตยังมีปัญหาอยู่หลายประการ เช่น ข้อมูลมีจำนวนมากและแหล่งที่มาหลากหลาย ข้อมูลด้านสรรพคุณ วิธีการใช้ การสนองตอบต่ออาการเจ็บป่วย บางครั้งอาจมีข้อขัดแย้งกัน ทำให้เกิดความสับสนและความเข้าใจผิดในการใช้สมุนไพรได้ เพราะเนื่องด้วยกระบวนการรวบรวมข้อมูล และสรรพคุณของสมุนไพรนั้น ผู้จัดทำและเผยแพร่อาจได้ข้อมูลมาจากช่องทางที่หลากหลาย ไม่ว่าจะเป็นเอกสารตำราสมัยใหม่ที่มีการพิสูจน์และยืนยันข้อมูลสรรพคุณตามหลักเภสัชวิทยา หรืออาจได้จากการบอกเล่าของผู้รู้ ผู้ใช้งาน หรือหมอพื้นบ้าน ซึ่งเป็นการถ่ายทอดด้วยคำพูดต่อ ๆ กันมา และด้วยเทคโนโลยีการเผยแพร่ข้อมูลในระบบอินเทอร์เน็ตซึ่งเป็นยุคของเว็บ 2.0 ซึ่งมีเทคโนโลยีที่

ช่วยให้ผู้ใช้งานทั่วไปสามารถสร้างเนื้อหาของตนเอง และเผยแพร่ผ่านเว็บไซต์ได้อย่างง่ายดาย เช่น เว็บบล็อก เว็บบอร์ด หรือสารานุกรมออนไลน์ เป็นต้น โดยเนื้อหาด้านสมุนไพรที่ถูกเผยแพร่ออกไปนี้อาจถ่ายทอดออกไป โดยที่ไม่ได้รับการตรวจสอบใด ๆ อีกทั้งการมีข้อมูลด้านสมุนไพรจำนวนมากอาศัยอยู่ในระบบอินเทอร์เน็ต ซึ่งมาจาก แหล่งข้อมูลที่หลากหลายอาจมีการนำเสนอข้อมูลสรรพคุณที่ขัดแย้งกัน หรือไม่ครบถ้วน การสืบค้นโดยทั่วไป ผ่านระบบเสิร์ชเอ็นจินธรรมดาอาจทำให้ได้ข้อมูลคำตอบเพียงผิวเผิน หรือไม่ตรงกับความต้องการของผู้สืบค้น ทำให้ผู้สืบค้นจำเป็นต้องศึกษาข้อมูลเชิงเปรียบเทียบจากผลลัพธ์ที่ได้มากกว่า 1 ผลลัพธ์ขึ้นไปด้วยตนเอง ซึ่งเป็นการเสียเวลาและทำให้เกิดความไม่แน่นอนในการตัดสินใจเพื่อเลือกใช้สมุนไพร

เทคนิคการทำเหมืองข้อความเป็นเทคนิคการค้นหาคำรู้ใหม่จากข้อมูลประเภทข้อความที่มีปริมาณ มากโดยอัตโนมัติ โดยการสกัดคำ คำนหารูปแบบ และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความเอกสาร เพื่อให้เกิด ความหมายและสามารถนำมาใช้ประโยชน์ได้ โดยสถาปัตยกรรมระบบในการค้นหาคำรู้ในฐานข้อมูลประเภท ข้อความ ประกอบด้วย ขั้นตอนการเตรียมข้อมูล ขั้นตอนการค้นหากฎความสัมพันธ์และขั้นตอนการนำเสนอ ผลลัพธ์ที่ชัดเจน ส่วนประโยชน์ของเหมืองข้อความ คือการทำให้สารสนเทศในรูปแบบข้อความ สามารถเข้าถึง วิเคราะห์ และประมวลผลภายในเวลาอันรวดเร็ว โดยผู้ใช้สามารถเข้าใจและใช้ประโยชน์จากเอกสารที่จัดเก็บ ไว้เป็นจำนวนมากอย่างมีประสิทธิภาพ (วรรณวิภา วงศ์วิไลสกุล, 2556) โดยที่ผ่านมามีงานวิจัยหลากหลายได้นำ เทคนิคเหมืองข้อความมาประยุกต์ใช้เพื่อค้นหาข้อสรุปเชิงสารสนเทศจากชุดเอกสารรูปแบบต่าง ๆ เช่น Mahgoub H. (2006) ได้ทำการพัฒนาระบบ EART (Extract Association Rules from Text) ซึ่งใช้เพื่อค้นหา กฎความสัมพันธ์จากชุดของเอกสารที่ไม่มีโครงสร้าง และถูกใช้เพื่อการค้นหาคำความสัมพันธ์จากชุดข้อมูลซึ่งเป็น เอกสารทางวิทยาศาสตร์จาก MEDLINE ที่เกี่ยวข้องกับการระบาดของเชื้อไวรัส H5N1 และในอีก 2 ปีต่อมา Mahgoub H. และคณะ (2008) เสนอการประยุกต์ใช้ระบบ EART เพื่อค้นหากฎความสัมพันธ์ระหว่างคำสำคัญ จากชุดข้อมูลที่เป็นเว็บไซต์ข่าว ที่เกี่ยวข้องกับการระบาดของเชื้อโรคไขหวัดนก หรือ Al-Daihanian SM. และ Abrahamsb A. (2016) เสนอการประยุกต์ใช้เหมืองข้อความสำหรับวิเคราะห์ข้อมูลการทวีต (Tweets) ของสำนักงานห้องสมุดในสถาบันการศึกษาระดับอุดมศึกษา เพื่อค้นหาข้อสรุปเชิงสารสนเทศที่ใช้เพื่อการทำ ความเข้าใจข้อมูลทางสังคมรวมของห้องสมุดสถาบันอุดมศึกษาเพื่อช่วยในการตัดสินใจและการวางแผนเชิง กลยุทธ์สำหรับการเผยแพร่ประชาสัมพันธ์แก่ผู้มีพระคุณและการตลาดของบริการ

ตลอดจนการวิจัยเพื่อการค้นหาคำรู้จากตำรายาสมุนไพรก็มีการประยุกต์ใช้เทคนิค เหมืองข้อความเช่นกัน เช่น Kang JH. และคณะ (2012) นำเสนอการใช้เหมืองข้อมูลเพื่อการวิเคราะห์และ อธิบายความสัมพันธ์ระหว่างสมุนไพรและประสิทธิภาพในการเยียวยารักษา โดยใช้ชุดข้อมูลต้นฉบับจากตำรา การแพทย์แผนตะวันออกของประเทศเกาหลี, He Y. และคณะ (2012) เสนอการใช้การค้นหากฎความสัมพันธ์ เพื่อค้นหารูปแบบของตำรายาสมุนไพรจีนที่ใช้ในการรักษาและป้องกันมะเร็งเต้านม โดยในการศึกษาได้ใช้ข้อมูล สมุนไพรที่ถูกบันทึกไว้ในฐานข้อมูลตำรายาของจีนจำนวน 3 ฐานข้อมูล, Marir F. และคณะ (2016) เสนอผล ของการใช้เหมืองข้อความเพื่อวิเคราะห์ข้อมูลโรคเบาหวานที่เผยแพร่อยู่ในสื่อสังคมออนไลน์ เว็บบล็อก กระตุสนทนา และบทความวิจัยจากฐานข้อมูล MEDLINE หรือ Chatree J. และคณะ (2016) นำเสนอวิธีการใช้ เหมืองข้อความเพื่อสกัดสารสนเทศจากข้อมูลสรรพคุณสมุนไพรต่าง ๆ ที่เผยแพร่อยู่บนเว็บไซต์ภาษาไทย

และใช้การค้นหากฎความสัมพันธ์เพื่อวิเคราะห์ความสัมพันธ์ระหว่างอาการเจ็บป่วยและสมุนไพรที่มีสรรพคุณในการรักษาอาการนั้น ๆ

ดังนั้นงานวิจัยนี้จึงได้ทำการประยุกต์วิธีการพัฒนาระบบสนับสนุนการตัดสินใจ ด้วยการผสมผสานเทคนิคของการสืบค้นข้อมูลเข้ากับการประมวลผลภาษาธรรมชาติ ทำให้สามารถสรุปความคิดเห็นที่หลากหลายบนเครือข่ายอินเทอร์เน็ตให้สามารถเข้าใจได้ง่ายยิ่งขึ้น หรือสกัดเป็นองค์ความรู้ที่น่าเชื่อถือมากยิ่งขึ้น โดยใช้เทคนิคเหมืองข้อความเพื่อการค้นหากฎความสัมพันธ์ของอาการเจ็บป่วยและสมุนไพรที่มีสรรพคุณตอบสนองต่ออาการนั้น ๆ ซึ่งช่วยให้สามารถค้นพบความรู้ที่แฝงอยู่ในข้อมูลจำนวนมากบนระบบอินเทอร์เน็ต และนำความสัมพันธ์ของคำที่เป็นอาการเจ็บป่วยและสมุนไพรที่ปรากฏอยู่ด้วยกันมาสร้างเป็นกฎความสัมพันธ์ และประยุกต์ออกมาในรูปแบบของระบบสนับสนุนการตัดสินใจ เพื่อให้ผู้ใช้งานสามารถสืบค้นข้อมูลที่นำไปสู่ข้อสรุปเกี่ยวกับการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยได้อย่างรวดเร็ว อันจะส่งผลต่อการตัดสินใจในการเลือกใช้ และสร้างความเชื่อมั่นในการใช้สมุนไพรเพื่อการรักษาหรือดูแลสุขภาพต่อไป

วัตถุประสงค์การวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ดังนี้

1. เพื่อพัฒนาระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยโดยเทคนิคเหมืองข้อความ
2. เพื่อประเมินประสิทธิภาพของระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยโดยเทคนิคเหมืองข้อความ

ขอบเขตการวิจัย

1. การวิจัยนี้เป็นการวิจัยเพื่อพัฒนาระบบโดยใช้กระบวนการและเทคนิคการทำเหมืองข้อความเพื่อพัฒนาระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วย ซึ่งประกอบไปด้วยขั้นตอน การเตรียมข้อมูล โดยการคัดเลือกข้อมูล การกรองข้อมูล การตัดคำภาษาไทย การกำจัดคำหยุด กระบวนการเลือกชุดคำ และเมื่อผ่านกระบวนการจัดเตรียมจนได้ข้อมูลที่มีความสมบูรณ์แล้ว คณะผู้วิจัยจึงนำข้อมูลดังกล่าวมาทำการศึกษาเพื่อหาความสัมพันธ์ของคำด้วยเทคนิคการค้นหากฎความสัมพันธ์ ซึ่งจะทำได้กฎความสัมพันธ์ที่แสดงถึงความสัมพันธ์ของอาการเจ็บป่วยและสมุนไพรที่มีสรรพคุณช่วยบำบัดอาการเจ็บป่วยทางร่างกาย
2. การวิจัยนี้ให้ความสนใจและพิจารณาเฉพาะข้อมูลความรู้เกี่ยวกับสมุนไพรที่เผยแพร่อยู่ในระบบอินเทอร์เน็ตซึ่งอยู่ในรูปแบบของเอกสารไฮเปอร์เท็กซ์ หรือเว็บเพจเป็นหลัก ซึ่งพิจารณาคัดเลือกเฉพาะเว็บไซต์ภาษาไทยเท่านั้น โดยระยะเวลาที่จัดเก็บข้อมูลจากแหล่งข้อมูลต้นฉบับอยู่ในช่วงเวลาตั้งแต่ 15 ธันวาคม 2558 ถึง 5 มกราคม 2559

3. กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ ประกอบด้วย 1) นักศึกษาสาขาวิชาสาธารณสุขชุมชน มหาวิทยาลัยราชภัฏศรีสะเกษ ที่ผ่านการลงทะเบียนเรียนในรายวิชาการแพทย์แผนไทย จำนวน 50 คน 2) อาจารย์สาขาวิชาสาธารณสุขชุมชน มหาวิทยาลัยราชภัฏศรีสะเกษ จำนวน 3 คน 3) เจ้าหน้าที่สาธารณสุข กลุ่มงานคุ้มครองผู้บริโภคและเภสัชสาธารณสุข งานแพทย์แผนไทย สำนักงานสาธารณสุขจังหวัดศรีสะเกษ จำนวน 3 คน รวมทั้งสิ้น 56 คน โดยทำการเลือกกลุ่มตัวอย่างแบบเจาะจง (Purposive Sampling)
4. รูปแบบการทำงานของระบบสนับสนุนการตัดสินใจในการวิจัยนี้เป็นแบบเว็บเบสแอปพลิเคชัน

วิธีดำเนินการวิจัย

ผู้วิจัยได้ดำเนินการตามขั้นตอนดังนี้

1) การคัดเลือกแหล่งข้อมูล (Data Source Selection)

คณะผู้วิจัยดำเนินการรวบรวมข้อมูลตามขอบเขตการวิจัยด้านข้อมูลที่ได้ตั้งไว้โดยพิจารณาแหล่งข้อมูลที่จะนำมาใช้ตามหลักเกณฑ์การคัดกรองแหล่งข้อมูลดังต่อไปนี้ 1) เป็นข้อมูลประเภทข้อความที่เกี่ยวข้องกับสมุนไพรไทย 2) สามารถเข้าถึงข้อมูลแบบออนไลน์ผ่านเครือข่ายอินเทอร์เน็ต 3) เป็นข้อมูลเนื้อหาที่เผยแพร่บนเว็บไซต์ภาษาไทย 4) ลักษณะของการเผยแพร่ข้อมูลบนเว็บไซต์เป็นในลักษณะให้ความรู้แก่สาธารณะมิใช่การโฆษณาประชาสัมพันธ์ผลิตภัณฑ์เพื่อการค้าจากเกณฑ์การคัดกรองแหล่งข้อมูลทั้ง 4 ข้อดังกล่าวประกอบกับการสืบค้นข้อมูลด้านสมุนไพรไทยผ่านระบบอินเทอร์เน็ต ทำให้ได้แหล่งข้อมูลต้นฉบับที่ผู้วิจัยนำมาใช้ในการวิจัยจำนวนทั้งสิ้น 40 เว็บไซต์

2) การเตรียมข้อมูล (Data Preparation)

เมื่อได้แหล่งข้อมูลต้นฉบับสำหรับการวิจัยตามเกณฑ์เป็นที่เรียบร้อยแล้ว กระบวนการต่อไปคือการนำข้อมูลที่อยู่ในรูปของเอกสารไฮเปอร์เท็กซ์ (Hypertext Documents) หรือเว็บเพจเข้ามาสู่กระบวนการเตรียมข้อมูลเพื่อเปลี่ยนจากข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) ให้เป็นข้อมูลที่อยู่ในรูปแบบมีโครงสร้าง (Structured Data) เพื่อให้สามารถนำไปสู่กระบวนการประมวลผล และวิเคราะห์ข้อมูลให้ได้คำตอบ ที่ต้องการตามวัตถุประสงค์ของการวิจัยต่อไป ตลอดจนการพัฒนาโปรแกรมเพื่อดำเนินกระบวนการในแต่ละขั้นตอนของการจัดเตรียมข้อมูลเพื่อการประมวลผล คณะผู้วิจัยใช้โปรแกรมภาษาไพทอน (Python) เวอร์ชัน 2.7 และภาษาจาวา (Java) เวอร์ชัน 1.7 เป็นภาษาโปรแกรมหลักในการพัฒนาโปรแกรม ใช้โปรแกรมมายเอสคิวแอล (MySQL) เวอร์ชัน 5 สำหรับบริหารจัดการฐานข้อมูล ตลอดจนใช้ไฟล์เอกสารประเภทซีเอสวี (CSV) ในการสื่อสารข้อมูลระหว่างโปรแกรม และใช้แฟ้มข้อมูลประเภทเพลนเท็กซ์ในการจัดเก็บข้อมูล ประเภทข้อความที่ได้จากการประมวลผลข้อมูลต้นฉบับจากเว็บไซต์ โดยคณะผู้วิจัยสามารถอธิบายขั้นตอนของกระบวนการนำข้อมูลจากแหล่งข้อมูลต้นฉบับเข้ามาสู่กระบวนการเตรียมข้อมูลพร้อมกับการทำงานของโปรแกรมประยุกต์ที่พัฒนาขึ้นมาได้ดังต่อไปนี้

ขั้นที่ 1 ทำดัชนีเว็บเพจโดยใช้เทคนิคเว็บครอว์เลอร์ (Web Crawler) ซึ่งคณะผู้วิจัยพัฒนาขึ้นโดยใช้ สแครปปี้ (Scrapy) เวอร์ชัน 1.0 ซึ่งเป็นเฟรมเวิร์คสำหรับการพัฒนาเว็บครอว์เลอร์

ขั้นที่ 2 ดาวน์โหลดเว็บเพจจากแหล่งข้อมูลต้นฉบับมาประมวลผลตามรายการดัชนี โดยใช้ โมดูล urllib เวอร์ชัน 2 ในภาษาไพทอน

ขั้นที่ 3 ประมวลผลเอกสารไฮเปอร์เท็กซ์เพื่อนำเอาที่เอ็มแอลแท็ก ภาษาจาวาสคริป และสไตร์ชีสออกจากเอกสาร เพื่อให้คงเหลือไว้เฉพาะข้อความ โดยใช้โมดูล BeautifulSOAP เวอร์ชัน 4 ในภาษาไพทอนสำหรับการประมวล และบันทึกผลลัพธ์ไว้ในไฟล์เอกสารประเภทเพลนเท็กซ์

ขั้นที่ 4 ตัดคำภาษาไทยเพื่อแยกคำที่มีความหมายออกจากประโยค โดยใช้โปรแกรมเล็กซ์โต หรือ Thai Lexeme Tokenizer: LexTo (ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ, 2537) ซึ่งเป็นโปรแกรมตัดคำภาษาไทยที่ใช้เทคนิคการเลือกคำที่ยาวที่สุด โดยอ้างอิงคำศัพท์จากพจนานุกรมเล็กซ์ตรอน (พจนานุกรมอิเล็กทรอนิกส์ไทย-อังกฤษ, 2552) จำนวน 40,000 คำ และคณะผู้วิจัยได้เพิ่มคำศัพท์ชื่อสมุนไพรไทยจำนวน 994 คำ และอาการเจ็บป่วยจำนวน 379 คำ โดยเก็บผลการตัดคำภาษาไทยในรูปแบบเพลนเท็กซ์

ขั้นที่ 5 กำจัดคำหยุดภาษาไทย ดำเนินการโดยพัฒนาโปรแกรมเพื่ออ่านไฟล์ที่ผ่านการตัดคำภาษาไทยแล้วเข้ามาประมวลผลเพื่อนำคำหยุดหรือคำที่ไม่สื่อความหมายออกไป โดยการพิจารณาจะใช้ข้อมูลคำหยุดภาษาไทยจำนวน 1,045 คำ

ขั้นที่ 6 คัดกรองเอกสารที่สนใจ โดยเริ่มจากการค้นหาคำรายชื่อสมุนไพรไทยภายในเอกสาร โดยเทียบจากพจนานุกรมชื่อสมุนไพรไทย หากค้นพบอย่างน้อยเพียงหนึ่งชื่อ จะทำการค้นหาต่อว่าในเอกสารนั้นมีคำอาการเจ็บป่วยปรากฏรวมอยู่ด้วยหรือไม่ ถ้าพบตามเงื่อนไขถือว่าเอกสารนั้นผ่านการคัดกรอง

ขั้นที่ 7 สร้างเวกเตอร์ (Word Vector) ซึ่งเวกเตอร์นี้หมายถึงโครงสร้างตารางที่ประกอบไปด้วย แถวที่แทนเอกสารที่พิจารณา และคอลัมน์คือคำที่พิจารณาประกอบด้วยคำที่เราให้ความสนใจ (Hofmann M., & Klinkenberg R., 2014) โดยงานวิจัยนี้คอลัมน์จะหมายถึงคำอาการเจ็บป่วยทั้งหมดที่อยู่ในพจนานุกรมและรายชื่อสมุนไพรไทยที่กำลังพิจารณาและสามารถแสดงโครงสร้างของเวกเตอร์ดังตารางที่ 1 สำหรับการให้น้ำหนักคำสำคัญ (Term Weighting) แก่สมาชิกในแต่ละเวกเตอร์ คณะผู้วิจัยเลือกใช้วิธีการให้น้ำหนักคำสำคัญแบบ TF-IDF (Term Frequency – Inverse Document Frequency) ซึ่งเป็นวิธีที่นิยมใช้มากเพราะทำให้สามารถลดจำนวนคำที่ไม่มีความสำคัญลงให้คงเหลือแต่คำสำคัญที่สามารถเป็นตัวแทนที่ดีของเอกสารอย่างแท้จริง (Mahgoub H., 2008; Chatree J. และคณะ, 2016) โดยวิธีการคำนวณแสดงได้ดังสมการต่อไปนี้

$$w_{ik} = tf_{ik} \times idf_k$$

โดย

$$idf_k = \log\left(\frac{N}{df_k}\right)$$

โดยที่

w_{ik}	คือ	ค่าน้ำหนักของคำ k ในเอกสาร i
tf_{ik}	คือ	ความถี่ของคำ k ที่ปรากฏในเอกสาร i
idf_k	คือ	ค่า Inverse Document Frequency ของคำ k
N	คือ	จำนวนเอกสารทั้งหมดที่นำมาพิจารณา
df_k	คือ	จำนวนเอกสารที่บรรจุคำหรือปรากฏพบคำ k

ตารางที่ 1 โครงสร้างของเวกเตอร์ของสมุนไพร j ที่เกิดจากการประมวลผลเอกสารที่ผ่านการกลั่นกรอง

	Symp ₁	Symp ₂	...	Symp _i	Herb _j
Document ₁	S ₁₁	S ₁₂	...	S _{1i}	h _{1j}
Document ₂	S ₂₁	S ₂₂	...	S _{2i}	h _{2j}
...	...	...	...	...	...
Document _n	S _{n1}	S _{n2}	...	S _{ni}	h _{nj}

จากตารางที่ 1

Symp_i หมายถึง คำแสดงอาการเจ็บป่วยลำดับที่ i

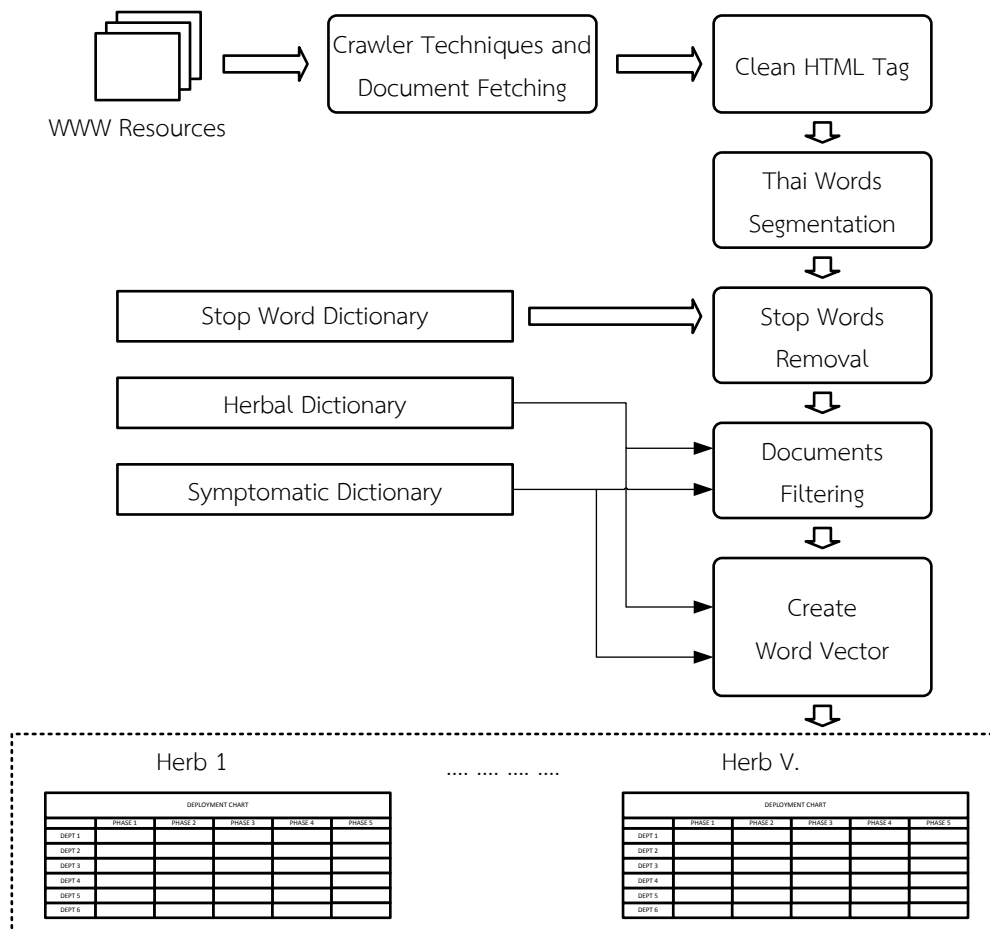
Herb_j หมายถึง ชื่อสมุนไพรไทยลำดับที่ j

Document_n หมายถึง เอกสารลำดับที่ n ที่ค้นพบชื่อสมุนไพร Herb_j

S_{ni} หมายถึง คำนวณน้ำหนักแบบ TF-IDF ของอาการเจ็บป่วยลำดับที่ i ในเอกสารที่ n

h_{nj} หมายถึง คำนวณน้ำหนักแบบ TF-IDF ของสมุนไพรลำดับที่ j ในเอกสารที่ n

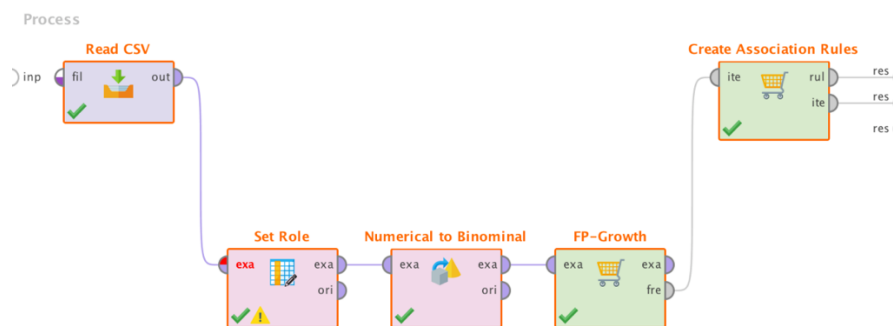
โดยผู้วิจัยสามารถอธิบายกระบวนการจัดเตรียมข้อมูลในขั้นที่ 1 ถึง 7 ที่มีการดำเนินการในงานวิจัยนี้ในรูปแบบของแผนภาพกระบวนการทำงานได้ดังภาพที่ 1



ภาพที่ 1 กระบวนการจัดเตรียมข้อมูลเพื่อการประมวลผล

3) การวิเคราะห์และค้นหากฎความสัมพันธ์

กระบวนการเตรียมข้อมูลจากแหล่งข้อมูลต้นฉบับ ทำให้ได้ผลลัพธ์ที่เรียกว่า เวิร์ดเวกเตอร์ ซึ่งเป็นข้อมูลที่มีโครงสร้างและมีความเหมาะสมในการประมวลผลเพื่อให้ได้สารสนเทศและค้นหากฎความสัมพันธ์ระหว่างสมุนไพรในแต่ละอาการตามวัตถุประสงค์ของการวิจัย โดยกระบวนการค้นหากฎความสัมพันธ์ของข้อมูลคณะผู้วิจัยใช้โปรแกรม RapidMiner Studio เวอร์ชันที่ 7.0 และใช้อัลกอริทึม FP Growth (Hunyadi D., 2011 และ Gadia K. & Bhowmick K., 2015) สำหรับค้นหากฎความสัมพันธ์จากชุดข้อมูล โดยสามารถนำเสนอลำดับของ Operators ของโปรแกรม RapidMiner Studio ที่ใช้ในกระบวนการวิเคราะห์หากฎความสัมพันธ์ทั้งหมดได้ดังภาพที่ 2 ซึ่งในงานวิจัยนี้ได้กำหนดค่าสนับสนุนขั้นต่ำ และค่าความเชื่อมั่นขั้นต่ำที่ทำให้เกิดกฎความสัมพันธ์อยู่ที่ 0.50 และ 0.90 ตามลำดับ



ภาพที่ 2 ลำดับโอเพอเรเตอร์ของโปรแกรม RapidMiner Studio ในกระบวนการค้นหาความสัมพันธ์

กระบวนการสร้างกฎความสัมพันธ์ดำเนินงานโดยการค้นหาความสัมพันธ์ระหว่างอาการเจ็บป่วยและสมุนไพรจากชุดเอกสารที่จำแนกหมวดหมู่ตามรายชื่อสมุนไพร โดยกำหนดให้ $S = \{s_1, s_2, \dots, s_m\}$ เป็นเซตคำสำคัญอาการเจ็บป่วยและ $H = \{h_1, h_2, \dots, h_n\}$ เป็นเซตคำสำคัญสมุนไพร และให้ $d_i h_j$ เป็นตัวแทนเอกสาร i ในหมวดหมู่สมุนไพร h_j ซึ่งภายในจะต้องปรากฏความถี่ที่พบ h_j อย่างน้อย 1 ครั้งและปรากฏความถี่ที่ค้นพบสมาชิกใน S อย่างน้อยหนึ่งตัวภายในเอกสาร และการค้นหาความสัมพันธ์จะกระทำที่ละชุดเอกสารในหมวดหมู่รายชื่อสมุนไพร

ผลลัพธ์ของกระบวนการนี้ทำให้ได้กฎความสัมพันธ์ที่ได้มีรูปแบบหลากหลาย คณะผู้วิจัยทำการคัดกรองกฎความสัมพันธ์ตามรูปแบบที่คณะผู้วิจัยให้ความสนใจ 2 รูปแบบ ซึ่งประกอบไปด้วยรูปแบบที่ 1) $\langle \text{symptom} \rangle \Rightarrow \langle \text{herb} \rangle$ เป็นกฎที่สามารถแสดงความสัมพันธ์ระหว่างอาการเจ็บป่วยหนึ่งอาการกับสมุนไพรหนึ่งรายชื่อ 2) $[\langle \text{symptom} \rangle] \Rightarrow \langle \text{herb} \rangle$ เป็นกฎที่แสดงความสัมพันธ์ระหว่างกลุ่มของอาการเจ็บป่วยและสมุนไพรหนึ่งรายชื่อ หลังจากนั้นคณะผู้วิจัยได้นำชุดของกฎความสัมพันธ์เหล่านี้เป็นข้อมูลป้อนเข้าให้แก่กระบวนการพัฒนาระบบสนับสนุนการตัดสินใจต่อไป

4) การวิเคราะห์และออกแบบระบบ

การวิเคราะห์และออกแบบระบบสนับสนุนการตัดสินใจ คณะผู้วิจัยได้ดำเนินการภายใต้กรอบคุณสมบัติด้านความสามารถของตัวระบบสนับสนุนการตัดสินใจ โดยคณะผู้วิจัยกำหนดความสามารถของระบบสนับสนุนการตัดสินใจ ไว้ดังต่อไปนี้ 1) สามารถในการนำเข้าข้อมูลกฎความสัมพันธ์ 2) สามารถรับข้อมูลคำสืบค้นจากผู้ใช้งาน 3) สามารถแสดงผลการสืบค้นข้อมูลจากกฎความสัมพันธ์ 4) ผู้ใช้งานสามารถแสดงความคิดเห็นของตนเองต่อกฎความสัมพันธ์ที่แสดงในระบบฯ 5) ผู้เชี่ยวชาญสามารถให้คะแนนการยอมรับกฎความสัมพันธ์เพื่อตรวจสอบข้อความถูกต้องของกฎความสัมพันธ์ 6) สามารถแสดงสารสนเทศทางสถิติ ของคำค้นจากการใช้งานระบบฯ

การดำเนินการในขั้นตอนนี้ คณะผู้วิจัยสามารถแสดงรายละเอียดการออกแบบการทำงานของระบบสนับสนุนการตัดสินใจในภาพรวมได้โดยใช้แผนภาพกระแสข้อมูลระดับสูงสุด (Context Diagram) ในการแสดงเส้นทางข้อมูลเข้าและออกจากแหล่งข้อมูลที่มีผลกระทบต่อระบบดังภาพที่ 3 และใช้แผนภาพกระแสข้อมูลระดับที่ 1 (Data Flow Diagram Level 1) แสดงกระบวนการทำงานหลักของระบบและข้อมูลที่เข้าและออกจากกระบวนการทำงานต่าง ๆ ดังภาพที่ 4

จากภาพที่ 4 แสดงให้เห็นถึงกระบวนการทำงานของระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยโดยเทคนิคเหมืองข้อความ ซึ่งมีการทำงานทั้งสิ้น 8 กระบวนการ คือ 1) ตรวจสอบสิทธิในการเข้าใช้งานระบบด้วยการป้อนข้อมูลชื่อและรหัสเข้าใช้งาน 2) บันทึกข้อมูลกฎความสัมพันธ์โดยผู้ดูแลระบบ 3) ผู้ดูแลระบบให้สิทธิสมาชิกที่ลงทะเบียนใช้งานในระบบเพื่อทำหน้าที่เป็นผู้เชี่ยวชาญในการตรวจสอบและให้คำแนะนำการยอมรับกฎความสัมพันธ์ 4) บันทึกข้อมูลสมัครผู้ใช้งานประเภทสมาชิก 5) บันทึกข้อมูลข้อคิดเห็นต่อกฎความสัมพันธ์โดยสมาชิกที่ได้รับสิทธิ 6) บันทึกข้อมูลคำแนะนำการยอมรับกฎความสัมพันธ์ 7) การสืบค้นกฎความสัมพันธ์ 8) ประมวลผลและแสดงผลสารสนเทศทางสถิติของคำค้นต่อผู้ใช้งาน

ผลการวิจัย

การวิจัยเรื่องการพัฒนากระบวนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยโดยเทคนิคเหมืองข้อความ สามารถนำเสนอผลวิจัยออกเป็น 2 ส่วน ดังต่อไปนี้

1) ผลของการพัฒนาระบบสนับสนุนการตัดสินใจ

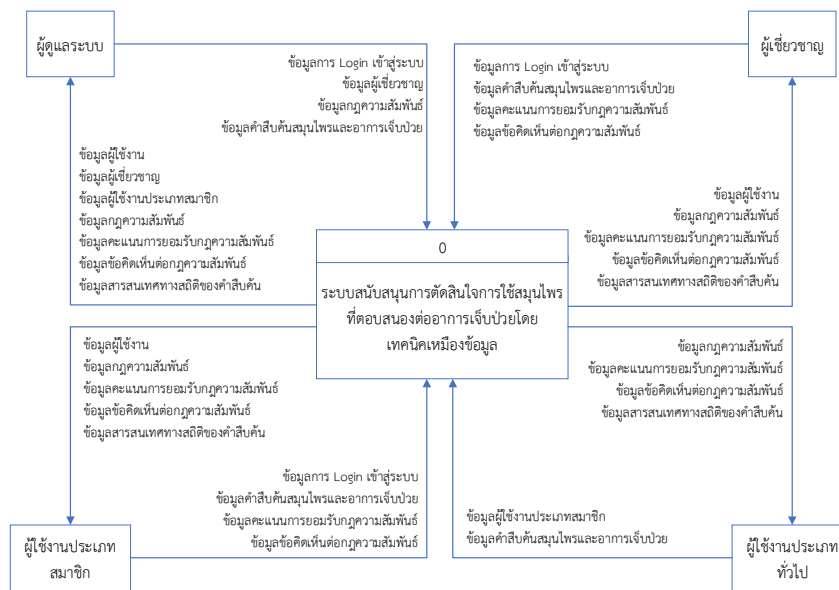
เมื่อคณะผู้วิจัยได้กฎความสัมพันธ์จากกระบวนการค้นหาความสัมพันธ์แล้ว จึงได้ทำการพัฒนาระบบสนับสนุนการตัดสินใจ ตามที่ได้วิเคราะห์และออกแบบระบบไว้และติดตั้งระบบไว้ในเครื่องคอมพิวเตอร์แม่ข่าย (Computer Server) เพื่อให้ผู้ใช้งานสามารถเข้าถึงผ่าน <http://www.sisaket-wisdoms.net> ซึ่งเป็นการต่อยอดจาก “ระบบสารสนเทศสารสนเทศภูมิปัญญาท้องถิ่นเพื่อการดูแลสุขภาพ” ที่พัฒนาโดยพิศาล สุขชี (พิศาล สุขชี, 2558) และสามารถนำเสนอผลการพัฒนาระบบสนับสนุนการตัดสินใจ ได้ดังต่อไปนี้

การนำเข้าข้อมูลกฎความสัมพันธ์

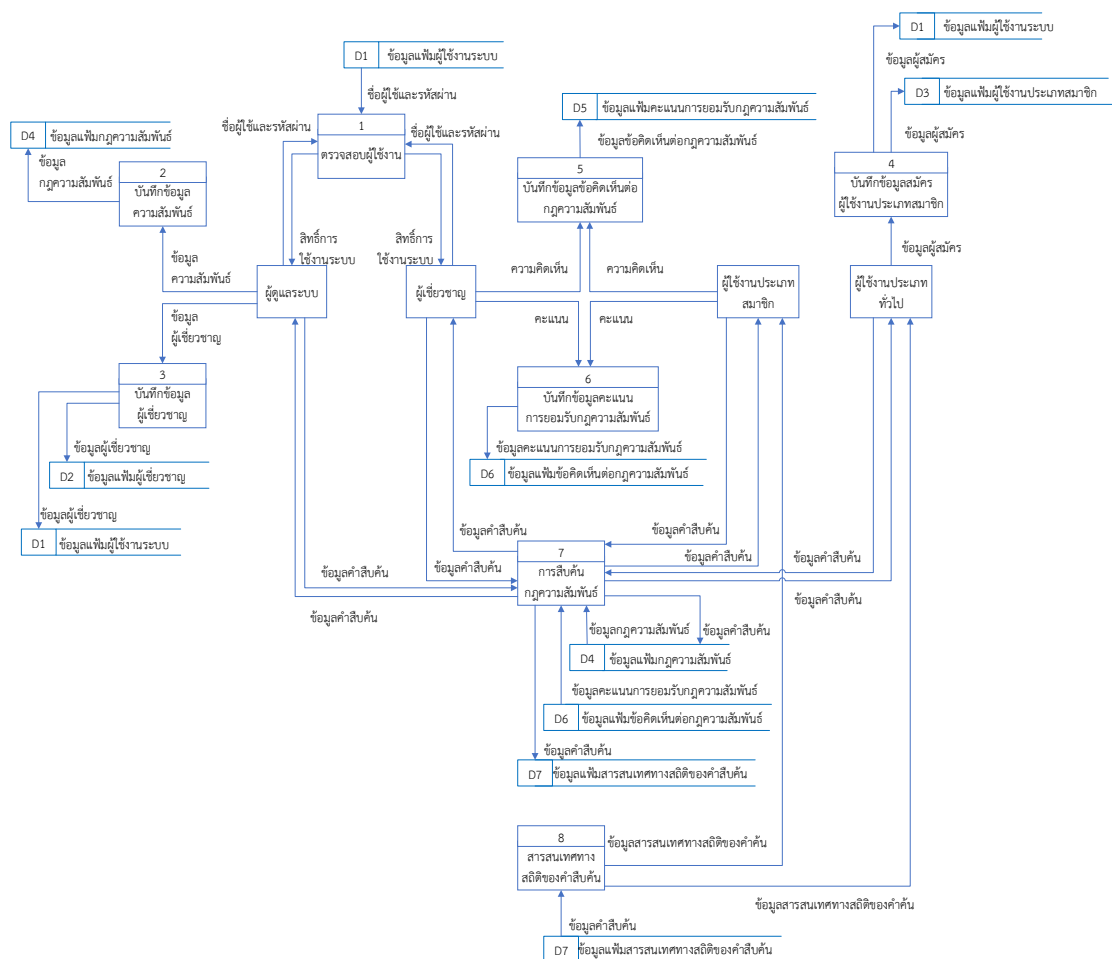
ระบบสนับสนุนการตัดสินใจ ที่พัฒนาขึ้นมีความสามารถในการนำเข้าข้อมูลกฎความสัมพันธ์ในรูปแบบไฟล์ประเภทซีเอสวี โดยดำเนินการกระบวนการอ่านไฟล์ซีเอสวี และทำการบันทึกข้อมูลเข้าสู่ระบบฐานข้อมูล และทำการนำเสนอข้อมูลกฎความสัมพันธ์ผ่านหน้าจอของระบบสนับสนุนการตัดสินใจ ในรูปแบบของตารางโดยเรียงลำดับตามค่าสนับสนุน และค่าความเชื่อมั่นของกฎแต่ละกฎจากมากไปหาน้อยดังภาพที่ 6 ซึ่งผู้ใช้งานระบบสนับสนุนการตัดสินใจ สามารถค้นหาข้อมูลกฎที่ตนเองต้องการโดยระบุค่าสำคัญลงในช่องทางการค้นหา 4 ช่องทางคือ อาการเจ็บป่วย, สมุนไพร, ค่าสนับสนุน และค่าความเชื่อมั่น

การสืบค้นข้อมูลจากผู้ใช้งาน

ระบบสนับสนุนการตัดสินใจ สามารถรองรับการสืบค้นข้อมูลจากผู้ใช้งานได้อย่างหลากหลายช่องทางเพื่อสร้างความสะดวกต่อการใช้ข้อมูลที่อยู่ในระบบไปประกอบการตัดสินใจ โดยผู้ใช้งานระบบสนับสนุนการตัดสินใจ สามารถเข้าถึงการสืบค้นรูปแบบต่าง ๆ ผ่านหน้าจอหลัก ซึ่งประกอบไปด้วยเมนู 1) ตัดสินใจเลือกใช้จากสมุนไพร 2) ตัดสินใจเลือกใช้จากอาการเจ็บป่วย 3) ค้นหาสมุนไพรที่สัมพันธ์กับอาการเจ็บป่วย 4) สารสนเทศจากการวิเคราะห์ชุดเอกสาร และ 5) ข้อมูลสถิติการใช้งานระบบสนับสนุนการตัดสินใจ



ภาพที่ 3 แผนภาพกระแสข้อมูลระดับสูงสุด



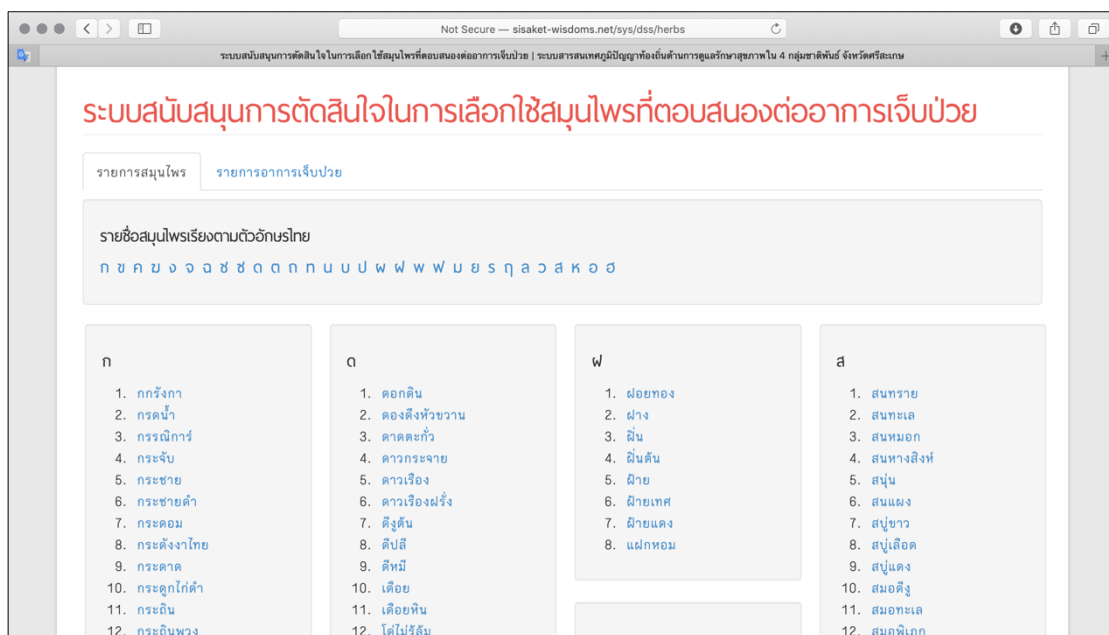
ภาพที่ 4 แผนภาพกระแสข้อมูลระดับที่ 1

เมนู “ตัดสินใจเลือกใช้จากสมุนไพร” และ “ตัดสินใจเลือกใช้จากอาการเจ็บป่วย” เมื่อคลิกเลือกที่เมนูแล้วจะปรากฏรายการสมุนไพรหรืออาการเจ็บป่วยเรียงตามตัวอักษรไทย ก-ฮ จำแนกหมวดหมู่ตามตัวอักษรดังแสดงตัวอย่างในภาพที่ 5 และผู้ใช้งานระบบสนับสนุนการตัดสินใจ สามารถคลิกเลือกเพื่อดูข้อมูลจากรายการอาการเจ็บป่วย หรือชื่อสมุนไพรแต่ละรายการได้

เมนู “ค้นหาสมุนไพรที่สัมพันธ์กับอาการเจ็บป่วย” ผู้ใช้สามารถค้นข้อมูลจากกฎความสัมพันธ์ โดยการระบุค่าค้นที่ตนเองต้องการสืบค้นไม่ว่าจะเป็นการระบุอาการเจ็บป่วย หรือรายชื่อสมุนไพรดังภาพที่ 6

การแสดงผลข้อมูลสารสนเทศเพื่อสนับสนุนการตัดสินใจ

เพื่อเป็นการนำเสนอสารสนเทศที่ได้จากการสกัดข้อมูลที่เผยแพร่อยู่ในระบบอินเทอร์เน็ตให้แก่ผู้ใช้งาน ให้สามารถทำความเข้าใจและนำสารสนเทศนั้นไปประกอบการตัดสินใจในการเลือกใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วยที่ต้องการได้ คณะผู้วิจัยจึงได้นำเสนอสารสนเทศในหลายรูปแบบดังแสดงในภาพที่ 7 โดยสารสนเทศเชิงความถี่ที่ค้นพบในหมวดหมู่ของสมุนไพรนั้น ๆ ประกอบไปด้วย 1) ความถี่ของอาการเจ็บป่วยที่ 2) ข้อมูลความถี่ของเอกสารที่ค้นพบสมุนไพร 3) จำนวนคำสำคัญอาการเจ็บป่วยที่ค้นพบ 4) จำนวนกฎความสัมพันธ์ที่ปรากฏ 5) กราฟวงกลมแสดงสัดส่วนความถี่ของคำสำคัญอาการเจ็บป่วย 6) กราฟสัดส่วนคำสำคัญอาการเจ็บป่วยจำแนกตามจำนวนเอกสาร และ 7) กฎความสัมพันธ์



ภาพที่ 5 หน้าจอแสดงรายชื่อสมุนไพรเรียงลำดับและจำแนกตามตัวอักษรไทย ก-ฮ

ค้นหาสมุนไพรที่สัมพันธ์กับอาการเจ็บป่วยจากกฎความสัมพันธ์

อาการเจ็บป่วย: สมุนไพร: ค่าสนับสนุน (มากกว่าหรือเท่ากับ):

ค่าความเชื่อมั่น (มากกว่าหรือเท่ากับ):

ลำดับ	กลุ่มอาการเจ็บป่วย	สมุนไพร	ค่าสนับสนุน	ค่าความเชื่อมั่น	Lift	การยอมรับกฎ	แสดง
1	พยาธิ	กกรังกา	0.20	1.00	1.00	5/5	แสดง
2	ผื่นคัน, คุดทะราด, อักเสบ, มะเร็ง, พยาธิ, กามโรค	มะหาด	0.13	1.00	1.00	0/5	แสดง
3	วัณโรค, คุดทะราด, หนอง, ร้อนใน, มะเร็ง, พยาธิ	มะหาด	0.13	1.00	1.00	0/5	แสดง
4	ผื่นคัน, วัณโรค, คุดทะราด, อักเสบ, หนอง, มะเร็ง, พยาธิ	มะหาด	0.13	1.00	1.00	0/5	แสดง
5	ผื่นคัน, คุดทะราด, อักเสบ, หนอง, มะเร็ง, พยาธิ, กามโรค	มะหาด	0.13	1.00	1.00	0/5	แสดง
6	ผื่นคัน, วัณโรค, อักเสบ, หนอง, ร้อนใน, มะเร็ง, พยาธิ, กามโรค	มะหาด	0.13	1.00	1.00	0/5	แสดง
7	พยาธิ, เป็นไข้	มะหาด	0.12	1.00	1.00	0/5	แสดง
8	พยาธิ, ถ่ายพยาธิ, ขับปัสสาวะ	มะหาด	0.12	1.00	1.00	0/5	แสดง

ภาพที่ 6 หน้าจอสำหรับสืบค้นข้อมูลจากกฎความสัมพันธ์ระหว่างอาการเจ็บป่วยและสมุนไพร

สารสนเทศเกี่ยวกับสมุนไพร "กระชาย"

อาการเจ็บป่วยที่พบในหมวดหมู่สมุนไพร "กระชาย"

เป็นรายการค่าสำคัญอาการเจ็บป่วยที่สามารถค้นพบได้ภายในเอกสารที่มีคำว่า 'กระชาย' ปรากฏร่วมอยู่ด้วย โดยข้อมูลด้านล่างสามารถนำมาใช้เป็นสารสนเทศสำหรับบ่งชี้ว่าอาการเจ็บป่วย และสมุนไพรใดบ้างที่มีความสัมพันธ์ต่อกันและปรากฏร่วมกันบ่อยครั้งในเอกสารที่เผยแพร่อยู่บนระบบอินเทอร์เน็ต...แสดงทั้งหมด

ลำดับ	อาการเจ็บป่วย	ความถี่ที่ปรากฏ	จำนวนเอกสารที่พบ
1	ปวด	1,098	245
2	เบาหวาน	671	121
3	ไต	620	96
4	กรดไหลย้อน	492	129
5	ปวดหัว	472	99
6	หัวใจ	427	119
7	ท้องอืด	426	169
8	อักเสบ	402	224
9	เรื้อรัง	338	183
10	ท้องเสีย	332	196

สารสนเทศความสัมพันธ์ระหว่างอาการเจ็บป่วยและสมุนไพร

จำนวนค่าสำคัญอาการเจ็บป่วย	244 ค่า
จำนวนกฎความสัมพันธ์	396 กฎ
จำนวนเอกสารที่นำมาวิเคราะห์	638 เอกสาร

สัดส่วนความถี่ของค่าสำคัญอาการเจ็บป่วยที่พบ

● ปวด
 ● เบาหวาน
 ● ไต
 ● กรดไหลย้อน
 ● ปวดหัว
 ● หัวใจ
 ● ท้องอืด
 ● อักเสบ
 ● เรื้อรัง
 ● ท้องเสีย
 ● อื่นๆ

ภาพที่ 7 หน้าจอแสดงสารสนเทศเกี่ยวกับสมุนไพรที่ได้จากการประมวลผลข้อมูล

การบันทึกข้อมูลสถิติการใช้งานด้านการสืบค้นข้อมูลของผู้ใช้

ความสามารถของระบบสนับสนุนการตัดสินใจ อีกด้านหนึ่งคือการบันทึกข้อมูลการเข้าชม หรือการบันทึกการสืบค้นข้อมูลในระบบสารสนเทศด้วยคำค้นต่าง ๆ และนำมาแสดงให้ผู้ดูแลระบบ หรือผู้ใช้งานทั่ว ๆ ไปว่ารายชื่อสมุนไพรหรืออาการเจ็บป่วยใดที่ผู้ใช้งานระบบใช้ในการสืบค้นมากที่สุด ซึ่งข้อมูลเหล่านี้ถือเป็นประโยชน์ที่จะสามารถสะท้อนพฤติกรรม หรือความสนใจของผู้ใช้งานส่วนมากที่มีต่อระบบสนับสนุนการตัดสินใจได้

การคัดกรองกฎความสัมพันธ์โดยผู้เชี่ยวชาญ

ผลของการทดลองค้นหาความสัมพันธ์ระหว่างคำสำคัญอาการเจ็บป่วย 379 อาการ ที่สัมพันธ์กับชื่อสมุนไพร 994 ชนิด จากชุดเอกสารจำนวน 31,452 เอกสาร ทำให้ได้กฎความสัมพันธ์เกิดขึ้นเป็นจำนวนมากภายใต้ค่าสนับสนุนขั้นต่ำที่ 0.50 และค่าความเชื่อมั่นขั้นต่ำที่ 0.90 โดยกฎความสัมพันธ์เหล่านี้เกิดขึ้นจากกระบวนการวิเคราะห์ข้อมูลทางคอมพิวเตอร์ที่อ้างอิงจากหลักทางสถิติ แต่การนำกฎความสัมพันธ์เหล่านี้ไปใช้งานยังคงต้องผ่านกระบวนการตรวจสอบซ้ำ และให้การยอมรับโดยผู้เชี่ยวชาญในอีกระดับก่อนระบบสนับสนุนการตัดสินใจ ที่พัฒนาขึ้นนี้จึงได้จัดเตรียมช่องทางสำหรับการให้ “ค่าคะแนนการยอมรับกฎ” ซึ่งเป็นการให้คะแนนกับกฎความสัมพันธ์ในแต่ละกฎโดยคณะผู้เชี่ยวชาญว่ากฎนั้นมีความน่าเชื่อถือมากน้อยเพียงใด

โดยผู้เชี่ยวชาญที่ทำหน้าที่นี้เป็นในลักษณะของอาสาสมัครที่มาจากเจ้าหน้าที่ด้านการแพทย์แผนไทยและอาจารย์ประจำสาขาสาธารณสุขชุมชน เป็นต้น ซึ่งผู้ดูแลระบบจะแต่งตั้งและให้สิทธิ์สมาชิกกลุ่มนี้ให้มีสิทธิ์ในการโหวตคะแนนการยอมรับกฎความสัมพันธ์ตลอดจนให้ข้อเสนอแนะในการใช้งานสมุนไพร โดยค่าคะแนนการยอมรับกฎจะเป็นคะแนนที่อยู่ในช่วง 1 – 5 และกฎความสัมพันธ์ใดที่มีค่าคะแนนการยอมรับกฎที่สูงจะได้รับการจัดแสดงผลลัพธ์ให้แก่ผู้ใช้งานเป็นลำดับต้นก่อนกฎความสัมพันธ์อื่น ๆ ดังแสดงในภาพที่ 6

2) ผลการประเมินประสิทธิภาพระบบสนับสนุนการตัดสินใจ

การประเมินประสิทธิภาพของระบบสนับสนุนการตัดสินใจ จากผู้ใช้งานกลุ่มเป้าหมายโดยใช้แบบประเมินประสิทธิภาพของระบบสารสนเทศฯ เป็นแบบมาตรวัดมาตราส่วนประมาณค่า (Rating Scale) 5 ระดับของลิเคิร์ตสเกล (Likert, 1967) จำนวน 20 ข้อ ประกอบด้วยด้านเนื้อหาที่นำเสนอจำนวน 5 ข้อ ด้านการใช้ประโยชน์จำนวน 6 ข้อ ด้านการใช้งานระบบจำนวน 5 ข้อ และด้านการออกแบบระบบจำนวน 4 ข้อ และคำถามปลายเปิด ในการแสดงข้อคิดเห็นและข้อเสนอแนะเพิ่มเติมจากผู้ใช้งาน คณะผู้วิจัยได้กำหนดเกณฑ์ ในการวิเคราะห์ตามเกณฑ์คุณภาพของเบสต์ (Best, 1981) ใช้สถิติพรรณนา (Descriptive Statistics) ในการอธิบายผลการประเมินประสิทธิภาพ ซึ่งมีรายละเอียดดังต่อไปนี้

ผู้ตอบแบบประเมินประสิทธิภาพระบบสนับสนุนการตัดสินใจ จำนวน 56 คน ส่วนใหญ่เป็นเพศหญิง 53 คน คิดเป็นร้อยละ 94.64 และเป็นเพศชาย 3 คน คิดเป็นร้อยละ 5.36 ตามลำดับ โดยจำแนกออกเป็นนักศึกษา 50 คน คิดเป็นร้อยละ 89.29 อาจารย์ 3 คน คิดเป็นร้อยละ 5.36 และเป็นเจ้าหน้าที่สาธารณสุข 3 คน คิดเป็นร้อยละ 5.36 ตามลำดับ เนื้อหาที่ใช้สำหรับประเมินประสิทธิภาพของระบบสนับสนุน การตัดสินใจ แบ่งกลุ่มในภาพรวมออกเป็น 3 ด้าน โดยสามารถแสดงผลประเมินจำแนกรายด้านได้ดังตารางที่ 2

ตารางที่ 2 ผลการประเมินประสิทธิภาพการใช้งานระบบสนับสนุนการตัดสินใจ ในภาพรวมทั้ง 4 ด้าน

หัวข้อประเมินในภาพรวม 3 ด้าน	ผลการประเมิน		
	X	S.D.	ความหมาย
ด้านเนื้อหาที่น่าสนใจ	4.13	0.49	มาก
ด้านการใช้ประโยชน์	4.14	0.56	มาก
ด้านการใช้งานระบบ	4.23	0.57	มาก
ด้านการออกแบบระบบ	4.16	0.53	มาก
ค่าเฉลี่ยรวมทุกด้าน	4.17	0.49	มาก

จากตารางที่ 2 เมื่อพิจารณาภาพรวม พบว่าผลการประเมินประสิทธิภาพการใช้งานระบบสนับสนุนการตัดสินใจ โดยภาพรวม มีค่าเฉลี่ยเท่ากับ 4.17 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.49 โดยด้านที่มีค่าเฉลี่ยมากที่สุดคือ ด้านการใช้งานระบบ โดยมีค่าเฉลี่ยเท่ากับ 4.23 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.57 รองลงมาคือ ด้านการออกแบบระบบ โดยมีค่าเฉลี่ยเท่ากับ 4.16 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.53 และค่าเฉลี่ยด้านที่น้อยที่สุดคือ ด้านเนื้อหาที่น่าสนใจ โดยมีค่าเฉลี่ยเท่ากับ 4.13 ค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.49 ตามลำดับ

นอกจากการประเมินประสิทธิภาพเชิงปริมาณผ่านแบบประเมินประสิทธิภาพแล้ว คณะผู้วิจัยได้ดำเนินการรวบรวมข้อมูลในเชิงคุณภาพเพื่อรับข้อมูลป้อนกลับในการพัฒนาปรับปรุง โดยรวบรวมข้อมูลผ่านกระบวนการสนทนากลุ่มย่อย (Focus Group) โดยใช้แบบสัมภาษณ์แบบมีโครงสร้าง (Structured Interview) ซึ่งมีผู้เชี่ยวชาญ จำนวน 5 ท่านมาร่วมแลกเปลี่ยนข้อคิดเห็น โดยสามารถสรุปผลได้ดังนี้

1. เนื่องจากอาการเจ็บป่วยมีการจำแนกออกเป็นหลายระดับ เช่น ระยะต้น ระยะกลาง และระยะสุดท้าย ควรวิเคราะห์ให้ลึกลงไปเพื่อป้องกันการใช้สมุนไพรรักษาอาการในแต่ละระยะนั้นก็จะเปลี่ยนไปตามระดับของอาการ
2. ระบบสนับสนุนการตัดสินใจ ควรมีการนำเสนอข้อมูลแบบเบ็ดเสร็จในที่เดียวโดยผู้ใช้งานไม่จำเป็นต้องสืบค้นต่อยอดในแหล่งข้อมูลอื่นอีก
3. ควรทำลิงค์ไปยังแหล่งข้อมูลที่หน้าเชื่อถืออื่น ๆ เพื่อให้ผู้ใช้ที่สนใจสามารถเข้าไปทำการศึกษาได้เองอย่างต่อเนื่อง
4. ระบบสนับสนุนการตัดสินใจ นี้สามารถนำไปใช้เป็นเครื่องมือประกอบการเรียนการสอนของรายวิชาการแพทย์แผนไทย ของสาขาวิชาสาธารณสุขชุมชน โดยมีการมอบหมายให้เด็กทำการสืบค้นข้อมูลสรรพคุณสมุนไพร และทำการแลกเปลี่ยนเรียนรู้ร่วมกันผ่านระบบฯ นี้ได้

สรุปผลและข้อเสนอแนะ

1) สรุปผล

การวิจัยเรื่อง การพัฒนาระบบสนับสนุนการตัดสินใจในการใช้สมุนไพรที่ตอบสนองต่ออาการเจ็บป่วย โดยใช้เทคนิคเหมืองข้อความ เป็นการผสานเทคนิคการสืบค้นข้อมูลเข้ากับการประมวลผลภาษาธรรมชาติ ซึ่งทำให้สามารถสรุปข้อมูลและความคิดเห็นด้านสรรพคุณของสมุนไพรที่มีหลากหลายบนเครือข่ายอินเทอร์เน็ต ให้สามารถเข้าใจได้ง่ายยิ่งขึ้น หรือสกัดเป็นองค์ความรู้ที่น่าเชื่อถือมากยิ่งขึ้น โดยเทคนิคเหมืองข้อความถูกใช้เพื่อประมวลผลข้อมูลที่ไม่มีโครงสร้างเพื่อการสกัดสารสนเทศและค้นหากฎความสัมพันธ์ระหว่างอาการเจ็บป่วย และสรรพคุณของสมุนไพรจากชุดเอกสารภาษาไทยที่มีแหล่งข้อมูลต้นฉบับจากอินเทอร์เน็ต โดยกระบวนการเริ่มต้นจากการคัดเลือกแหล่งข้อมูลต้นฉบับจากเว็บไซต์ภาษาไทย การประมวลผลเอกสารไฮเปอร์เท็กซ์ให้เป็นเพอร์เท็กซ์ การตัดคำภาษาไทย การกำจัดคำหยุด การคัดกรองเอกสารที่สนใจ การแปลงข้อมูลในเอกสารให้อยู่ในรูปแบบข้อมูลที่มีโครงสร้าง การใช้เทคนิคการให้น้ำหนักคำแบบ TF-IDF สำหรับคัดกรองคำสำคัญที่ไม่สามารถเป็นตัวแทนที่ดีของเอกสารออกไป และทำการค้นหากฎความสัมพันธ์ที่แสดงความสัมพันธ์ระหว่างอาการเจ็บป่วย และสรรพคุณของสมุนไพรภายใต้ค่าสนับสนุนขั้นต่ำ และค่าความเชื่อมั่นขั้นต่ำเท่ากับ 0.50 และ 0.90 ตามลำดับ โดยการนำกฎความสัมพันธ์ไปใช้งานนั้น คณะผู้วิจัยได้คัดกรองกฎความสัมพันธ์ที่สนใจใน 2 รูปแบบคือ 1) $\langle symptom \rangle \Rightarrow \langle herb \rangle$ แสดงความสัมพันธ์ระหว่างอาการเจ็บป่วยหนึ่งอาการกับสมุนไพรหนึ่งรายชื่อ และ 2) $[\langle symptom \rangle] \Rightarrow \langle herb \rangle$ แสดงความสัมพันธ์ระหว่างกลุ่มของอาการเจ็บป่วยและสมุนไพรหนึ่งรายชื่อ จากนั้นนำผลลัพธ์ของกระบวนการทำเหมืองข้อความไปใช้เป็นข้อมูลป้อนเข้าแก่ระบบสนับสนุนการตัดสินใจฯ ต่อไป

โดยระบบสนับสนุนการตัดสินใจที่ได้ทำการพัฒนาขึ้นนี้ ทำหน้าที่เป็นสื่อกลางในการเชื่อมต่อระหว่างข้อมูลที่มีความซับซ้อน หลากหลาย และปริมาณมากเข้ากับผู้ใช้งานที่สนใจจะใช้สมุนไพรให้สามารถเข้าถึงหรือสืบค้นข้อมูลได้ง่าย ตลอดจนสามารถเห็นข้อมูลผ่านการสกัดจนได้สารสนเทศเชิงสถิติ ซึ่งสามารถทำความเข้าใจได้ทันที อันจะส่งผลต่อความสะดวกในการตัดสินใจ และสร้างความเชื่อมั่นต่อการเลือกใช้สมุนไพรในการรักษาบำบัดอาการเจ็บป่วยหรือดูแลสุขภาพของตนเอง

2) ข้อเสนอแนะ

ข้อมูลต้นฉบับที่นำมาวิเคราะห์ และประมวลผลในการวิจัยครั้งนี้ อยู่ในรูปแบบข้อมูลที่ไม่มีโครงสร้าง ซึ่งมีความไม่เป็นระเบียบของข้อมูลอยู่มาก อาจส่งผลให้การประมวลผลข้อมูลเกิดความผิดพลาดได้ง่าย เช่น การเขียนผิดหรือการสะกดผิด คำพ้องรูป หรือคำที่มีความหมายกำกวม ตลอดจนการจัดเรียงข้อมูลปะปนกันอย่างไม่เป็นระบบ หรืออาจมีข้อมูลโฆษณาเกินจริงปนรวมอยู่ภายในเนื้อหา สารสนเทศหรือกฎความสัมพันธ์ที่ได้จึงเป็นการประมวลผลข้อมูลโดยพิจารณาจากกระบวนการทางสถิติเป็นหลัก ซึ่งเป็นการมองหาความสัมพันธ์ระหว่างคำสำคัญที่ปรากฏร่วมกันในชุดเอกสารที่พิจารณา กฎความสัมพันธ์ที่เกิดขึ้นจึงไม่ได้ใช้บริบทของคำอื่น ๆ ในเอกสารมาพิจารณาไปด้วย ทำให้ยังไม่สามารถบ่งชี้ได้ว่ากฎความสัมพันธ์ที่แสดงความสัมพันธ์ระหว่างอาการเจ็บป่วย และสมุนไพรนั้นเป็นความสัมพันธ์ในทางบวก หรือทางลบ เช่น จากกฎที่แสดงความสัมพันธ์ระหว่างคำสำคัญ “เบาหวาน” กับ “น้ำมะพร้าว” ซึ่งเป็นกฎความสัมพันธ์ที่มีค่าระดับความเชื่อมั่นที่สูง แต่ก็ยัง

ไม่สามารถบ่งชี้อย่างเจาะจงพื้นที่ว่าระหว่าง “เบาหวาน” กับ “น้ำมะพร้าว” มีความสัมพันธ์กันในทางบวก หรือว่าทางลบ คือเมื่อมีอาการเบาหวานแล้วควรดื่มน้ำมะพร้าว หรือไม่ควรดื่มน้ำมะพร้าว เป็นต้น

ดังนั้นการนำข้อมูลที่อยู่ในระบบสนับสนุนการตัดสินใจ ไปใช้งานยังมีความจำเป็นต้องใช้ร่วมกับการให้คำแนะนำจากผู้เชี่ยวชาญร่วมด้วยเป็นหลัก ซึ่งระบบสนับสนุนการตัดสินใจ นี้จะสามารถเป็นเครื่องมือช่วยประกอบคำแนะนำที่ดีแก่ผู้เชี่ยวชาญในการทบทวนถึงสรรพคุณของสมุนไพรต่าง ๆ ที่มีการเผยแพร่อยู่บนระบบอินเทอร์เน็ตได้เป็นอย่างดีก่อนที่จะให้คำแนะนำแก่ผู้สนใจ หรือสำหรับผู้ที่ต้องการใช้สมุนไพรในการบำบัดอาการเจ็บป่วยหรือเสริมสร้างสุขภาพให้แข็งแรง ก็สามารถใช้ระบบสนับสนุนการตัดสินใจ นี้เป็นเครื่องมือช่วยตัดสินใจประกอบกับคำแนะนำจากผู้เชี่ยวชาญ หรือประกอบกับความรู้จากแหล่งอื่นด้วยเช่นกัน

References

- Al-Daihan, SM., & Abrahamsb, A. (2016). A Text Mining Analysis of Academic Libraries Tweets. *Journal of Academic Librarianship*. 2016; 42: 135-143.
- Best, J.W. (1981). *Research in Education*. New Jersey: Prentice - Hall.
- Chatree, J., Sookkhee, P., & Sookkhee, S. (2016). Text Mining Approach to Extract Information of Symptoms and Herbs from Data on Thai Web Sites. *The International Conference on Applied Statistics 2016 (ICAS 2016)*. Phuket Graceland Resort & Spa, Phuket, Thailand, 13-15 July 2016.
- Gadia, K., & Bhowmick, K. (2015). Parallel Text Mining in Multicore Systems Using FP-tree Algorithm. *Procedia Computer Science*. 2015; 45: 111-117.
- He, Y., Zheng, X., Sit, C., Loo, WT., Wang, Z., Xie, T., et al. (2012). Using Association Rules Mining to Explore Pattern of Chinese Medicinal Formulae (Prescription) in Treating and Preventing Breast Cancer Recurrence and Metastasis. *Journal of Translational Medicine*. 2012; 10(1): 1-8.
- Hofmann, M., & Klinkenberg R. (2014). *LapidMiner Data Mining Use Cases and Business Analytics Applications*. USA: CRC Press.
- Hunyadi, D. (2011). Performance Comparison of Apriori and FP-growth Algorithms in Generating Association Rules. *ECC'1 Proceedings of The 5th European Conference on European Computing Conference; 2011*, Stevens Point, Wisconsin, USA.
- Kang, JH., Yang, DH., Park, YB., & Kim, SB. (2012). A Text mining Approach to Find Patterns Associated with Diseases and Herbal Materials in Oriental Medicine. *International Journal of Information and Education Technology*. 2012; 2(3): 224-226.
- Likert, R. (1967). The Method of Constructing and Attitude Scale. In *Reading in Fishbein, M (Ed.), Attitude Theory and Measurement* (pp. 90 - 95). New York: Wiley & Son.
- Mahgoub, H. (2006). Mining Association Rules from Unstructured Documents. *International Journal of Applied Mathematics and Computer Science*. 2006; 1(4): 201-206.
- Mahgoub, H., Rosner, D., Ismail, N. & Torkey, F. (2008). A Text Mining Technique Using Association Rules Extraction. *International Journal of Information and Mathematical Science* 4(1).
- Marir, F., Said, H., & Al-Obeidat, F. (2016). Mining the Web and Literature to Discover New Knowledge about Diabetes. *Procedia Computer Science* 83 (2016): 1256–1261.

- Nak khoṃmūn læ pramoēnphon krom phatthanākān phǣt phǣnthai læ kānphǣt thāng lūāk krasuāng sāthāranasuk. (2013). rāingān kān sāthāranasuk Thai dān kānphǣt phǣnthai kānphǣt phūnbān læ kānphǣt thāng lūāk 2011-2013 [Thai Traditional and Alternative Health Profile, 2011-2013]. nonhaburi: krom phatthanākān phǣt phǣnthai læ kānphǣt thāng lūāk krasuāng sāthāranasuk.
- Photchanānukrom ‘ilekthronik Thai - ‘Angkrit. (2009). LEXITRON [Online]. sūpkhon čāk : http://lexitron.nectec.or.th/2009_1
- Sukhi, P. (2015). rabop sārasonthet phūmpanya thōngthīn dān kāndulæ raksā sukkhaphāp nai sī klum chattiphan čangwat sīsakēt [Information System for The Local Wisdom of Health Care in Four Ethnic Groups Sisaket Province]. kān prachum wichākān radap chāt ratchaphat wichai khrang thī 3. mahāwitthayalai ratchaphat nakhoṇ sī tham rāt, nakhoṇ sī tham rāt, yīsip - yīsipsōng Phrutsaphākhom 2015.
- Sun theknōyī ‘ilekthronik læ khoṃphiutōē hæng chāt. (1994). prokræm tat khoṃhwām phāsā Thai (Thai Lexeme Tokenizer: LexTo) [prokræm khoṃphiutōē]. sūpkhon čāk <http://www.sansarn.com/lexto>
- Wongwilaisakun, W. (2013). mūang khoṃhwām læ kānprayukchai [Text Mining and Its Applications]. Panyapiwat Journal, 157- 165.