

การพัฒนาเว็บแอปพลิเคชันเพื่อการจัดประเภทผลิตภัณฑ์สุขภาพ

กฤติพล จันทรานันต์¹, วีรยุทธ์ เลิศนที²

¹นักศึกษาปริญญาโท หลักสูตรเภสัชศาสตรมหาบัณฑิต สาขาสารสนเทศศาสตร์ทางสุขภาพ

คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร

²สาขาสุขภาพดิจิทัล คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร

บทคัดย่อ

วัตถุประสงค์: เพื่อพัฒนาเว็บแอปพลิเคชันสำหรับการจัดประเภทผลิตภัณฑ์สุขภาพที่อยู่ในการกำกับของสำนักงานคณะกรรมการอาหารและยา (อย.) โดยเป็นเครื่องมือสนับสนุนการตัดสินใจที่ช่วยให้ผู้ประกอบการและนักวิจัยสามารถจัดประเภทผลิตภัณฑ์ของตนและเตรียมการขึ้นทะเบียนได้ด้วยตนเอง **วิธีการวิจัย:** เว็บแอปพลิเคชันนี้พัฒนาโดยใช้แพลตฟอร์ม Google Apps Script (GAS) โดยมีส่วน rule-based system สำหรับจัดประเภทผลิตภัณฑ์ตามกฎหมายและตามแผนภูมิการจัดประเภทของ อย. นอกจากนี้ ยังมีส่วนให้คำแนะนำเบื้องต้นด้วย generative AI (artificial intelligence) ประเภท large language model (LLM) โดยใช้โมเดล gpt-5 จาก OpenAI ผ่าน API ผู้วิจัยประเมินประสิทธิภาพของระบบโดยเปรียบเทียบผลการจัดประเภทโดยใช้ rule-based system, generative AI และผู้เชี่ยวชาญจาก อย. 2 ท่าน ตลอดจนประเมินความพึงพอใจของผู้ใช้งาน 30 คนที่มารับบริการที่ศูนย์บริการผลิตภัณฑ์สุขภาพเบ็ดเสร็จของ อย. **ผลการวิจัย:** การจัดประเภทผลิตภัณฑ์โดย rule-based system และ generative AI มีความสอดคล้องกับผลการจัดประเภทของผู้เชี่ยวชาญ 2 ท่านในระดับสูงมาก (Fleiss' Kappa = 0.862) และไม่มี ความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างผลการประเมินในแต่ละวิธี rule-based system ให้ผลสอดคล้องกับผู้เชี่ยวชาญในระดับสูงมาก (K = 0.918) ขณะที่ generative AI สอดคล้องในระดับดี (K = 0.751) และมี overall accuracy อยู่ที่ 0.933 สำหรับ rule-based system และ 0.800 สำหรับ generative AI ผู้เชี่ยวชาญจาก อย. ประเมินคุณภาพคำแนะนำจาก generative AI ด้วยคำถามที่พบบ่อย 5 ข้อ (ถามซ้ำ 3 รอบต่อข้อ) พบว่า คุณภาพโดยรวมอยู่ในระดับปานกลาง (3.20 จาก 5 คะแนนเต็ม) ความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชันอยู่ในระดับมากที่สุด (ค่าเฉลี่ย 4.38–4.47 จากคะแนนเต็ม 5) **สรุป:** เว็บแอปพลิเคชันที่พัฒนาขึ้นมีประสิทธิภาพในการเป็นเครื่องมือสนับสนุนการตัดสินใจเพื่อการจัดประเภทผลิตภัณฑ์ โดยสามารถช่วยเตรียมความพร้อมและเป็นแหล่งเรียนรู้เบื้องต้นสำหรับผู้ประกอบการหรือนักวิจัยในการเตรียมการขึ้นทะเบียนผลิตภัณฑ์สุขภาพได้ อย่างไรก็ตามควรใช้เว็บแอปพลิเคชันที่พัฒนาขึ้นร่วมกับข้อแนะนำจากผู้เชี่ยวชาญ

คำสำคัญ: เว็บแอปพลิเคชัน การจัดประเภทผลิตภัณฑ์สุขภาพ เครื่องมือประเภทอิงกฎ ปัญญาประดิษฐ์

รับต้นฉบับ: 10 เม.ย. 2569, ได้รับบทความฉบับปรับปรุง: 15 พ.ค. 2569, รับลงตีพิมพ์: 21 พ.ค. 2569

ผู้ประสานงานบทความ: วีรยุทธ์ เลิศนที คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร เลขที่ 6 ถนนราชมรรคาใน ตำบลพระปฐมเจดีย์ อำเภอเมือง จังหวัดนครปฐม 73000 **E-mail:** LERTNATTEE_V@su.ac.th

Development of a Web Application for Health Product Classification

Kridiphol Janthranant¹, Verayuth Lertnattee¹Student in Master of Pharmacy Program in Health Informatics, Faculty of Pharmacy, Silpakorn University²Department of Digital Health, Faculty of Pharmacy, Silpakorn University

Abstract

Objective: To develop a web application for the classification of health products regulated by the Food and Drug Administration (FDA). The application serves as a decision support tool, enabling entrepreneurs and researchers to self-classify their products and prepare for the product registration process. **Methods:** The web application was developed on the Google Apps Script (GAS) platform. It incorporates a rule-based system designed to classify products in accordance with relevant legislation and FDA classification flowcharts. Additionally, it integrates a generative artificial intelligence (AI), specifically a Large Language Model (LLM) utilizing the OpenAI gpt-5 model via API, to provide preliminary guidance. The system's performance was evaluated by comparing the classification outcomes among the rule-based system, the generative AI, and two FDA experts. Furthermore, user satisfaction was assessed among 30 individuals utilizing services at the FDA's One Stop Service Center. **Results:** The product classifications derived from the rule-based system and the generative AI demonstrated a very high level of agreement with the two experts (Fleiss' Kappa = 0.862), with no statistically significant differences observed between the assessment methods. The rule-based system exhibited an almost perfect agreement with the experts (K = 0.918), whereas the generative AI showed a substantial level of agreement (K = 0.751). The overall accuracy was 0.933 for the rule-based system and 0.800 for the generative AI. The FDA experts' evaluation of the generative AI's guidance, based on 5 frequently asked questions (repeated for 3 iterations per question), indicated an overall quality at a moderate level (mean score: 3.20 out of 5). User satisfaction with the web application was rated at the highest level (mean score: 4.38 – 4.47 out of 5). **Conclusion:** The developed web application is an effective decision support tool for health product classification. It serves as a valuable preliminary learning resource and assists entrepreneurs or researchers in preparing for health product registration. Nonetheless, the application should be utilized in conjunction with professional guidance from experts.

Keywords: web application, health product classification, rule-based tool, artificial intelligence

บทนำ

สำนักงานคณะกรรมการอาหารและยา (อย.) มีหน้าที่ในการกำกับดูแลคุณภาพ ความปลอดภัย และประสิทธิภาพตามคำกล่าวอ้างของผลิตภัณฑ์สุขภาพ ตามกรอบที่กำหนดไว้ในกฎกระทรวงสาธารณสุข ปี พ.ศ. 2563 (1) อย. แบ่งผลิตภัณฑ์สุขภาพออกเป็นหมวดหมู่ตามระดับความเสี่ยงของตัวผลิตภัณฑ์เองและความเปราะบางของกลุ่มผู้ใช้งาน (risk-based approach) โดยมีพระราชบัญญัติควบคุมผลิตภัณฑ์แต่ละประเภท ได้แก่ พระราชบัญญัติยา พ.ศ. 2510 พระราชบัญญัติอาหาร พ.ศ. 2522 พระราชบัญญัติวัตถุอันตราย พ.ศ. 2535 พระราชบัญญัติเครื่องสำอาง พ.ศ. 2558 พระราชบัญญัติเครื่องมือแพทย์ พ.ศ. 2562 พระราชบัญญัติผลิตภัณฑ์สมุนไพร พ.ศ. 2562 และพระราชบัญญัติให้ใช้ประมวลกฎหมายยาเสพติด พ.ศ. 2564 (2-8) โดยในแต่ละประเภทผลิตภัณฑ์ต้องการข้อมูลเพื่อพิสูจน์ถึงประสิทธิภาพ คุณภาพ และความปลอดภัยที่แตกต่างกันไปตามลักษณะความเสี่ยง ทำให้การยื่นขึ้นทะเบียนผลิตภัณฑ์มีความจำเพาะและรายละเอียดปลีกย่อยที่แตกต่างกัน ความหลากหลายของตัวบทกฎหมายนี้เองเป็นอุปสรรคสำคัญที่ทำให้ผู้ประกอบการเกิดความสับสนในการจำแนกประเภทผลิตภัณฑ์

ด้วยความต้องการผลิตภัณฑ์ด้านสุขภาพเพิ่มขึ้นอย่างมาก (9) ส่งผลให้เกิดการพัฒนาเทคโนโลยีด้านเภสัชภัณฑ์และการคิดค้นผลิตภัณฑ์ใหม่ ๆ ควบคู่กับนโยบายภาครัฐที่เน้นส่งเสริมการพัฒนาสมุนไพร (10) และส่งเสริมการผลิตผลิตภัณฑ์สุขภาพภายในประเทศเพื่อการพึ่งพาตนเอง (11) ปัจจัยดังกล่าวนำไปสู่การเกิดผู้ประกอบการรายใหม่และนักวิจัยเพื่อพัฒนาผลิตภัณฑ์สุขภาพนวัตกรรมอย่างต่อเนื่อง จากข้อมูลการปฏิบัติงานของกลุ่มให้คำปรึกษาผลิตภัณฑ์สุขภาพ กองผลิตภัณฑ์สุขภาพ นวัตกรรมและการบริการ (สนบ.) พบว่า ตั้งแต่ช่วงเดือนตุลาคม พ.ศ. 2566 จนถึงเดือนมิถุนายน พ.ศ. 2567 มีผู้ขอรับบริการจัดประเภทผลิตภัณฑ์สุขภาพทั้งสิ้น 878 คำขอสะท้อนว่าผู้ประกอบการและนักวิจัยจำนวนมากยังขาดความเข้าใจเกี่ยวกับแนวทางการจัดแบ่งประเภทผลิตภัณฑ์สุขภาพ ซึ่งอาจทำให้เกิดปัญหาในการขึ้นทะเบียน เช่น การกล่าวอ้างสรรพคุณของผลิตภัณฑ์ที่มีลักษณะคาบเกี่ยวกับพระราชบัญญัติมากกว่าหนึ่งฉบับ ปัญหาการพัฒนาผลิตภัณฑ์ที่ไม่สอดคล้องกับข้อกำหนด หรือไม่มีข้อมูลที่เหมาะสมในการขึ้นทะเบียน บางส่วนติดปัญหาการนำเข้า

เนื่องจากเอกสารไม่สอดคล้องตามข้อกำหนด ก่อให้เกิดความล่าช้า ต้นทุนทางธุรกิจที่สูงขึ้นและการสูญเสียโอกาสทางการตลาดจากการยื่นขึ้นทะเบียนผิดประเภท

ปัจจุบันกระบวนการจัดประเภทผลิตภัณฑ์สุขภาพผ่านระบบ E-consultation ใช้ระยะเวลาดำเนินการประมาณ 30 – 120 วันทำการ (12) ทั้งนี้ขึ้นกับความซับซ้อนของประเด็นที่สอบถาม รวมถึงความครบถ้วนสมบูรณ์ของเอกสารที่ชื่อยื่นประกอบ ถึงแม้จะสามารถช่วยแก้ไขปัญหาในการยื่นขึ้นทะเบียนที่ผิดประเภทได้ในระดับหนึ่ง แต่ยังพบว่ามีความจำกัดในการดำเนินการคือความล่าช้าในการให้บริการจากผู้ใช้งานที่มีปริมาณมาก ความซับซ้อนของการเรียกเอกสารในการจัดประเภท และข้อกังวลต่อความลับทางการค้า จากช่องว่างดังกล่าวนำไปสู่ความจำเป็นในการพัฒนาเครื่องมือที่ช่วยให้ผู้ประกอบการสามารถจัดประเภทผลิตภัณฑ์สุขภาพเบื้องต้นได้ด้วยตนเองอย่างรวดเร็ว ซึ่งจะช่วยเพิ่มความถูกต้องในการเตรียมเอกสารให้เป็นไปตามข้อกำหนดและลดความสูญเสียทางด้านเวลา ลดความกังวลในด้านความลับทางการค้า และลดการใช้ทรัพยากรเงินทุน

จากการทบทวนวรรณกรรมพบว่า การพัฒนาในรูปแบบเว็บแอปพลิเคชัน (web application) จะช่วยลดอุปสรรคในการเข้าถึงเครื่องมือของผู้ประกอบการโดยไม่จำเป็นต้องติดตั้งซอฟต์แวร์เพิ่มเติม (13, 14) ขณะเดียวกันเทคโนโลยีปัญญาประดิษฐ์ (artificial intelligence: AI) ได้รับการพัฒนาเป็นอย่างมากจนในปัจจุบันมีศักยภาพในการทำหน้าที่เป็นระบบช่วยเหลือการตัดสินใจ (15) ผ่านความสามารถด้านการประมวลภาษาธรรมชาติ (natural language processing: NLP) (16, 17) และมีพื้นฐานการเรียนรู้จากฐานข้อมูลขนาดใหญ่ (big data) จึงมีศักยภาพในการประยุกต์ใช้เป็นเครื่องมือให้คำแนะนำเบื้องต้นในลักษณะผู้ช่วยเสมือน (virtual assistant) ซึ่งการใช้ระบบจะให้ข้อมูลและคำแนะนำที่มีความสม่ำเสมอเมื่อเทียบกับการให้คำแนะนำจากเจ้าหน้าที่ (18) ในปัจจุบันองค์กรควบคุมยาในระดับนานาชาติ เช่น U.S. Food and Drug Administration (USFDA), European Medicines Agency และ World Health Organization (WHO) ได้ให้ความสำคัญกับการนำ AI มาใช้สนับสนุนการทำงานด้านกฎระเบียบและการตัดสินใจเชิงวิชาการ โดยเฉพาะในด้านวิเคราะห์ข้อมูลขนาดใหญ่ การสกัดข้อมูลจากเอกสารจำนวนมาก การช่วยตีความขอบเขตของการกำกับดูแลผลิตภัณฑ์ การจัดระดับความเสี่ยง และการสนับสนุนการตัดสินใจบนฐานข้อมูล

อย่างไรก็ตามแนวทางจากองค์กรสากลไม่ได้มุ่งเน้นการใช้ AI เพื่อทดแทนผู้เชี่ยวชาญทั้งหมด แต่เน้นในการนำ AI มาใช้เพื่อเป็นเครื่องมือช่วยเหลือการตัดสินใจภายใต้การกำกับดูแลที่เหมาะสม โดยมุ่งเน้นความโปร่งใส ความน่าเชื่อถือ และการตรวจสอบย้อนกลับได้ เพื่อให้การนำ AI มาประยุกต์ใช้บรรลุวัตถุประสงค์ควบคู่กับความปลอดภัยและความเหมาะสมเชิงจริยธรรม (19-21)

ด้วยเหตุนี้ผู้วิจัยได้พัฒนาเว็บแอปพลิเคชันเพื่อช่วยให้ผู้ประกอบการสามารถจัดประเภทผลิตภัณฑ์สุขภาพได้ด้วยตนเอง พร้อมทั้งมีส่วนให้คำแนะนำเบื้องต้นในการขึ้นทะเบียนด้วยเทคโนโลยีปัญญาประดิษฐ์ และทำการประเมินทั้งในด้านความถูกต้องของการจัดประเภทผลิตภัณฑ์ผลิตภัณฑ์ คุณภาพของคำแนะนำจาก generative AI และความพึงพอใจของผู้ใช้งาน

วิธีการวิจัย

การศึกษานี้เป็นการศึกษาเชิงวิจัยพัฒนาและประเมินประสิทธิภาพ ผ่านการพิจารณาจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์มหาวิทยาลัยศิลปากร (เลขที่โครงการ REC68.0506-085-3987) หมายเลขใบรับรอง COE 68.0507-048 ขั้นตอนการดำเนินการวิจัยแบ่งออกเป็น 2 ระยะ คือ ระยะพัฒนาเว็บแอปพลิเคชันและระยะประเมินประสิทธิภาพเว็บแอปพลิเคชันที่พัฒนาขึ้น

ระยะที่ 1 ระยะพัฒนาเว็บแอปพลิเคชัน

ผู้วิจัยได้นำทฤษฎีวงจรการพัฒนาระบบ (system development life cycle: SDLC) ในรูปแบบ waterfall model (22) มาประยุกต์ใช้ เนื่องจากเป็นการพัฒนาเครื่องมือนวัตกรรมที่ยังไม่มีความซับซ้อนสูง มีลำดับขั้นตอนการพัฒนาที่ชัดเจน และเอื้อต่อการตรวจสอบความถูกต้องในแต่ละขั้นตอน โดยในการพัฒนาระบบ ผู้วิจัยทวนสอบและปรับปรุงรายละเอียดบางส่วนระหว่างดำเนินงานตามความเหมาะสม โดยยังคงยึดโครงสร้างการดำเนินงานตามแนวคิด waterfall model เป็นหลักเพื่อให้ได้ระบบที่มีคุณภาพและลดความเสี่ยงต่อการพัฒนาระบบที่ไม่สามารถนำไปใช้งานได้จริง แนวคิดดังกล่าวมี 6 ขั้นตอนย่อยดังนี้

ขั้นตอนที่ 1. การวางแผน: ผู้วิจัยศึกษานิยามผลิตภัณฑ์ตามพระราชบัญญัติผลิตภัณฑ์สุขภาพทุกฉบับ รวมทั้งแนวทางการจัดประเภทผลิตภัณฑ์สุขภาพของกลุ่มให้คำปรึกษาผลิตภัณฑ์สุขภาพ สนบ.

ขั้นตอนที่ 2. การวิเคราะห์ระบบงาน: ผู้วิจัยสัมภาษณ์ผู้เชี่ยวชาญกลุ่มงานให้คำปรึกษา ถึงลักษณะการทำงาน ระยะเวลาดำเนินการในแต่ละขั้นตอนและเอกสารประกอบที่จำเป็นต้องใช้ในการดำเนินการ ระบบสารสนเทศที่รองรับและความต้องการในการพัฒนาระบบก่อนนำไปออกแบบตัวระบบ

ขั้นตอนที่ 3. การออกแบบระบบ: ผู้วิจัยออกแบบระบบเว็บแอปพลิเคชันโดยเลือกพัฒนาบนแพลตฟอร์ม Google Apps Script (GAS) ซึ่งพัฒนาขึ้นด้วยภาษา JavaScript ซึ่งมีข้อดีคือสามารถเชื่อมต่อกับบริการของ Google ได้ง่าย เช่น Gmail, Google Sheets, Google Drive เป็นต้น (23) ซึ่งเจ้าหน้าที่ในหน่วยงานคุ้นเคยกับการใช้งานอยู่เดิม และแพลตฟอร์มสามารถใช้พัฒนาระบบตัวต้นแบบได้ง่าย การพัฒนาใช้ rule-based system ซึ่งจัดเป็นเครื่องมือปัญญาประดิษฐ์ประเภท rule-based AI ในการจัดประเภทผลิตภัณฑ์ พร้อมทั้งมีส่วนให้คำแนะนำเบื้องต้นในประเด็นที่เกี่ยวข้องกับการขึ้นทะเบียนผลิตภัณฑ์สุขภาพด้วย generative AI ประเภท large language model (LLM) โดยสถาปัตยกรรมของระบบ (system architecture) ประกอบด้วย 3 ส่วนหลัก ได้แก่ ส่วนการแสดงผลและปฏิสัมพันธ์กับผู้ใช้ (front-end) ส่วนประมวลผลตรรกะ (application logic layer หรือ back-end) และส่วนจัดเก็บฐานข้อมูล (data layer) ผู้วิจัยวางแผนการออกแบบระบบดังนี้

ส่วนที่ 1 การแสดงผลและปฏิสัมพันธ์กับผู้ใช้ (front-end): ผู้วิจัยออกแบบให้ผู้ใช้สามารถเข้าถึงข้อมูลพื้นฐานเบื้องต้นในการศึกษาข้อมูลเกี่ยวกับการขึ้นทะเบียนผลิตภัณฑ์ เช่น ข้อมูลเกี่ยวกับการจัดประเภทผลิตภัณฑ์ ลิงก์ติดต่อหน่วยงาน เป็นต้น อย่างไรก็ตามหากผู้ใช้งานต้องการเข้าใช้งานฟังก์ชันการจัดประเภทผลิตภัณฑ์ จำเป็นต้องลงทะเบียนก่อน โดยผู้ใช้งานต้องตั้งชื่อผู้ใช้งาน รหัสผ่าน และให้ความยินยอมในเงื่อนไขการใช้งาน ข้อมูลเหล่านี้ถูกเก็บไว้ในฐานข้อมูลของระบบ โดยข้อมูลรหัสผ่านจะมีการ encrypt รหัสเพื่อความปลอดภัยเมื่อผู้ใช้งานลงทะเบียนเรียบร้อยแล้ว ก็จะสามารถ log-in เข้าใช้งานฟีเจอร์ต่าง ๆ ได้อย่างเต็มประสิทธิภาพ โดยมีฟังก์ชันภายในเว็บแอปพลิเคชันดังนี้

ส่วนที่ 1.1 ส่วนการจัดประเภทผลิตภัณฑ์ ผู้ใช้งานสามารถกรอกข้อมูลพื้นฐานของผลิตภัณฑ์ เช่น ชื่อผลิตภัณฑ์ วัตถุประสงค์การใช้งาน รูปแบบผลิตภัณฑ์ หรือ

ส่วนประกอบสำคัญ จากนั้นผู้ใช้งานจัดประเภทผลิตภัณฑ์ โดยการตอบแบบฟอร์มที่มีลักษณะเป็นตัวเลือกแบบเลือกได้เพียงอย่างเดียว (radio) เมื่อตอบจนครบทุกข้อแล้ว ระบบจะส่งข้อมูลเข้าสู่ส่วนประมวลผลตรรกะโดย rule-based system เพื่อจัดประเภทผลิตภัณฑ์ เมื่อประมวลผลเรียบร้อยแล้ว ระบบจะส่งผลลัพธ์กลับมายังหน้าจอผู้ใช้งาน ซึ่งผู้ใช้งานสามารถใช้ฟังก์ชันส่งรายงานเข้าอีเมลที่ต้องการ เพื่อเก็บเป็นข้อมูลส่วนตัวได้

ส่วนที่ 1.2 ส่วนให้คำแนะนำเบื้องต้นเพื่อการขึ้นทะเบียนผลิตภัณฑ์สุขภาพโดย generative AI เริ่มจากผู้ใช้งานกรอกคำถามที่ต้องการ จากนั้นระบบส่งข้อมูลต่อไปยัง generative AI เมื่อประมวลผลเรียบร้อยแล้ว ระบบจะส่งข้อมูลกลับมาในหน้าจอผู้ใช้งาน ซึ่งฟังก์ชันนี้จะช่วยให้ผู้ใช้งานสามารถถามคำถามที่มีเกี่ยวข้องกับการขึ้นทะเบียนผลิตภัณฑ์ ทำให้สามารถเลือกคำตอบใน rule-based system ได้ถูกต้องมากขึ้น และเรียนรู้ประเด็นต่าง ๆ ที่เกี่ยวข้องได้ ทั้งนี้ ส่วนให้คำแนะนำนี้ถูกกำหนดให้ตอบกลับเฉพาะประเด็นที่เกี่ยวข้องกับการขึ้นทะเบียนผลิตภัณฑ์เท่านั้น เพื่อป้องกันการใช้งานที่ผิดวัตถุประสงค์

ส่วนที่ 1.3 ส่วนสอบถามข้อมูลเพิ่มเติมกับเจ้าหน้าที่โดยตรง ผู้ใช้งานสามารถกรอกคำถามเพิ่มเติมเพื่อให้ระบบส่งข้อมูลผลิตภัณฑ์และประเด็นคำถามไปยังเจ้าหน้าที่ที่เกี่ยวข้องโดยตรง ทำให้เจ้าหน้าที่ได้รับข้อมูลเพื่อพิจารณาได้อย่างรวดเร็วขึ้น

ในส่วนการแสดงผลผู้วิจัยพัฒนาโดยใช้ภาษา HTML, CSS และ JavaScript ร่วมกับเครื่องมือตกแต่งเช่น Bootstrap (24) และ Canva เพื่อให้เกิดความสวยงาม

ส่วนที่ 2 การประมวลผลตรรกะ (application logic layer หรือ back-end): ผู้วิจัยได้ออกแบบส่วนประมวลผลตรรกะ โดยแบ่งออกเป็น 2 ส่วนย่อย คือ

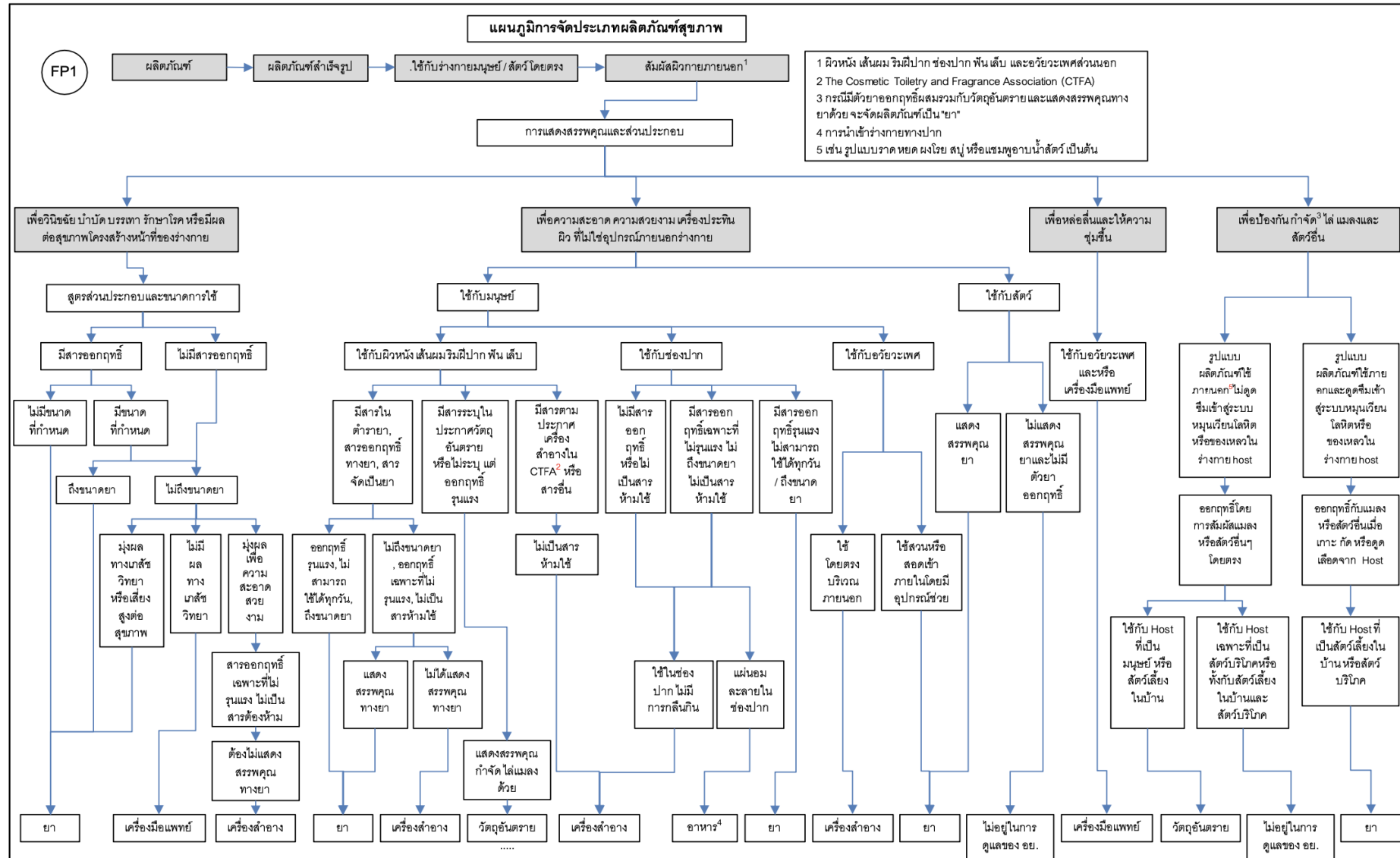
ส่วนที่ 2.1 ส่วนตรรกะการจัดประเภทผลิตภัณฑ์ (product classification rule-based system): ผู้วิจัยออกแบบส่วน rule-based system ของระบบเว็บแอปพลิเคชัน ซึ่งผู้วิจัยได้สกัดตรรกะการตัดสินใจจากพระราชบัญญัติที่เกี่ยวกับผลิตภัณฑ์สุขภาพทุกฉบับร่วมกับแผนภูมิการจัดประเภทผลิตภัณฑ์สุขภาพของกลุ่มให้คำปรึกษา สนบ. และจัดทำให้ออกมาในรูปแบบโครงสร้างลำดับแบบต้นไม้ตัดสินใจ (decision tree) จากนั้นนำไปสร้าง rule-based system เพื่อใช้เป็นตรรกะหลักในการทำงานของระบบ รูปที่ 1 แสดงตัวอย่างบางส่วนของแผนภูมิการจัดประเภท

ผลิตภัณฑ์สุขภาพ ซึ่งคัดเลือกเฉพาะบางกลุ่มผลิตภัณฑ์มาแสดงเพื่อความกระชับในการนำเสนอ แผนภูมิการจัดประเภทผลิตภัณฑ์สุขภาพฉบับเต็มได้ครอบคลุมผลิตภัณฑ์สุขภาพทุกประเภทที่อยู่ในขอบเขตการวิจัย รวมถึงผลิตภัณฑ์สมุนไพรและยาเสพติด

ผู้วิจัยได้พัฒนาให้ผลลัพธ์ในการจัดประเภทจาก rule-based system สามารถแสดงผลได้ 9 ประเภท คือ 1) ยา 2) เครื่องมือแพทย์ 3) ยาเสพติด 4) วัตถุอันตราย 5) อาหาร 6) เครื่องสำอาง 7) ผลิตภัณฑ์สมุนไพร 8) ผลิตภัณฑ์ที่ไม่อยู่ในการค้ากับดูแลของ อย. และ 9) แนะนำให้ติดต่อเจ้าหน้าที่

ส่วนที่ 2.2 ส่วนการให้คำแนะนำจากเครื่องมือปัญญาประดิษฐ์ (generative AI-Based guidance section): ผู้วิจัยออกแบบส่วนของระบบให้มีการประยุกต์ใช้ generative AI ประเภท LLM สำหรับให้คำแนะนำเบื้องต้นแก่ผู้ใช้งาน โดยเลือกใช้แพลตฟอร์ม ChatGPT (OpenAI) เนื่องจากในขณะดำเนินงานวิจัย แพลตฟอร์มนี้มีความพร้อมของ API ที่มีเสถียรภาพ มีข้อมูลเอกสารประกอบการพัฒนาที่ครบถ้วน สามารถใช้งานได้อย่างสะดวกในการเชื่อมต่อกับ GAS ซึ่งเป็นแพลตฟอร์มหลักในการพัฒนาเครื่องมือ นอกจากนี้ยังมีความสามารถในการประมวลผลภาษาธรรมชาติและการให้เหตุผลเชิงบริบทที่ครอบคลุมในระดับที่เหมาะสมต่อการเป็นเครื่องมือให้คำแนะนำเบื้องต้นในด้านการขึ้นทะเบียนผลิตภัณฑ์ ดังนั้น จึงทำให้สามารถนำแพลตฟอร์มดังกล่าวมาพัฒนาเป็นเครื่องมือต้นแบบให้พร้อมใช้งานและทดสอบได้ภายในกรอบระยะเวลาของการศึกษา ผู้วิจัยใช้งานผ่าน API เมื่อวันที่ 27 มิถุนายน พ.ศ. 2568 โดยใช้ OpenAI model identifier: gpt-5 และกำหนดบทบาทให้ทำหน้าที่เป็นผู้เชี่ยวชาญด้านการขึ้นทะเบียนผลิตภัณฑ์สุขภาพและให้ตอบคำถามเฉพาะประเด็นที่เกี่ยวข้องกับการขึ้นทะเบียนผลิตภัณฑ์เท่านั้นเพื่อป้องกันการใช้งานที่ผิดวัตถุประสงค์ รูปที่ 2 แสดงคำสั่งในการกำหนดบทบาทของ Generative AI (system prompt) เพื่อกำหนดขอบเขตในการตอบและป้องกันการใช้งานที่ผิดวัตถุประสงค์

ส่วนที่ 3 การจัดเก็บฐานข้อมูล (data layer): ผู้วิจัยออกแบบฐานข้อมูลในการเก็บข้อมูลของผู้ใช้งาน เช่น ชื่อผู้ใช้งาน รหัสผ่าน ข้อมูลผลิตภัณฑ์สุขภาพของผู้ใช้งาน และส่วนตรรกะของการจัดประเภทผลิตภัณฑ์ ทั้งนี้ การเลือกพัฒนาเว็บแอปพลิเคชันโดยใช้แพลตฟอร์ม GAS ทำให้



รูปที่ 1. ตัวอย่างแผนภูมิการจัดประเภทผลิตภัณฑ์สุขภาพ

```
const payload = JSON.stringify({
  model: model,
  messages: [
    {
      role: 'developer',
      content: `
        You are a highly strict filter assistant. Your sole purpose is to respond ONLY to questions
        clearly and explicitly about health product registration or classification.

        Rules:
        1. If the question is directly and explicitly about:
          - Regulatory classification of a health product
          - Registration process for a health product
          - Required documents, standards, or compliance for health product registration
          Then: answer concisely and accurately.

        2. If the question is even partially outside this scope,
          or the intention is unclear,
          or it involves other domains (medicine use, health advice, unrelated products, etc.),
          reply EXACTLY:
          "ขอภัยครับ!! ระบบนี้รองรับเฉพาะคำถามที่เกี่ยวข้องกับประเด็นการขึ้นทะเบียนผลิตภัณฑ์สุขภาพเท่านั้น
          กรุณาถามคำถามที่เกี่ยวข้องในประเด็นดังกล่าวเท่านั้น"

        3. Do not guess. If you are unsure whether it is strictly about
          health product registration/classification, refuse.

        You must never answer anything outside the approved scope.
      `
    }
  ]
}).trim()
```

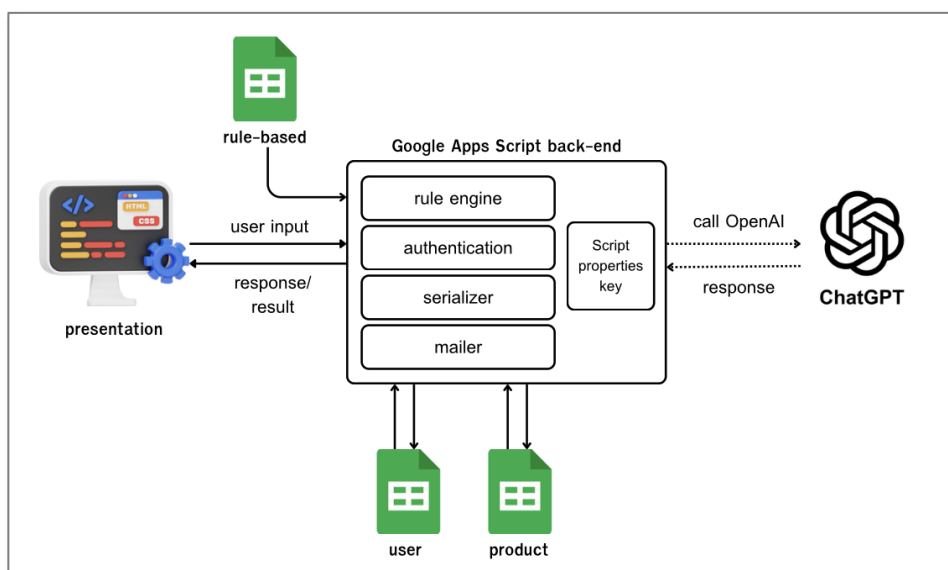
รูปที่ 2. คำสั่งในการกำหนดบทบาทของ Generative AI (system prompt) เพื่อกำหนดขอบเขตในการตอบและป้องกันการใช้งานที่ผิดวัตถุประสงค์

สามารถเก็บข้อมูลไว้ในโปรแกรม Google Sheets แล้วทำการเชื่อมข้อมูลเข้าสู่เว็บแอปพลิเคชันได้อย่างสะดวก รูปที่ 3 แสดงลักษณะสถาปัตยกรรมของระบบ

ขั้นตอนที่ 4 การพัฒนาระบบ ผู้วิจัยได้พัฒนาระบบเว็บแอปพลิเคชันโดยใช้แพลตฟอร์ม GAS และพัฒนาตามโครงร่างการออกแบบที่วางไว้ในขั้นตอนก่อนหน้า ผลการพัฒนาแสดงไว้ในภาคผนวกในภาพที่ ก ถึง จ

ขั้นตอนที่ 5 การทดสอบระบบ ผู้วิจัยแบ่งการทดสอบระบบออกเป็น 2 ส่วน

ส่วนที่ 5.1 การประเมินความถูกต้องของ rule-based system ที่สร้างขึ้น ทำโดยการเลือกผลิตภัณฑ์ที่จำหน่ายในท้องตลาดจำนวน 40 รายการมาด้วยวิธีการเลือกแบบสะดวกจากห้างสรรพสินค้า โรงพยาบาลและร้านค้า โดยรายการที่เลือกครอบคลุมกลุ่มผลิตภัณฑ์ 8 กลุ่ม ได้แก่ ยา อาหาร เครื่องมือแพทย์ เครื่องสำอาง วัตถุอันตราย วัตถุเสพติด ผลิตภัณฑ์สมุนไพร และผลิตภัณฑ์ที่ไม่ได้อยู่ในการกำกับดูแลของ อย. กลุ่มละ 5 รายการ ผลิตภัณฑ์ทั้งหมดมีการขึ้นทะเบียนถูกต้องตามกฎหมาย มี



รูปที่ 3. ลักษณะสถาปัตยกรรมของระบบ

ฉลาก/เอกสารกำกับถูกต้องตามกฎหมาย และมีการจำหน่ายจริงในท้องตลาด

ข้อมูลผลิตภัณฑ์ที่ใช้ในการจัดประเภทประกอบด้วยชื่อผลิตภัณฑ์ วิธีการใช้งาน ข้อบ่งใช้และสรรพคุณ พร้อมทั้งภาพถ่ายผลิตภัณฑ์ ผลการทดสอบเบื้องต้นพบว่า ระบบสามารถจัดประเภทได้ถูกต้อง 36 จาก 40 ตัวอย่าง (ร้อยละ 92.5) โดยพบความคลาดเคลื่อน 4 ตัวอย่าง ผู้วิจัยได้นำผลการทดสอบดังกล่าวมาใช้ในการทบทวนและปรับปรุงกฎการตัดสินใจ (rule refinement) เพื่อเพิ่มความครอบคลุมของกฎให้เหมาะสมยิ่งขึ้นกับกรณีตัวอย่างที่ศึกษา ส่งผลให้ระบบสามารถจำแนกตัวอย่างชุดข้อมูลทดสอบดังกล่าวได้ถูกต้องครบถ้วนหลังการปรับปรุง

ส่วนที่ 5.2 การประเมินประสิทธิภาพก่อนการนำไปใช้จริง โดยผู้เชี่ยวชาญจำนวน 3 ท่าน ประกอบด้วยผู้เชี่ยวชาญด้านการจัดประเภทผลิตภัณฑ์สุขภาพ 1 คนซึ่งมีประสบการณ์ที่เกี่ยวข้องมากกว่า 2 ปี ผู้เชี่ยวชาญด้านเทคโนโลยีและสารสนเทศ 1 คนที่มีประสบการณ์ที่เกี่ยวข้องมากกว่า 1 ปี หรือสำเร็จการศึกษาอย่างน้อยในระดับปริญญาโทในสาขาที่เกี่ยวข้องกับเทคโนโลยีสารสนเทศ และเจ้าหน้าที่ผู้รับผิดชอบกระบวนการจัดประเภทผลิตภัณฑ์สุขภาพ 1 คนที่มีประสบการณ์ที่เกี่ยวข้องมากกว่า 1 ปี การทดสอบทำโดยให้ผู้เชี่ยวชาญลองใช้เว็บแอปพลิเคชันและประเมินประสิทธิภาพเว็บแอปพลิเคชัน (ก่อนการนำไปใช้จริง) โดยมีหัวข้อการประเมินด้านการทำงานและประสิทธิภาพของระบบ ด้านประสบการณ์ผู้ใช้งาน และข้อเสนอแนะเพิ่มเติม แบบประเมินที่ใช้ผ่านการทดสอบหาความตรงเชิงเนื้อหาโดยคำนวณค่า Index of item objective congruence (IOC)

ขั้นตอนที่ 6 การบำรุงรักษา ผู้วิจัยปรับปรุงเว็บแอปพลิเคชันให้สอดคล้องกับผลการประเมินประสิทธิภาพก่อนการนำไปใช้จริงและจัดทำคู่มือการใช้งาน เว็บแอปพลิเคชัน

ระยะที่ 2 การประเมินประสิทธิภาพของแอปพลิเคชัน ความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชัน

การประเมินความพึงพอใจในการใช้งานเว็บแอปพลิเคชันทำในผู้รับบริการจำนวน 30 คนที่เข้ามาสอบถามข้อมูลเพื่อการขึ้นทะเบียนผลิตภัณฑ์ที่ศูนย์บริการผลิตภัณฑ์สุขภาพเบ็ดเสร็จ (one stop service center: OSSC) ของ ออย. ซึ่งถูกคัดเลือกด้วยวิธีการเลือกตัวอย่าง

แบบตามสะดวกจากผู้ที่มีคุณสมบัติดังนี้ 1) สามารถใช้งานเว็บแอปพลิเคชันได้ 2) เป็นผู้ที่มีประสบการณ์เกี่ยวกับการขึ้นทะเบียนหรือขอรับคำปรึกษาเกี่ยวกับการดำเนินการด้านกฎหมายในประเด็นของผลิตภัณฑ์สุขภาพ และ 3) เป็นผู้ที่ใช้บริการที่ศูนย์บริการผลิตภัณฑ์สุขภาพเบ็ดเสร็จ ออย. เกณฑ์การคัดเลือกตัวอย่างออกจากการวิจัย คือ ผู้ที่ไม่สามารถเข้าร่วมงานวิจัยได้จนจบกระบวนการเก็บข้อมูล

การศึกษาให้อาสาสมัครทดลองใช้งานเว็บแอปพลิเคชัน และตอบแบบประเมินความพึงพอใจในรูปแบบมาตราส่วนประมาณค่า 5 ระดับ หัวข้อการประเมินความพึงพอใจแบ่งเป็น 3 ด้าน ได้แก่ 1) ด้านประสิทธิภาพการทำงานของเว็บแอปพลิเคชัน 2) ด้านการออกแบบและจัดรูปแบบของเว็บแอปพลิเคชัน และ 3) ด้านการตอบสนองต่อความต้องการของผู้ใช้และประโยชน์ที่ได้รับ แบบประเมินดังกล่าวได้ผ่านการทดสอบหาค่า IOC โดยผู้ทรงคุณวุฒิ

ความถูกต้องของการจัดประเภท

วิธีการ

การประเมินความถูกต้องของการจัดประเภทผลิตภัณฑ์โดยเว็บแอปพลิเคชัน ทำโดยการวิเคราะห์ความสอดคล้องและความแตกต่างระหว่างผลการจัดประเภทผลิตภัณฑ์จากผู้ประเมิน 3 กลุ่ม คือ 1) rule-based system 2) generative AI ซึ่งถูกใช้ในฐานะคู่เปรียบเทียบกับประเมินศักยภาพในการนำเทคโนโลยี generative AI มาใช้ในอนาคต และ 3) ผู้เชี่ยวชาญจากกลุ่มงานให้คำปรึกษา 2 คน ผลิตภัณฑ์ที่ใช้ทดสอบ คือ ผลิตภัณฑ์ 30 รายการที่อาสาสมัครใช้ในการทดลองใช้งานเว็บแอปพลิเคชัน

การจัดประเภทโดย generative AI และผู้เชี่ยวชาญทำโดยใช้ข้อมูลของผลิตภัณฑ์ที่จัดเก็บอยู่ในฐานข้อมูลซึ่งปกปิดชื่อผลิตภัณฑ์ สรรพคุณ และชื่อเจ้าของผลิตภัณฑ์ โดยเป็นข้อมูลที่เหมือนกับการใช้ rule-based system ทั้งข้อมูลและจำนวนข้อมูล ในการจัดประเภทผลิตภัณฑ์โดย generative AI ผู้วิจัยวางโครงสร้างของคำสั่ง (prompt) ให้มีโครงสร้างของข้อมูลสอดคล้องกับของ rule-based system ดังนี้ “(1) รูปแบบผลิตภัณฑ์ + (2) ข้อบ่งใช้ + (3) กลุ่มประชากรที่ใช้ + (4) ลักษณะสุขภาพของประชากร + (5) กลไกการออกฤทธิ์ + (6) สารออกฤทธิ์หลัก + (7) แหล่งที่มา + (8) ประวัติการใช้ + (9) เป็นวัตถุที่ รมต. ประกาศเป็นยา แต่ได้รับยกเว้น” ตัวอย่างเช่น ผลิตภัณฑ์ A

มีจุดประสงค์ใช้เพื่อทำความเข้าใจการแปลเป็น prompt ว่า “(1) ผลิตภัณฑ์สำเร็จรูป (2) ใช้เพื่อทำความเข้าใจ (3) โดยมีเป้าหมายใช้ในมนุษย์ จัดเป็นผลิตภัณฑ์ใด” เป็นต้น

การวิเคราะห์ข้อมูล

ความสอดคล้องของการจัดประเภทผลิตภัณฑ์วิเคราะห์ด้วย Fleiss' Kappa ซึ่งใช้วัดความสอดคล้องระหว่างวิธีการจัดประเภทมากกว่า 2 วิธี และใช้ Pairwise Cohen's Kappa วัดความสอดคล้องระหว่างวิธีการจัดประเภทรายคู่ ส่วนความแตกต่างของผลการจัดประเภทผลิตภัณฑ์ระหว่างวิธีการวิเคราะห์ด้วย Bowker's test of symmetry ที่พัฒนามาจาก McNemar test เนื่องจากข้อมูลมีลักษณะเป็นข้อมูลกลุ่ม (nominal data) ที่ไม่เป็นอิสระระหว่างตัวอย่าง จำนวนข้อมูลไม่มาก (30 ตัวอย่าง) และมีหมวดหมู่ของประเภทผลิตภัณฑ์หลายกลุ่ม ทำให้พบว่าบางกลุ่มมีความถี่น้อยมาก จึงเลือกใช้สถิติดังกล่าวเพื่อลดปัญหาด้านเสถียรภาพของเมทริกซ์ความแปรปรวนร่วม

ความสอดคล้องในการจัดประเภทระหว่างวิธีการวิเคราะห์ด้วย Weighted Cohen's Kappa การประเมินประสิทธิภาพของระบบในการจัดประเภทผลิตภัณฑ์ ใช้ตัวชี้วัด accuracy, precision, recall, F-1 score โดยค่า overall accuracy (micro-averaged) ใช้วัดประสิทธิภาพในการจัดประเภทโดยรวม ค่า macro-average ใช้วัดประสิทธิภาพโดยเฉลี่ยในภาพรวมของทุกกลุ่มประเภทผลิตภัณฑ์โดยไม่ถ่วงน้ำหนักตามจำนวนข้อมูล ส่วนค่าเฉลี่ยแบบ weighted average ใช้วัดประสิทธิภาพโดยรวมของระบบเมื่อคำนึงถึงสัดส่วนจำนวนข้อมูลในแต่ละประเภทผลิตภัณฑ์

การแปลผล Kappa ใช้เกณฑ์ที่ Landis และ Koch (26) แนะนำ คือ ค่า Kappa (K) ที่ < 0.00 , $0.00-0.20$, $0.21-0.40$, $0.41-0.60$, $0.61-0.80$, และ $0.81-1.00$ บ่งบอกขนาดความสอดคล้องที่แย่มาก, น้อย, พอใช้, ปานกลาง, ดี, ดีมาก/ค่อนข้างสมบูรณ์ ตามลำดับ การแปลผลค่า Cronbach's alpha ใช้เกณฑ์ของ Nunnally และ Bernstein (28) ที่ระบุว่าค่า Cronbach's $\alpha < 0.60$, $0.60-0.69$, $0.70-0.79$, $0.80-0.89$ และ ≥ 0.9 หมายถึง ต่ำ/ไม่ดี, น่าสงสัย, ยอมรับได้, ดี และดีมาก/สูงมาก ตามลำดับ

การศึกษากำหนดระดับนัยสำคัญทางสถิติที่ $\alpha = 0.05$ การวิเคราะห์ข้อมูลทั้งหมดในการศึกษาดำเนินการโดยใช้โปรแกรม R version 2025.09.2+418

คุณภาพของคำแนะนำจาก Generative AI

การประเมินคุณภาพของคำแนะนำจาก Generative AI ประยุกต์ใช้วิธีการศึกษาของนันทพงศ์ บุญฤทธิ์และคณะ (25) โดยผู้วิจัยสัมภาษณ์เจ้าหน้าที่กลุ่มให้คำปรึกษา จำนวน 4 คนเพื่อรวบรวมประเด็นคำถามที่มักพบในการปฏิบัติงานจริง (FAQ) นอกจากนี้ยังได้รวบรวมคำถามที่พบจากการให้บริการผ่านระบบ E-Consult และทางโทรศัพท์ในช่วงเดือน มกราคม - มิถุนายน 2568 จากนั้นนำข้อมูลที่ได้อาวิเคราะห์และคัดเลือกเฉพาะประเด็นที่ผู้รับบริการหลายรายกล่าวถึง ประเด็นหลักที่พบในการปฏิบัติงาน คือ 1) ขอบเขตการกำกับดูแลตามกลุ่มผู้ใช้งาน 2) แนวทางและขั้นตอนการขึ้นทะเบียนตามกฎหมาย 3) การจัดประเภทผลิตภัณฑ์ตามนิยามทางกฎหมาย 4) ขอบเขตการกำกับดูแลจากวัตถุประสงค์การใช้หรือคำกล่าวอ้างของผลิตภัณฑ์ และ 5) ข้อกำหนดด้านการนำเข้าและการปฏิบัติตามข้อกำหนด ชุดคำถามดังกล่าวเป็นตัวแทนคำถามที่มุ่งสะท้อนคำถามที่พบในการปฏิบัติงานจริงและเหมาะสมสำหรับการประเมินประสิทธิภาพของระบบในระยะเบื้องต้น การกำหนดคำถามที่เจาะจงทำให้ผู้เชี่ยวชาญสามารถประเมินคุณภาพของคำตอบจาก generative AI ได้อย่างเหมาะสมในเชิงลึก อีกทั้งยังช่วยควบคุมภาระในการประเมินไม่ให้เกิดความล้าจากการประเมิน (rater fatigue) ซึ่งอาจส่งผลต่อความสม่ำเสมอและความน่าเชื่อถือของคะแนนประเมิน คำถามที่ได้มีจำนวนทั้งสิ้น 5 คำถาม ดังนี้ FAQ 1: แชมพูสำหรับสัตว์จัดเป็นผลิตภัณฑ์ในการกำกับดูแลของ ออย. หรือไม่; FAQ 2: ลูกประคบสมุนไพรต้องยื่นขึ้นทะเบียนที่ไหน; FAQ 3: เครื่องนวดคลายกล้ามเนื้อจัดเป็นเครื่องมือแพทย์หรือไม่; FAQ 4: ชุดตรวจสอคุณภาพน้ำในสิ่งแวดล้อม เป็นผลิตภัณฑ์ที่จัดอยู่ภายใต้การดูแลของ ออย.หรือไม่; FAQ 5: การนำเข้าวัตถุดิบเครื่องสำอางต้องดำเนินการขึ้นทะเบียนหรือไม่อย่างไร

จากนั้นผู้วิจัยป้อนคำถาม (prompt) ไปยังส่วนให้คำแนะนำโดย generative AI ที่ฝังอยู่ในเว็บแอปพลิเคชันซ้ำในบริบทเดิมทันที (same-session prompting) เป็นจำนวน 3 ครั้ง ซึ่งสะท้อนลักษณะการใช้งานจริงของผู้ใช้งานที่มักมีการถามต่อเนื่อง ถามซ้ำ และขยายความในประเด็นเดิม หลังจากนั้นเก็บรวบรวมผลและนำไปให้ผู้เชี่ยวชาญจากกลุ่มงานให้คำปรึกษา สนบ. จำนวน 2 คน ประเมินความพึงพอใจในคำตอบที่ได้ ผู้เชี่ยวชาญทำงานในกลุ่มงาน

ให้คำปรึกษาฯ มีประสบการณ์ในการทำงานให้คำปรึกษาผลิตภัณฑ์สุขภาพกับผู้ประกอบการอย่างน้อย 1 ปีและเป็นเภสัชกรระดับปฏิบัติการขึ้นไป การประเมินใช้แบบวัดมาตราส่วนประมาณค่า 5 ระดับ (Likert scale) โดยประเมินคำตอบที่ได้ใน 3 มิติ คือ มิติด้านความเหมาะสมของภาษามิติด้านความถูกต้องของข้อมูล และมิติด้านความสามารถในการนำไปปฏิบัติตามได้

ความแตกต่างของคะแนนความพึงพอใจในการประเมิน FAQ แบบรายรอบ prompt วิเคราะห์โดยใช้สถิติ Friedman test

ผลการวิจัย

ความถูกต้องของการจัดประเภท

การเปรียบเทียบผลการการจัดประเภทผลิตภัณฑ์สุขภาพ 30 รายการของอาสาสมัครในงานวิจัยโดยการประมวลผลของ rule-based system, generative AI และผู้เชี่ยวชาญที่เจ้าหน้าที่กลุ่มงานให้คำปรึกษาฯ สนบ. จำนวน 2 คน (ผลของการจัดประเภทมีทั้งหมด 4 ชุด) มีรายละเอียด ดังนี้ ความสอดคล้องของผลการจัดประเภททั้ง 4 ชุดมีค่า Fleiss' Kappa = 0.862 (95% CI: 0.795 – 0.929) ซึ่งถือว่ามีความสอดคล้องอยู่ในระดับดีมาก/ค่อนข้างสมบูรณ์ (almost perfect)

เมื่อเปรียบเทียบความสอดคล้องของผลการจัดประเภทผลิตภัณฑ์สุขภาพทั้ง 4 ชุดเป็นรายคู่โดยใช้ Pairwise Cohen's Kappa พร้อมทั้งวิเคราะห์หาความแตกต่างโดยใช้ Bowker's test of symmetry ได้ผลการวิเคราะห์ดังแสดงในตารางที่ 1

จากตารางที่ 1 การจัดประเภทของผู้เชี่ยวชาญ 1 เทียบกับผู้เชี่ยวชาญ 2 พบว่ามีค่า $K = 1.00$ แสดงว่าผู้เชี่ยวชาญทั้งสองจัดประเภทผลิตภัณฑ์ได้สอดคล้องกันทั้งหมด ส่วนผลการจัดประเภทของผู้เชี่ยวชาญท่านที่ 1

หรือ 2 เทียบกับ rule-based system พบว่ามีค่า $K = 0.918$ แสดงว่า rule-based system มีการจัดประเภทผลิตภัณฑ์ที่สอดคล้องกับผู้เชี่ยวชาญทั้งสอง อยู่ในระดับดีมาก/ค่อนข้างสมบูรณ์

ส่วนผลการจัดประเภทของผู้เชี่ยวชาญท่านที่ 1 หรือ 2 เทียบกับ generative AI พบว่ามีค่า $K = 0.751$ แสดงว่า generative AI มีการจัดประเภทผลิตภัณฑ์ที่สอดคล้องกับผู้เชี่ยวชาญทั้งสองในระดับดี ขณะที่ผลการจัดประเภทของ rule-based system เทียบกับ generative AI พบว่ามีค่า $K = 0.834$ แสดงว่า มีความสอดคล้องอยู่ในระดับดีมาก/ค่อนข้างสมบูรณ์ (ตารางที่ 1)

การวิเคราะห์ความแตกต่างของผลการจัดประเภทโดยผู้เชี่ยวชาญกับ rule-based system ได้ค่า $P = 0.368$ ขณะที่ความแตกต่างระหว่างผลของผู้เชี่ยวชาญกับ generative AI มีค่า $P = 0.199$ ความแตกต่างระหว่างผลของ rule-based system กับ generative AI ได้ค่า $P = 0.135$ แสดงให้เห็นว่า ผลของการจัดประเภทผลิตภัณฑ์ด้วยวิธีการที่ใช้ในการวิจัยและผู้เชี่ยวชาญไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ ($P > 0.05$) จึงสามารถสรุปได้ว่า การจัดประเภทโดย rule-based system, generative AI และผู้เชี่ยวชาญมีความสอดคล้องกันในระดับดีมาก/ค่อนข้างสมบูรณ์

ประสิทธิภาพของเครื่องมือในการจัดประเภท

การทดสอบประสิทธิภาพของเว็บแอปพลิเคชันที่ใช้ข้อมูลผลิตภัณฑ์สุขภาพของอาสาสมัคร 30 รายการที่ทดลองใช้เว็บแอปพลิเคชัน อย่างไรก็ตาม ตัวอย่างทั้ง 30 รายการไม่มีผลิตภัณฑ์ที่เป็น ยาเสพติด วัตถุอันตราย ผลิตภัณฑ์สมุนไพร และกลุ่มที่ไม่อาจจัดประเภทได้ จึงไม่มีผลการทดสอบผลิตภัณฑ์ในกลุ่มนี้ ผลการทดสอบประสิทธิภาพของเว็บแอปพลิเคชันแสดงอยู่ในตารางที่ 2

ตารางที่ 1. ผลการประเมินความสอดคล้องในการจัดประเภทผลิตภัณฑ์ของแต่ละการประเมิน

ที่	คู่เปรียบเทียบ	Kappa	95% CI	χ^2	P
1	ผู้เชี่ยวชาญ 1 เทียบ ผู้เชี่ยวชาญ 2	1.00	0.837 - 1.000	0	1.000
2	ผู้เชี่ยวชาญ 1 เทียบ rule-based system	0.918	0.757 - 1.000	2	0.368
3	ผู้เชี่ยวชาญ 2 เทียบ rule-based system	0.918	0.757 - 1.000	2	0.368
4	ผู้เชี่ยวชาญ 1 เทียบ generative AI	0.751	0.587 - 0.916	6	0.199
5	ผู้เชี่ยวชาญ 2 เทียบ generative AI	0.751	0.587 - 0.916	6	0.199
6	rule-based system เทียบ generative AI	0.834	0.670 - 0.999	4	0.135

พบว่า rule-based system สามารถจำแนกประเภทผลิตภัณฑ์เครื่องสำอาง กลุ่มที่ต้องติดต่อเจ้าหน้าที่ ยา และอาหารได้อย่างถูกต้องโดยมีค่า precision, recall และ F1-score เท่ากับ 1.00 ส่วนในเครื่องมือแพทย์ rule-based system มีค่า precision เท่ากับ 0.670 และ F1-score เท่ากับ 0.800

generative AI สามารถจำแนกประเภทผลิตภัณฑ์ได้อย่างถูกต้องในกลุ่มเครื่องสำอาง กลุ่มที่ต้องติดต่อเจ้าหน้าที่ และอาหาร โดยมีค่า precision, recall และ F1-score เท่ากับ 1.00 ขณะที่ในกลุ่มเครื่องมือแพทย์ มีค่า precision และ F1-score ที่ 0.670 และ 0.800 ตามลำดับ ในกลุ่มผลิตภัณฑ์ยามีค่า precision และ F1-score เท่ากับ 0.750 และ 0.860 ตามลำดับ (ตารางที่ 2)

ผลการเปรียบเทียบประสิทธิภาพของ rule-based system และ generative AI ในการจัดประเภทผลิตภัณฑ์สุขภาพ พิจารณาได้จากค่า overall accuracy (micro-averaged) พบว่ามีค่าเท่ากับ 0.933 และ 0.800 ตามลำดับ (ตารางที่ 2) แสดงให้เห็นว่า rule-based system มีประสิทธิภาพในการจัดประเภทผลิตภัณฑ์สุขภาพโดยรวมสูงกว่า generative AI เมื่อพิจารณาจากข้อมูลเฉพาะในหมวดเครื่องมือแพทย์ เครื่องสำอาง กลุ่มที่ต้องติดต่อเจ้าหน้าที่ ยา และอาหาร

เมื่อพิจารณาค่าเฉลี่ยแบบ macro-average พบว่า rule-based system มีค่า precision, recall และ F1-score เท่ากับ 0.833, 0.875 และ 0.971 ตามลำดับ ขณะที่ generative AI มีค่าเท่ากับ 0.631, 0.625 และ 0.931 ตามลำดับ สะท้อนให้เห็นว่า rule-based system มีผลการ

ประเมินเฉลี่ยโดยรวมของทุกกลุ่มประเภทผลิตภัณฑ์สูงกว่า generative AI เมื่อพิจารณาจากข้อมูลเฉพาะในหมวดเครื่องมือแพทย์ เครื่องสำอาง กลุ่มที่ต้องติดต่อเจ้าหน้าที่ ยา และอาหาร

เมื่อพิจารณาค่าเฉลี่ยแบบ weighted average พบว่า rule-based system มีค่า precision, recall และ F1-score เท่ากับ 0.976, 0.933 และ 0.986 ตามลำดับ ขณะที่ generative AI มีค่าเท่ากับ 0.840, 0.800 และ 0.948 ตามลำดับ แสดงให้เห็นว่า rule-based system สามารถจำแนกประเภทผลิตภัณฑ์เฉพาะที่พบในชุดข้อมูลทดสอบได้อย่างมีประสิทธิภาพโดยรวมสูงกว่า generative AI อย่างไรก็ตามเมื่อพิจารณาผลการทดสอบด้วย Bowker's test of symmetry พบว่า วิธีการทั้งสองมีประสิทธิภาพไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

คุณภาพของคำแนะนำที่ได้จาก generative AI

ความพึงพอใจต่อคำตอบที่ได้

ผู้เชี่ยวชาญที่เป็นเจ้าหน้าที่กลุ่มให้คำปรึกษาฯ 2 ท่านให้คะแนนความพึงพอใจต่อคำแนะนำที่ได้จาก generative AI ในด้านความเหมาะสมของภาษา ด้านความถูกต้องของข้อมูล และด้านความสามารถในการนำไปใช้ได้จริง โดยมีค่าเฉลี่ยเท่ากับ 3.23 ± 0.94 , 3.20 ± 1.19 และ 3.17 ± 1.18 ตามลำดับ (จากคะแนนเต็ม 5) ซึ่งถือว่าเป็นความพึงพอใจในระดับปานกลาง

คะแนนความพึงพอใจจำแนกตามคำถามที่ใช้ทดสอบแสดงอยู่ในตารางที่ 3 ความพึงพอใจในด้านความเหมาะสมของภาษาต่อคำตอบในคำถามทั้งหมดอยู่ในระดับ

ตารางที่ 2. ผลการคำนวณค่า accuracy, precision, recall และ F1-score เปรียบเทียบระหว่าง rule-based system และ generative AI

ประเภทผลิตภัณฑ์	rule-based system			generative AI		
	precision	recall	F1-score	precision	recall	F1-score
เครื่องมือแพทย์	0.670	1.000	0.800	0.670	1.000	0.800
เครื่องสำอาง	1.000	1.000	1.000	1.000	1.000	1.000
กลุ่มที่ต้องติดต่อเจ้าหน้าที่	1.000	1.000	1.000	1.000	1.000	1.000
ยา	1.000	1.000	1.000	0.750	1.000	0.860
อาหาร	1.000	1.000	1.000	1.000	1.000	1.000
overall accuracy (micro-averaged)			0.933			0.800
macro-average	0.833	0.875	0.971	0.631	0.625	0.931
weighted average	0.976	0.933	0.986	0.840	0.800	0.948

ตารางที่ 3. ความพึงพอใจของเจ้าหน้าที่กลุ่มให้คำปรึกษา 2 ท่านต่อคำแนะนำที่ได้จาก generative AI จำแนกตามคำถามที่ใช้ทดสอบ

คำถาม ¹ (FAQ)	ด้านความเหมาะสมของภาษา			ด้านความถูกต้องของข้อมูล			ด้านความสามารถในการนำไปใช้ได้จริง		
	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)
1	4.17	0.41	4 (4-4)	3.50	1.22	4 (2.5-4)	3.33	1.21	4 (2.25-4)
2	3.17	1.17	3 (2.25-3.75)	2.33	1.75	1.5 (1-3.5)	2.33	1.75	1.5 (1-3.5)
3	3.17	0.75	3 (3-3.75)	3.67	0.82	4 (4-4)	3.67	0.82	4 (4-4)
4	3.00	0.89	3 (2.25-3.75)	3.33	0.82	3.5 (3-4)	3.33	0.82	3.5 (3-4)
5	2.67	0.82	2.5 (2-3)	1.67	0.98	3.5 (2.25-4)	3.17	0.98	3.5 (2.25-4)

1: คำถามที่ใช้ทดสอบ FAQ 1: แชมพูสำหรับสัตว์จัดเป็นผลิตภัณฑ์ในการกำกับดูแลของ ออย. หรือไม่; FAQ 2: ลูกประคบสมุนไพรต้องยี่นขึ้นทะเบียนที่ไหน; FAQ 3: เครื่องนวดคลายกล้ามเนื้อจัดเป็นเครื่องมือแพทย์หรือไม่; FAQ 4: ชุดตรวจสอบคุณภาพน้ำในสิ่งแวดล้อม เป็นผลิตภัณฑ์ที่จัดอยู่ภายใต้การดูแลของ ออย.หรือไม่; FAQ 5: การนำเข้าวัตถุอันตรายต้องดำเนินการขึ้นทะเบียนหรือไม่อย่างไร; 2: IQR: interquartile range (พิสัยของควอไทล์)

ปานกลางถึงน้อย ยกเว้นในคำถามข้อแรก (แชมพูสำหรับสัตว์จัดเป็นผลิตภัณฑ์ในการกำกับดูแลของ ออย. หรือไม่) ที่มีความพึงพอใจในระดับที่ดี (4.17±0.41 จากคะแนนเต็ม 5) ความพึงพอใจในด้านความถูกต้องของคำตอบจากคำถามทั้งหมดที่ใช้ทดสอบอยู่ในระดับที่ปานกลางถึงน้อย โดยเฉพาะในคำถามข้อที่ 5 (การนำเข้าวัตถุอันตรายต้องดำเนินการขึ้นทะเบียนหรือไม่อย่างไร) ที่มีความพึงพอใจในระดับที่น้อยมาก (1.67±0.98 จากคะแนนเต็ม 5) ความพึงพอใจในด้านความสามารถในการนำไปใช้ได้จริงของคำตอบจากคำถามทั้งหมดที่ใช้ทดสอบอยู่ในระดับที่ปานกลางถึงน้อย (ตารางที่ 3)

ความพึงพอใจในแต่ละรอบของการถาม

การวิจัยใช้คำสั่งเพื่อขอคำตอบจาก generative AI จำนวน 3 รอบในแต่ละคำถาม ผู้เชี่ยวชาญที่เป็นเจ้าหน้าที่กลุ่มให้คำปรึกษา 2 ท่านให้คะแนนความพึงพอใจต่อคำตอบที่ได้ในแต่ละรอบดังแสดงในตารางที่ 4 การวิเคราะห์

ด้วย Friedman test พบว่า คะแนนความพึงพอใจต่อคำตอบที่ได้ในแต่ละรอบมีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติในด้านความถูกต้องของข้อมูลและความสามารถในการนำไปใช้ได้จริง ($P = 0.005$ และ 0.005 ตามลำดับ) โดยความพึงพอใจต่อคำตอบที่ได้ในรอบที่ 3 มีระดับที่สูงกว่าในรอบที่ 1 และ 2 (ตารางที่ 4) ขณะที่ด้านความเหมาะสมของภาษาไม่พบความแตกต่างของความพึงพอใจอย่างมีนัยสำคัญทางสถิติระหว่างรอบการประเมิน ($P = 0.066$)

ความสอดคล้องกันระหว่างผู้เชี่ยวชาญ

ผู้วิจัยตรวจสอบความคล้อยกันในการให้คะแนนความพึงพอใจต่อคำแนะนำจาก generative AI ระหว่างผู้เชี่ยวชาญ 2 ท่าน พบว่า การประเมินความพึงพอใจในประเด็นด้านความเหมาะสมของภาษา ด้านความถูกต้องของข้อมูล และด้านความสามารถในการนำไปใช้ได้จริง มีค่า Kappa เท่ากับ 0.252 (95%CI: -0.188 – 0.531), 0.571 (95%CI: 0.110 – 0.808) และ 0.498 (95%CI: 0.000 –

ตารางที่ 4. ความพึงพอใจของผู้เชี่ยวชาญ 2 ท่านต่อคำแนะนำที่ได้จาก generative AI ในแต่ละรอบของการถาม

รอบที่	ด้านความเหมาะสมของภาษา			ด้านความถูกต้องของข้อมูล			ด้านความสามารถในการนำไปใช้ได้จริง		
	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)	ค่าเฉลี่ย	SD	มัธยฐาน (IQR)
1	2.90	0.88	3 (2-3.75)	3.00	1.15	3.5 (2-4)	2.80	1.03	3 (2-3.75)
2	3.10	0.88	3 (2.25-4)	2.50	1.18	2 (2-3.75)	2.60	1.26	2 (2-4)
3	3.70	0.95	4 (3-4)	4.10	0.57	4 (4-4)	4.10	0.57	4 (4-4)
Friedman's $\chi^2 = 5.43$, $P=0.066$			Friedman's $\chi^2 = 10.70$, $P=0.005$			Friedman's $\chi^2 = 10.60$, $P=0.005$			

หมายเหตุ: ค่าเฉลี่ยและมัธยฐานของความพึงพอใจในแต่ละรอบคำนวณมาจากความพึงพอใจต่อคำตอบของคำถามทั้ง 5 ข้อ

0.788) ตามลำดับ นั่นคือ มีความสอดคล้องในระดับพอใช้ปานกลาง และปานกลาง ตามลำดับ การวิเคราะห์ความแตกต่างของความพึงพอใจระหว่างผู้เชี่ยวชาญทั้งสองโดยใช้สถิติ Wilcoxon signed rank test พบว่า ไม่มีความแตกต่างกันอย่างมีนัยสำคัญทางสถิติในทุกด้านของการประเมิน ($P=0.114, 0.132$ และ 0.169 ตามลำดับ)

ดัชนีคุณภาพคำตอบรวม

ผู้วิจัยคำนวณดัชนีคุณภาพคำตอบรวม (quality index: QI) จากค่าเฉลี่ยของคะแนนความพึงพอใจทั้งสามด้านต่อคำตอบของ generative AI ต่อคำถามที่พบบ่อยจำนวน 5 ประเด็น รวมทั้งสิ้น 30 คำตอบ ได้ค่า ค่าเฉลี่ย QI = 3.20 ± 1.01 (พิสัย: 1.33 - 5.00) จากคะแนนเต็ม 5 แสดงให้เห็นว่า คุณภาพของคำตอบโดยรวมในเฉพาะประเด็นที่ศึกษาอยู่ในระดับปานกลาง ค่าพิสัยบ่งชี้ว่าคำตอบมีความแปรปรวนค่อนข้างสูง โดยมีทั้งคำตอบที่มีคุณภาพอยู่ในระดับต่ำและระดับสูงอยู่ปะปนกัน ซึ่งแสดงให้เห็นว่า generative AI ยังให้คุณภาพของคำตอบไม่สม่ำเสมอในบริบทของคำถามที่ศึกษา

ความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชัน

ตัวอย่างผู้รับบริการจากศูนย์บริการสุขภาพเบ็ดเสร็จ ออย. จำนวน 30 รายได้ทดลองใช้งานเว็บแอปพลิเคชัน ตัวอย่างเป็นหญิง 24 ราย มีอายุอยู่ในช่วง 20-29 ปี, 30-39 ปี และ 40-49 ปี จำนวน 12, 15 และ 3 คนตามลำดับ ตัวอย่างประกอบธุรกิจส่วนตัว 12 รายและเป็นพนักงานบริษัท 18 ราย ประสบการณ์การทำงานด้านทะเบียนผลิตภัณฑ์ คือ น้อยกว่า 1 ปี, 1-3 ปี และมากกว่า 3 ปี จำนวน 17, 5 และ 8 ราย ตามลำดับ

ค่า Cronbach's alpha ของแบบประเมินความพึงพอใจด้านประสิทธิภาพการทำงานของเว็บแอปพลิเคชันด้านการออกแบบและจัดรูปแบบของเว็บแอปพลิเคชันและด้านการตอบสนองต่อความต้องการของผู้ใช้และประโยชน์ที่ได้รับ เท่ากับ 0.82, 0.85 และ 0.91 ตามลำดับ ซึ่งถือว่ามีความเที่ยงในระดับที่ดีถึงดีมาก ตารางที่ 5 แสดงค่าเฉลี่ยของความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชันดังกล่าว ผู้ใช้งานเว็บแอปพลิเคชันมีความพึงพอใจอย่างมากทั้งในด้านประสิทธิภาพการทำงาน การออกแบบและจัดรูปแบบ

ตารางที่ 5. ความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชันเพื่อการจัดประเภทผลิตภัณฑ์ (N=30)

ความพึงพอใจ	ค่าเฉลี่ย	SD	ต่ำสุด	สูงสุด
1. ด้านประสิทธิภาพการทำงานของเว็บแอปพลิเคชัน	4.38	0.60	3	5
1.1 เว็บแอปพลิเคชันมีเมนูการใช้งานครบถ้วนและการนำทางที่เหมาะสม	4.40	0.67	3	5
1.2 การประมวลผลของระบบมีประสิทธิภาพ	4.23	0.77	2	5
1.3 สามารถใช้งานผ่านอุปกรณ์หลากหลายชนิด และแสดงผลได้อย่างถูกต้อง	4.50	0.63	3	5
2. ด้านการออกแบบและจัดรูปแบบของเว็บแอปพลิเคชัน	4.42	0.51	3	5
2.1 เว็บแอปพลิเคชันมีความสวยงาม จัดวางองค์ประกอบได้ดี	4.37	0.72	3	5
2.2 ตัวอักษรในเว็บแอปพลิเคชัน มีรูปแบบตัวอักษร ขนาด และสี ชัดเจนอ่านได้ง่าย	4.67	0.55	3	5
2.3 ปุ่มกดต่าง ๆ มีรูปแบบ ขนาดและสีชัดเจน ใช้งานได้ง่าย	4.47	0.63	3	5
2.4 ข้อความโต้ตอบจากระบบถึงผู้ใช้งานมีความเหมาะสม	4.33	0.71	3	5
2.5 การใช้คำศัพท์ เนื้อหาหรือรูปประโยคอธิบายต่าง ๆ ในเว็บแอปพลิเคชัน สามารถเข้าใจได้ง่าย	4.27	0.64	3	5
3. ด้านการตอบสนองต่อความต้องการของผู้ใช้และประโยชน์ที่ได้รับ	4.47	0.66	3	5
3.1 เว็บแอปพลิเคชันมีส่วนเตรียมความพร้อมก่อนการยื่นขึ้นทะเบียนผลิตภัณฑ์ หรือเตรียมเอกสาร	4.27	0.78	3	5
3.2 เว็บแอปพลิเคชันเป็นเครื่องมือที่ช่วยในการเรียนรู้เบื้องต้นเกี่ยวกับการขึ้นทะเบียนผลิตภัณฑ์สุขภาพได้ด้วยตนเอง	4.60	0.62	3	5
3.3 ผลลัพธ์การจัดประเภทผลิตภัณฑ์สุขภาพจากเว็บแอปพลิเคชันนำไปใช้ประโยชน์ได้	4.53	0.73	3	5

ตารางที่ 6. ข้อเสนอแนะเพิ่มเติมจากผู้ใช้งานเว็บแอปพลิเคชันเพื่อการจัดประเภทผลิตภัณฑ์ (N=30)

หัวข้อ	ข้อเสนอแนะ
ด้านการใช้งาน	- เครื่องมือนี้มีประโยชน์ เหมาะกับผู้ใช้งานที่เริ่มต้นศึกษาข้อมูล เข้าใช้ได้ง่าย - ผู้ใช้ต้องสามารถเลือกข้อมูลได้อย่างถูกต้อง เนื่องจากการจัดประเภทผลิตภัณฑ์โดยใช้รูปแบบตัวเลือก (radio) มิเช่นนั้นจะส่งผลต่อผลลัพธ์ - เครื่องมือนี้อาจไม่ตอบโจทย์สำหรับผู้ที่ยอมรับค่าปรึกษาเชิงลึกที่มากกว่าการจัดประเภท การเข้าพบเจ้าหน้าที่ทำให้บรรลุเป้าหมายได้มากกว่า - เครื่องมือปัญญาประดิษฐ์ที่ฝังอยู่ในเว็บแอปพลิเคชันมีการประมวลผลที่ช้าในบางช่วง
ด้านความเข้าใจในเนื้อหา	- คำศัพท์เฉพาะทางและศัพท์ตามกฎหมายเข้าใจได้ยาก จำเป็นต้องศึกษาข้อมูลเพิ่มเติม - ควรเลือกใช้คำหรือนิยามที่กระชับ เข้าใจได้ง่าย
ด้านความสมบูรณ์ของฟังก์ชัน	- ควรพัฒนาเครื่องมือให้สามารถจัดประเภทผลิตภัณฑ์ระดับหมวดหมู่ย่อยได้ - ควรพัฒนาระบบให้สามารถตรวจสอบความถูกต้องและเหมาะสมของช่องกรอกข้อมูลการกล่าวอ้างหรือสรรพคุณ รวมถึงช่องชื่อผลิตภัณฑ์
ด้านการสนับสนุนความรู้	- ควรมีลิงก์ไปยังแหล่งข้อมูลหรือเว็บไซต์ที่ อ ย. เชื่อถือเพื่อให้สามารถนำข้อมูลมาใช้ในการประกอบการขึ้นทะเบียนได้ เช่น แหล่งข้อมูลของส่วนประกอบที่น่าเชื่อถือ หรือแนวทางการตั้งชื่อผลิตภัณฑ์ เป็นต้น - ควรมีข้อมูลติดต่อเจ้าหน้าที่หรือหน่วยงานที่เกี่ยวข้องเพิ่มเติม
ด้านการออกแบบ	- ควรปรับปรุงหน้าเว็บแอปพลิเคชันให้มีสีสันสดใส - ควรปรับการแจ้งผลในอีเมลให้มีสีสันที่เด่นชัด และเพิ่มช่องทางการติดต่อเจ้าหน้าที่ - ควรปรับให้ส่วนที่เป็น generative AI ให้มีความโดดเด่นขึ้น เนื่องจากเป็นจุดเด่นของระบบและมีประโยชน์มาก

ของเว็บแอปพลิเคชัน และการตอบสนองต่อความต้องการของผู้ใช้และประโยชน์ที่ได้รับ โดยการประเมินในทั้ง 11 ประเด็น (ตารางที่ 5) ได้คะแนนมากกว่า 4 (จากคะแนนเต็ม 5) ผู้ใช้งานมีข้อเสนอแนะต่อระบบดังแสดงในตารางที่ 6

การอภิปรายผล

งานวิจัยนี้พัฒนาเว็บแอปพลิเคชันที่ใช้ในการจัดประเภทผลิตภัณฑ์สุขภาพที่อยู่ภายใต้การกำกับดูแลของอย. โดยออกแบบให้มีทั้งส่วนของระบบประมวลผลแบบ rule-based system สำหรับการตัดสินใจเชิงตรรกะตามหลักกฎหมายและส่วนของ generative AI ในรูปแบบ LLM เพื่อให้คำแนะนำเบื้องต้นเกี่ยวกับกระบวนการขึ้นทะเบียนผลิตภัณฑ์ ทั้งนี้เครื่องมือที่พัฒนาขึ้นนี้มีเป้าหมาย คือ เพื่อสนับสนุนผู้ประกอบการและนักวิจัยรายใหม่ให้สามารถเข้าใจและจัดประเภทผลิตภัณฑ์สุขภาพได้ด้วยตนเอง

ในด้านความปลอดภัยของข้อมูล ระบบมีการจัดเก็บข้อมูลของผู้ใช้งานบางส่วน เช่น ชื่อผู้ใช้งาน

รหัสผ่าน และข้อมูลผลิตภัณฑ์ภายใต้ระบบการยืนยันตัวตน (authentication) โดยจำกัดสิทธิ์การเข้าถึงข้อมูลไว้ภายใต้บัญชีของผู้วิจัยเท่านั้น อย่างไรก็ตาม เนื่องจากระบบมีการเชื่อมต่อกับ ChatGPT API ซึ่งเป็นบริการบน third-party cloud จึงมีความเป็นไปได้ที่ข้อมูลที่ผู้ใช้งานป้อนเข้าสู่ระบบอาจถูกส่งไปประมวลผลภายนอก ดังนั้นในการพัฒนาในอนาคตควรมีการแสดงคำเตือนแก่ผู้ใช้งานเพิ่มเติม และพัฒนา generative AI ให้สอดคล้องกับข้อกำหนดด้านความปลอดภัยของหน่วยงาน

ผลการจัดประเภทผลิตภัณฑ์สุขภาพระหว่าง rule-based system ผู้เชี่ยวชาญ 2 คน และ generative AI พบว่ามีความสอดคล้องกันสูง (Fleiss' Kappa = 0.862) และไม่มีพบความแตกต่างกันอย่างมีนัยสำคัญทางสถิติ แสดงให้เห็นว่า rule-based system และ generative AI มีความสามารถในการจัดประเภทผลิตภัณฑ์ไม่แตกต่างจากผู้เชี่ยวชาญ อย่างไรก็ตามผลการจัดประเภทโดยผู้เชี่ยวชาญมีความสอดคล้องกับผลจาก จาก rule-based system ($K = 0.918$) มากกว่าผลการจัดประเภทโดย generative AI ($K = 0.751$)

เพราะ rule-based system อาศัยตรรกะที่อ้างอิงตาม กฎหมายและเงื่อนไขที่กำหนดไว้อย่างชัดเจน สอดคล้องกับ การทำงานจริงของเจ้าหน้าที่ ในขณะที่ generative AI มีความยืดหยุ่นสูงกว่าจึงมีความแปรปรวนของผลลัพธ์ได้ มากกว่า ผลลัพธ์ดังกล่าวสอดคล้องกับการศึกษาของ Halvoník และ Kapusta (29) ที่พบว่าในงานเฉพาะทางที่มี ขอบเขตองค์ความรู้ชัดเจน ระบบแบบ rule-based อาจให้ ผลลัพธ์ที่แม่นยำและมีความสม่ำเสมอเหนือกว่า LLM ได้ เมื่อมีการออกแบบฐานความรู้ที่เหมาะสม

ในการประเมินประสิทธิภาพของระบบพบว่า rule-based system มีค่า overall accuracy เท่ากับ 0.933 ในการจัดประเภทเครื่องมือแพทย์ เครื่องสำอาง ผลิตภัณฑ์ กลุ่มที่ต้องติดต่อเจ้าหน้าที่ ยา และ อาหาร ซึ่งแม้จะอยู่ใน ระดับสูง แต่ยังไม่สามารถให้ผลลัพธ์ได้ถูกต้องสมบูรณ์ใน ทุกกรณี ทั้งนี้ ระบบที่อาศัยตรรกะแบบกฎของความสัมพันธ์ จะให้ผลลัพธ์ที่มีความแม่นยำสูง เมื่อโครงสร้างของกฎมีความถูกต้องครบถ้วน ครอบคลุม และผ่านการตรวจสอบ ความถูกต้องจากผู้เชี่ยวชาญอย่างเป็นระบบ อย่างไรก็ตาม ในการศึกษานี้อาจมีข้อจำกัดด้านความครบถ้วนของ โครงสร้างกฎ (rule incompleteness) ทำให้เกิดช่องว่างของ กฎ ระบบจึงไม่สามารถนำผู้ใช้งานไปสู่ตัวเลือกข้อสรุป สุดท้ายที่ชัดเจนได้ในบางกรณี และเกิดตัวเลือก “ไม่แน่ใจ” ขึ้น ข้อจำกัดดังกล่าวสอดคล้องกับการศึกษาของ Rashid และคณะ (30) ที่ชี้ให้เห็นว่าความครบถ้วนและความ สอดคล้องขององค์ความรู้เป็นปัจจัยสำคัญต่อคุณภาพของ ระบบฐานความรู้

ตัวชี้วัดด้านประสิทธิภาพของเครื่องมือชี้ว่า generative AI ยังคงมีประสิทธิภาพในการจัดประเภท ผลิตภัณฑ์ได้อย่างถูกต้องในหลายกรณี จึงสามารถใช้เป็น เครื่องมือช่วยตัดสินใจเบื้องต้นได้ ทั้งนี้ ความถูกต้องของ ผลลัพธ์อาจไม่ได้ขึ้นกับความสามารถของตัวโมเดลแต่เพียง อย่างเดียว แต่ยังขึ้นกับคุณภาพ ความครบถ้วน และ โครงสร้างของข้อมูลที่ใช้ในการ prompt ซึ่งสอดคล้องกับ แนวคิดด้าน prompt engineering ของ White และคณะ (31) ซึ่งเสนอแนวคิดว่าการออกแบบ prompt อย่างเป็น ระบบและให้ข้อมูลแวดล้อมที่เพียงพอมีส่วนสำคัญต่อ คุณภาพของคำตอบจาก LLM ทั้งยัง สอดคล้องกับการวิจัย ที่ผู้วิจัยกำหนดโครงสร้างของการ prompt ให้สอดคล้องกับ การออกแบบ rule-based system ดังนั้นผู้ใช้งานจึงควรมี ความเข้าใจเบื้องต้นเกี่ยวกับตรรกะการจัดประเภท

ผลิตภัณฑ์และระบุข้อมูลที่สำคัญของผลิตภัณฑ์ให้ครบถ้วน เพื่อเพิ่มความถูกต้องของผลลัพธ์ที่ได้จากระบบ ทั้งนี้ในการ ใช้งานยังคงควรได้รับคำแนะนำจากผู้เชี่ยวชาญเพื่อป้องกัน ความคลาดเคลื่อนของข้อมูลเสมอ

ในส่วนการประเมินคุณภาพของคำแนะนำจาก generative AI เมื่อถูกถามด้วย 5 คำถามถือเป็นการ ประเมินประสิทธิภาพในระยะเริ่มต้น จำนวนตัวอย่างที่ใช้ยัง ถือว่ามีค่อนข้างจำกัดสำหรับงานวิจัยด้าน natural language processing (NLP) และสารสนเทศทางสุขภาพ ซึ่งโดยทั่วไปให้ความสำคัญกับความครอบคลุมและความเป็น ตัวแทน ของโจทย์ที่ใช้ประเมินมากกว่าจำนวนตัวอย่าง เพียงอย่างเดียว เนื่องจากปัจจุบันยังไม่มีข้อกำหนด มาตรฐานที่ชัดเจนเกี่ยวกับจำนวนตัวอย่างที่เหมาะสม สำหรับการประเมินประสิทธิภาพของ LLM ทั้งนี้ การศึกษา ครั้งนี้ได้ออกแบบให้มีการประเมินซ้ำหลายรอบรวมกับการ ประเมินโดยผู้เชี่ยวชาญหลายคน เพื่อเพิ่มความน่าเชื่อถือ ของผลการประเมินภายใต้บริบทของการศึกษาเชิงพัฒนา เครื่องมือต้นแบบ อย่างไรก็ตาม ด้วยข้อจำกัดด้านจำนวน และความหลากหลายของชุดคำถาม ผลการประเมินอาจยัง ไม่สามารถสะท้อนศักยภาพของระบบ AI ต่อคำถามด้าน การขึ้นทะเบียนผลิตภัณฑ์ได้อย่างรอบด้าน โดยเฉพาะ คำถามที่มีโครงสร้างซับซ้อนหรือมีความกำกวมสูง ดังนั้น การศึกษาในอนาคตควรเพิ่มจำนวนและความหลากหลาย ของชุดคำถามให้ครอบคลุมบริบทการใช้งานมากยิ่งขึ้น เพื่อให้สามารถประเมินความสม่ำเสมอของคำตอบและ ประสิทธิภาพของระบบได้อย่างเหมาะสมยิ่งขึ้น

ในการเก็บข้อมูลผู้วิจัยทำการ prompt คำถามที่ พบบ่อยซ้ำเป็นจำนวน 3 ครั้งแบบ zero-shot prompting ภายใต้บริบทเดิมเพื่อจำลองรูปแบบการใช้งานจริงของ ผู้ใช้งาน ซึ่งมักมีการสอบถามประเด็นเดิมซ้ำภายในบริบท การสนทนาเดิม เพื่อให้ได้คำตอบที่ละเอียด ชัดเจน และ เหมาะสมต่อการนำไปใช้งานได้มากยิ่งขึ้น อย่างไรก็ตามวิธีการ ดังกล่าวอาจส่งผลต่อรูปแบบคำตอบของ generative AI เนื่องจากระบบยังคงรับรู้บริบทจากบทสนทนาด้านหน้า ทำให้คำตอบในรอบถัดไปอาจได้รับอิทธิพลจากข้อมูลใน context window ส่งผลให้ผลการประเมินอาจสะท้อน ประสิทธิภาพของระบบภายใต้การใช้งานในเชิงโต้ตอบ มากกว่าความสม่ำเสมอของระบบโดยตรง ซึ่งสอดคล้องกับ งานวิจัยของ Huang และคณะ (32) ที่พบว่า LLM มักพึ่งพา คำตอบเดิมที่เคยได้รับจนอาจก่อให้เกิด context pollution

ซึ่งส่งผลต่อคุณภาพคำตอบที่อาจพบข้อผิดพลาด การหลอนของข้อมูล (hallucination) หรือรูปแบบของภาษาที่ผิดเพี้ยนไป ซึ่งอาจสะสมต่อเนื่องข้ามรอบการสนทนาได้ ทั้งนี้ การศึกษาในอนาคตอาจพิจารณาเปรียบเทียบการ prompt แบบเริ่มบริบทใหม่ในแต่ละรอบ (independent session) เพื่อประเมินความสอดคล้องของคำตอบจากระบบได้อย่างชัดเจนยิ่งขึ้น

อย่างไรก็ตามเมื่อพิจารณาคะแนนในภาพรวมพบว่า คำแนะนำจาก generative AI ได้คะแนนการประเมินอยู่ในระดับปานกลางในทุกมิติของการประเมินและมีค่าความแปรปรวนของคะแนนค่อนข้างสูง สะท้อนว่าคำแนะนำของ generative AI ยังมีความไม่สม่ำเสมอ ซึ่งอาจขึ้นกับบริบทของคำถามและรูปแบบการ prompt อย่างไรก็ตามหากผู้ใช้งานใช้รูปแบบและโครงสร้างที่ชัดเจนรวมถึงให้บริบทได้อย่างครอบคลุม generative AI ก็จะให้ผลลัพธ์ที่มีประสิทธิภาพและแม่นยำยิ่งขึ้น (33)

เมื่อทำการวิเคราะห์ข้อมูลจากคำถามที่ใช้รายข้อพบว่า คำถามข้อ 1, 3 และ 4 ได้คะแนนด้านความถูกต้องและความสามารถในการนำไปใช้ได้จริงในระดับปานกลางค่อนข้างมาก ขณะที่คำถามที่ 2 และ 5 ได้คะแนนต่ำกว่า โดยลักษณะสำคัญของคำถามที่ได้คะแนนสูง คือ คำถามเชิงจำแนกประเภทผลิตภัณฑ์ที่มีโครงสร้างของคำตอบชัดเจนและมีลักษณะการตอบแบบใช้หรือไม่ ทำให้ generative AI สามารถให้คำตอบได้ค่อนข้างตรงประเด็นและอธิบายเหตุผลประกอบได้ดี ในทางตรงกันข้าม คำถามที่ได้คะแนนต่ำมักเป็นคำถามเชิงกระบวนการและขั้นตอนปฏิบัติ ซึ่งมีขอบเขตเนื้อหาที่กว้างและต้องการคำอธิบายเชิงบรรยายที่ซับซ้อน ส่งผลให้คำตอบที่ได้มีรายละเอียดมากขึ้น และมีโอกาสเกิดความคลาดเคลื่อนของข้อมูลได้มากกว่า ซึ่งแสดงให้เห็นว่า generative AI ยังมีข้อจำกัดในการให้คำแนะนำในประเด็นที่ซับซ้อนหรือมีโครงสร้างของคำตอบที่ไม่ชัดเจน

การวิเคราะห์ความแตกต่างของรอบการ prompt พบว่า การ prompt หลายครั้งมีความสัมพันธ์กับคุณภาพคำแนะนำที่ได้จาก generative AI ที่ดีขึ้น โดยเฉพาะในมิติด้านความถูกต้องและด้านความสามารถในการนำไปใช้ได้จริง อย่างไรก็ตามคะแนนที่ไม่ได้เพิ่มขึ้นอย่างสม่ำเสมอในทุกมิติอาจสะท้อนถึงความแปรผันของคุณภาพคำแนะนำที่ได้ตามบริบทของคำถามและรูปแบบการ prompt ดังนั้นควรนำคำแนะนำจาก AI ไปใช้เป็นเพียงแนวทางเบื้องต้นหรือเครื่องมือช่วยสนับสนุนการตัดสินใจเท่านั้น โดยก่อน

นำไปดำเนินการจริง ผู้ใช้งานควรได้รับคำแนะนำหรือการยืนยันจากผู้เชี่ยวชาญหรือเจ้าหน้าที่ที่เกี่ยวข้องซึ่งสอดคล้องกับผลการวิจัยของนันทพงศ์ บุญฤทธิ์และคณะ (25)

ผลการวิเคราะห์ด้วย weighted Cohen's Kappa พบว่า ความเห็นของผู้เชี่ยวชาญ 2 ท่านต่อคุณภาพของคำแนะนำจาก generative AI ในด้านความเหมาะสมของภาษามีค่าความสอดคล้องน้อยที่สุด ($K = 0.252$) ในขณะที่ด้านความถูกต้องของข้อมูล ($K = 0.571$) และด้านความสามารถในการนำไปใช้ได้จริง ($K = 0.498$) ความเห็นของผู้เชี่ยวชาญ 2 ท่านมีความสอดคล้องในระดับปานกลาง สะท้อนว่าผู้เชี่ยวชาญมีความเห็นสอดคล้องกันในมิติความถูกต้องและด้านความสามารถในการนำไปใช้ได้จริงมากกว่าด้านความเหมาะสมของภาษา เนื่องจากเป็นมิติที่สามารถตรวจสอบได้จากหลักเกณฑ์และข้อกำหนดที่ชัดเจน ในทางกลับกันประเด็นด้านความเหมาะสมของภาษามีความเป็นอัตวิสัยสูงกว่า โดยขึ้นกับการตีความด้านถ้อยคำ โครงสร้างประโยค และประสบการณ์ของผู้ประเมิน ซึ่งสอดคล้องกับงานวิจัยของ Bükür และคณะ (34) ที่พบว่า การประเมินด้าน readability มีความซับซ้อนและแปรผันตามมุมมองของผู้ประเมินมากกว่ามิติด้านความถูกต้องของข้อมูลและมีดัชนีชี้วัดค่อนข้างมาก เนื่องจากต้องผลานความซับซ้อนของโครงสร้างประโยครวมถึงความเข้าใจของผู้ประเมิน ซึ่งนำไปสู่ความแตกต่างของคะแนนมากกว่ามิติด้านความถูกต้องของข้อมูล นอกจากนี้ค่าความสอดคล้องที่อยู่ในระดับไม่สูงมากอาจเกิดจากเกณฑ์การประเมิน (rubric score) ที่ยังไม่ละเอียดเพียงพอ และขาดกระบวนการปรับเทียบผู้ประเมิน (rater calibration) ก่อนการประเมินจริง ดังนั้นในการศึกษาครั้งต่อไป ควรมีการพัฒนา rubric score ให้มีความชัดเจนและเป็นมาตรฐาน พร้อมทั้งจัดทำ scoring guideline และดำเนินการปรับเทียบผู้ประเมินก่อนการประเมิน เพื่อเพิ่มความสอดคล้องและความน่าเชื่อถือของผลการประเมิน

ในการประเมินความพึงพอใจของผู้ใช้งานเว็บแอปพลิเคชันโดยอาสาสมัครจำนวน 30 คนที่มาใช้บริการที่ OSSC ออย. พบว่า ตัวอย่างส่วนใหญ่เป็นเพศหญิง และช่วงอายุที่มากที่สุดคือ 30-39 ปี ประกอบอาชีพพนักงานบริษัทหรือมีธุรกิจส่วนตัว และกว่าครึ่งมีประสบการณ์ทำงานด้านทะเบียนผลิตภัณฑ์น้อยกว่า 1 ปี สะท้อนว่ากลุ่มตัวอย่างมีพื้นฐานความรู้ด้านข้อกำหนดที่เกี่ยวกับงานขึ้นทะเบียนใน

ระดับเริ่มต้นถึงปานกลาง ซึ่งสอดคล้องกับวัตถุประสงค์ในการพัฒนาเครื่องมือนี้เพื่อช่วยให้ผู้ประกอบการและนักวิจัยหน้าใหม่ สามารถเรียนรู้การจัดประเภทผลิตภัณฑ์สุขภาพได้ด้วยตนเอง ผลการประเมินพบว่าผู้ใช้งานมีความพึงพอใจต่อเว็บแอปพลิเคชันที่พัฒนาขึ้นในระดับสูงทุกด้าน โดยได้คะแนนเฉลี่ย 4.38 - 4.47 จาก 5 คะแนน

ข้อจำกัดของงานวิจัยนี้ได้แก่ การใช้โมเดล AI ที่เป็นโมเดลเชิงภาษาทั่วไป (general-purpose AI) ซึ่งยังไม่ได้ถูกปรับแต่งให้เหมาะสมเฉพาะด้านผลิตภัณฑ์สุขภาพ จึงทำให้พบข้อจำกัดในด้านความถูกต้อง ความครบถ้วน และความสม่ำเสมอของคำแนะนำในบางกรณี โดยเฉพาะอย่างยิ่งคำถามที่มีโครงสร้างหรือบริบทไม่ชัดเจนหรือคำถามในเชิงกระบวนการขั้นตอน นอกจากนี้ การประเมินประสิทธิภาพของเครื่องมือในการจัดประเภทผลิตภัณฑ์มีข้อจำกัดด้านจำนวนตัวอย่าง เนื่องจากใช้ข้อมูลทดสอบเพียง 30 กรณีและไม่ได้ครอบคลุมผลิตภัณฑ์ครบทั้ง 9 ประเภท ส่งผลให้ผลการทดสอบอาจยังไม่เพียงพอในการสะท้อนความสามารถของระบบในแต่ละหมวดหมู่อย่างครบถ้วน

ข้อเสนอแนะ

การวิจัยในอนาคตควรใช้ generative AI ที่มีความเฉพาะทางมากขึ้นหรือพัฒนา generative AI ขึ้นใช้ภายในหน่วยงานเอง โดยประยุกต์ใช้เทคนิค retrieval-augmented generation (RAG) โดยพัฒนาให้สามารถเชื่อมโยงกับฐานข้อมูลขององค์กร เพื่อเพิ่มความแม่นยำและความน่าเชื่อถือของคำแนะนำ ซึ่งจะช่วยเพิ่มประสิทธิภาพและความปลอดภัย นอกจากนี้ควรตรวจสอบความพร้อมของระบบก่อนการนำไปใช้งานจริง ปรับปรุงข้อมูลและตรรกะการประมวลผลให้สอดคล้องกับกฎระเบียบที่มีการปรับปรุงอยู่เสมอ พร้อมทั้งพัฒนาแนวทางการกำกับดูแลการใช้ generative AI ให้สอดคล้องกับข้อกำหนดของภาครัฐ นอกจากนี้ ควรปรับปรุง rule-based system และขยายขอบเขตการจำแนกอย่างต่อเนื่อง โดยพัฒนาร่วมกับผู้เชี่ยวชาญ เพื่อลดปัญหา rule gap และเพิ่มความละเอียดของการจัดประเภทในระดับย่อยให้ครอบคลุมมากยิ่งขึ้น

ระบบที่พัฒนาควรยกระดับความปลอดภัยของระบบและลดความเสี่ยงจากการใช้ third-party API โดยอาศัยเทคนิค data minimization และ anonymization ในการรับ-ส่งข้อมูล และในการพัฒนาระยะยาวควรพัฒนาระบบ AI ของหน่วยงานเองแบบ private หรือ on-premise

ขององค์กรเองเพื่อเพิ่มความปลอดภัยและประสิทธิภาพ นอกจากนี้ ควรเพิ่มการแจ้งเตือนผู้ใช้งานก่อนการกรอกข้อมูลในช่องที่ส่งข้อมูลให้ generative AI โดยไม่ควรส่งข้อมูลที่เป็นความลับทางการค้า หรือข้อมูลที่อ่อนไหวเพื่อป้องกันความลับรั่วไหล

ในการใช้ generative AI เพื่อการจัดประเภทผลิตภัณฑ์ควรมีการแนะนำการใส่ keyword ในการจัดประเภทให้ครบถ้วน เพื่อให้ผลลัพธ์ในการจัดประเภทถูกต้องแม่นยำยิ่งขึ้น นอกจากนี้ การศึกษาในอนาคตควรเพิ่มจำนวนและความหลากหลายของทั้งตัวอย่างข้อมูลที่ใช้ในการวิเคราะห์ ทั้งตัวอย่างผลิตภัณฑ์ในการวิเคราะห์และประเด็นปัญหาที่ใช้ในการวิเคราะห์ โดยอาจใช้เทคนิคแบ่งชั้นเพื่อให้ได้ข้อมูลที่ครอบคลุมกับสถานการณ์จริงซึ่งจะช่วยให้เพิ่มความน่าเชื่อถือของผลการประเมิน

ในการพัฒนาการประเมินและการใช้งาน generative AI ในเชิงลึก ควรศึกษาปัจจัยรูปแบบของ prompt และบริบทที่มีผลต่อคุณภาพคำตอบ รวมถึงประเมินผลกระทบของระบบที่เป็นรูปธรรม ซึ่งจะสามารถสะท้อนประสิทธิภาพการทำงานของระบบได้อย่างชัดเจนมากขึ้น เช่น การประเมินความถูกต้องในการจัดประเภทของผู้ใช้งานเอง ระยะเวลาที่ลดลงในการดำเนินงาน หรือประสิทธิภาพการใช้งานระบบเมื่อเทียบกับวิธีการดั้งเดิม เพื่อให้สามารถประเมินประสิทธิภาพของระบบในบริบทการใช้งานจริงได้อย่างครอบคลุมยิ่งขึ้น นอกจากนี้ ในอนาคตอาจทำการศึกษา generative AI จากหลากหลายแบบจำลองภาษา เพื่อเปรียบเทียบความสามารถด้านการประมวลผลภาษาไทย การให้เหตุผลเชิงกฎหมาย ต้นทุนการประมวลผล และความเหมาะสมต่อการประยุกต์ใช้ในงานสนับสนุนการตัดสินใจด้านการขึ้นทะเบียนผลิตภัณฑ์สุขภาพ

สรุป

งานวิจัยครั้งนี้แสดงให้เห็นว่า เว็บแอปพลิเคชันเพื่อการจัดประเภทผลิตภัณฑ์สุขภาพที่พัฒนาขึ้นสามารถจัดประเภทผลิตภัณฑ์ได้โดยมีความสอดคล้องกับผู้เชี่ยวชาญในระดับดีมาก ทั้งในภาพรวมเมื่อเทียบกับผู้ประเมินทั้งหมดและในระดับรายคู่ โดย rule-based system ที่พัฒนาขึ้นจากกฎหมายมีการจัดประเภทผลิตภัณฑ์ได้ใกล้เคียงกับผู้เชี่ยวชาญที่ดีกว่าการจัดประเภทด้วย generative AI ส่วนฟังก์ชันการให้คำแนะนำโดย generative

AI ซึ่งฝังอยู่ในระบบ สามารถให้คำแนะนำในประเด็นด้านการขึ้นทะเบียนได้ในระดับปานกลางค่อนข้างดี และมีประสิทธิภาพสูงขึ้นเมื่อใช้ prompt คำถามที่มีลักษณะโครงสร้างและบริบทที่ชัดเจน อย่างไรก็ตามยังคงมีข้อจำกัดด้านความสม่ำเสมอของคุณภาพในการให้คำแนะนำ โดยเฉพาะอย่างยิ่งกรณี prompt คำถามที่มีโครงสร้างหรือบริบทไม่ชัดเจน หรือคำแนะนำที่ต้องอธิบายเป็นกระบวนการขั้นตอน ดังนั้นในการใช้งานจึงควรใช้เพื่อเป็นเพียงส่วนช่วยสนับสนุนการตัดสินใจหรือเป็นแนวทางเบื้องต้นเท่านั้น ยังคงต้องได้รับคำแนะนำจากผู้เชี่ยวชาญควบคู่ด้วย ผู้ใช้งานมีความพึงพอใจสูงต่อเว็บแอปพลิเคชันที่พัฒนาขึ้น ระบบดังกล่าวมีศักยภาพที่จะพัฒนาไปเป็นเครื่องมือจัดประเภทผลิตภัณฑ์สุขภาพสำหรับผู้ประกอบการหรือนักวิจัยรายใหม่ได้ ซึ่งจะสามารถช่วยลดภาระงานของเจ้าหน้าที่ และเตรียมความพร้อมให้ผู้ประกอบการดำเนินการได้สอดคล้องกับข้อกำหนดของอย.

กิตติกรรมประกาศ

การวิจัยครั้งนี้สำเร็จได้ด้วยความอนุเคราะห์จาก ผู้ให้ข้อมูลทุกท่านตลอดจนผู้เกี่ยวข้องทุกรายที่ผู้วิจัยขอข้อมูลเพิ่มเติม งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยประเภท นักศึกษาระดับบัณฑิตศึกษาจากกองทุนวิจัยและสร้างสรรค์ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร

การใช้ปัญญาประดิษฐ์ในบทความ

ในการพัฒนาองค์ประกอบของเว็บแอปพลิเคชัน ผู้วิจัยเลือกใช้ ChatGPT (OpenAI) ผ่าน API โดยใช้ OpenAI model identifier : gpt-5 เพื่อเป็นองค์ประกอบของระบบสำหรับการให้คำแนะนำเบื้องต้นเพื่อการขึ้นทะเบียนผลิตภัณฑ์สุขภาพภายในเว็บแอปพลิเคชัน และใช้เป็นส่วนหนึ่งของกระบวนการทดลองตามระเบียบวิธีวิจัย โดยเข้าใช้งานผ่าน API key เมื่อวันที่ 27 มิถุนายน 2568 ในการดำเนินการจัดทำต้นฉบับงานวิจัย ผู้วิจัยเลือกใช้ ChatGPT (OpenAI) ตั้งแต่ในช่วงเดือน กันยายน พ.ศ.2568 และมีการใช้งานอย่างต่อเนื่องในช่วงที่จัดทำต้นฉบับ โดยใช้เฉพาะเพื่อช่วยในการปรับปรุงความชัดเจนของภาษา ตรวจสอบไวยากรณ์ ช่วยเรียบเรียงข้อความ และจัดทำภาพประกอบ ผู้เขียนขอยืนยันว่าได้ตรวจสอบความถูกต้องของเนื้อหาทั้งหมดที่ AI สร้างขึ้น และผู้เขียนขอรับผิดชอบต่อความ

ถูกต้อง ความคิดริเริ่ม และความน่าเชื่อถือของเนื้อหาในบทความฉบับสมบูรณ์

เอกสารอ้างอิง

1. Public Health Ministerial Declaration in 2020 on the organization of the office of the Thai Food and Drug Administration. Royal Gazette. No. 137 (Part 8A) (Jan 29, 2020).
2. Food Act B.E. 2522. Royal Gazette. No. 79 (May 8, 1979).
3. Narcotics Code. Royal Gazette. No. 138, Part 73A (Nov 8, 2021).
4. Hazardous Substance Act B.E. 2535. Royal Gazette. No. 57A (Mar 29, 1992).
5. Drug Act B.E. 2510. Royal Gazette. No. 101 (Oct 20, 1967).
6. Herbal Product Act B.E. 2562. Royal Gazette. No. 56A (Apr 30, 2019).
7. Cosmetic Act B.E. 2558. Royal Gazette. No. 132, Part 86A (Sep 8, 2015).
8. Medical Device Act B.E. 2562. Royal Gazette. No. 2, Part 56A (Apr 30, 2019).
9. Kasikornbank. The growth of the health business as a new income opportunity in the COVID-19 era [online]. 2021 [cited Apr 4, 2024]. Available from: www.kasikornbank.com/th/personal/the-wisdom/articles/Pages/BIZ-Empowerment_Onward57_2021.aspx.
10. Notification of the National Herbal Policy Committee on measures to promote herbal production for use as raw materials in herbal product manufacturing under safe agriculture guidelines. Royal Gazette. No.139, Part 287D special 9Dec 8, 2022).
11. Department of Thai Traditional and Alternative Medicine. National herbal development action plan, phase 2 (2023–2027). Nonthaburi: Ministry of Public Health; 2023.
12. Division of Innovative Health Products and Services. Infographic: Guidelines for selecting the appropriate application in the e-consult system [online]. 2023

- [cited Apr 20, 2024]. Available from: innovative.fda.moph.go.th/online-consultation/infographic-02.
13. Moore J, Clark A, McKinley P. Evolutionary robotics on the web with WebGL and Javascript [online]. 2014 [cited Apr 20, 2024]. Available from: doi.org/10.48550/arXiv.1406.3337
14. Ginige A, Murugesan S. Web engineering: An introduction. *IEEE Multimedia*. 2001; 8: 14–8.
15. Kovari A. AI for decision support: balancing accuracy, transparency, and trust across sectors. *Information*. 2024; 15: 725.
16. Ahmed N, Saha AK, Al Noman MA, Jim JR, Mridha MF, Kabir MA. Deep learning-based natural language processing in human–agent interaction: Applications, advancements and challenges. *Nat Lang Proc J* 2024; 9: 100112.
17. Cordeschi R. AI turns fifty: revisiting its origins. *J Exp Theor Artif Intell* 2007; 19: 257-65.
18. Levin C, Zaboli A, Turcato G, Saban M. Nursing judgment in the age of generative artificial intelligence: A cross-national study on clinical decision-making performance among emergency nurses. *Int J Nurs Stud* 2025; 172: 105216.
19. U.S. Food and Drug Administration. Considerations for the use of artificial intelligence to support regulatory decision-making for drug and biological products draft guidance for industry and other interested parties. U.S. Food and Drug Administration; 2025. Report No.: FDA-2024-D-4689.
20. European Medicines Agency. Reflection paper on the use of Artificial Intelligence (AI) in the medicinal product lifecycle. Amsterdam, The Netherlands: European Medicines Agency; 2024 Sep 9. Report No.: EMA/CHMP/CVMP/83833/2023.
21. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva, Switzerland: World Health Organization; 2021.
22. Gurung G, Shah R, Jaiswal DP. Software development life cycle models—A comparative study. *Int J Sci Res Comput Sci Eng Inf Technol* 2020; 6: 30–7.
23. Google LLC. Apps Script [online]. 2024 [cited Jan 22, 2024]. Available from: developers.google.com/apps-script?hl=th.
24. Yang L. Responsive website design and development based on Bootstrap -- taking digital media technology professional learning resource website as an example. In: *Proceedings of the 2024 International Conference on Computer and Multimedia Technology (ICCMT '24)*. New York: Association for Computing Machinery; 2024. p. 190-194. Available from: doi.org/10.1145/3675249.3675284
25. Boonrit N, Hopkins AM, Ruanglertboon W. Assessment of the appropriateness of responses in Thai from ChatGPT on the questions for recommendations of drug uses in common illnesses. *Thai Journal of Pharmacy Practice*. 2023; 15: 1134–48.
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33: 159–74.
27. Likert R. A technique for the measurement of attitudes. New York: Columbia University Press; 1932.
28. Nunnally JC, Bernstein IH. *Psychometric theory*. New York: McGraw-Hill; 1994.
29. Halvoník D, Kapusta J. Large language models and rule-based approaches in domain-specific communication. *IEEE Access*. 2024;12:107046-58
30. Rashid MRA, Rizzo G, Torchiano M, Mihindukulasooriya N, Corcho O, García-Castro R. Completeness and consistency analysis for evolving knowledge bases. *Web Semant* 2019; 54: 48-71.
31. White J, Fu Q, Hays S, Sandborn M, Olea C, Gilbert H, et al. A prompt pattern catalog to enhance prompt engineering with ChatGPT. *arXiv* 2023: 2302.11382. doi.org/10.48550/arXiv.2302.11382

32. Huang JY, Choshen L, Astudillo R, Broderick T, Andreas J. Do LLMs benefit from their own words? arXiv[Preprint]. 2026: 2602.24287.
33. Anam RK. Prompt engineering and the effectiveness of large language models in enhancing human productivity. arXiv [Preprint]. 2025: 2507.18638.
34. Bükür M, Mercan G. Readability, accuracy and appropriateness and quality of AI chatbot responses as a patient information source on root canal retreatment: A comparative assessment. Int J Med Inform 2025; 201: 105948.

ภาคผนวก

เว็บแอปพลิเคชันการจัดประเภทผลิตภัณฑ์สุขภาพ



ยินดีต้อนรับทุกท่านเข้าสู่เว็บแอปพลิเคชันเพื่อการจัดประเภทผลิตภัณฑ์สุขภาพ ซึ่งเป็นหนึ่งในเครื่องมือสำหรับการวิจัยของ ภก.กฤตพล จันทราบัณฑิต นักศึกษาหลักสูตรเภสัชศาสตรมหาบัณฑิต สาขาวิชา สารสนเทศศาสตร์ทางสุขภาพ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร โดยเว็บแอปพลิเคชันนี้ได้รับการสนับสนุนทางวิชาการจากทั้งสำนักงานคณะกรรมการอาหารและยา และมหาวิทยาลัยศิลปากร



ระบบจัดประเภทผลิตภัณฑ์สุขภาพ
ทดลองจัดประเภทผลิตภัณฑ์สุขภาพด้วยตัวท่านเอง เพื่อการประเมินความเป็นไปได้ทางธุรกิจและแนวทางการเตรียมเอกสารขึ้นทะเบียนผลิตภัณฑ์สุขภาพของท่าน

[เข้าสู่ระบบ](#)



ทำไมต้องจัดประเภทผลิตภัณฑ์สุขภาพ
ค้นพบความสำคัญของการจัดประเภทผลิตภัณฑ์สุขภาพ เคล็ดลับของความเข้าใจเริ่มต้นที่ก้าวแรก

[ที่มาและความสำคัญ](#)



เว็บแอปพลิเคชันนี้ใช้งานอย่างไร
ไม่เข้าใจ สงสัย หลงทาง เราเปิดใจออก ศึกษาคู่มือการใช้งานเว็บแอปพลิเคชันได้ที่นี่

[คู่มือการใช้งานเว็บแอปฯ](#)

หน่วยงานที่เกี่ยวข้องกับการขึ้นทะเบียนผลิตภัณฑ์

กองยา	กลุ่มควบคุมเครื่องสำอาง	กลุ่มควบคุมวัตถุอันตราย	กองเครื่องมือแพทย์
กองอาหาร	กองควบคุมวัตถุเสพติด	กองผลิตภัณฑ์สมุนไพร	กองผลิตภัณฑ์สุขภาพนวัตกรรมและการบริการ

รูป ก. ภาพหน้าแรกของเว็บแอปพลิเคชัน



สวัสดี AdminFN ลงชื่อออก

ผลการวินิจฉัยเบื้องต้น: **ผลิตภัณฑ์สมุนไพร** | path : d2286da8-b2e0-4387-bd76-578ac6975359

คำถาม: รูปแบบและลักษณะ (format) ของผลิตภัณฑ์สุขภาพของท่านเป็นอย่างไร ?

คำตอบ:

2. รูปแบบ (format) ของผลิตภัณฑ์ดังกล่าวเป็นผลิตภัณฑ์สำเร็จรูป (finished product) หมายถึงเป็นผลิตภัณฑ์ที่ผู้บริโภคสามารถใช้ประโยชน์ได้โดยตรง

คำถาม: ผลิตภัณฑ์ดังกล่าวมีจุดมุ่งหมายใช้กับสิ่งใด ?

คำตอบ: 1. มุ่งหมายใช้กับคน

คำถาม: หลักการทำงานหรือกลไกการออกฤทธิ์ของผลิตภัณฑ์นี้ (mode of actions)

คำตอบ:

1. ผลิตภัณฑ์ดังกล่าวอาศัยกลไกการออกฤทธิ์ทางเภสัชวิทยา (pharmacological) วิทยาภูมิคุ้มกัน (immunological) หรือปฏิกิริยาเผาผลาญให้เกิดพลังงาน (metabolic activity)

คำถาม: ผลิตภัณฑ์ดังกล่าวมีส่วนประกอบหลัก (active ingredient) แบบใด ?

คำตอบ:

2. ผลิตภัณฑ์ดังกล่าวมีส่วนประกอบหลักเพียงอย่างเดียว ตัวอย่างเช่น เป็นสารเคมี หรือเป็นสารจากสมุนไพร อย่างใดอย่างหนึ่ง

คำถาม: ส่วนประกอบหลักดังกล่าวเป็นสารที่มีต้นกำเนิดจากสิ่งใด (origin) ?

คำตอบ: 1. ผลิตภัณฑ์ดังกล่าวมีส่วนประกอบหลักจากสมุนไพร

รูป ง. ภาพหน้าแสดงผลการจัดประเภทผลิตภัณฑ์สุขภาพของเว็บแอปพลิเคชัน

ผลการวินิจฉัยจัดประเภทผลิตภัณฑ์เบื้องต้น(id: e14cfd37-a3ca-4460-93fc-9f24fadca7ba):
GMT+0700 (เวลาอินโดจีน) 🔍 🗖



ถึง จัน

เรียนคุณ

ตามที่ท่านได้ขอร้องให้ผลการวินิจฉัยจัดประเภทผลิตภัณฑ์เบื้องต้น

ขอแจ้งผลการวินิจฉัย รายละเอียดดังนี้:

Response ID: e14cfd37-a3ca-4460-93fc-9f24fadca7ba

วันที่ทำการบันทึก: Tue Nov 11 2025 13:57:58 GMT+0700 (เวลาอินโดจีน), API version: 5.0, Last Update On: 2025-09-23T08:41:36.354Z

ชื่อผลิตภัณฑ์: สเปรย์บรรเทาอาการปวดเมื่อยกล้ามเนื้อ

ข้อบ่งใช้/สรรพคุณ: แก้อาการปวดเมื่อยกล้ามเนื้อ

ผลการวินิจฉัยเบื้องต้น: ผลิตภัณฑ์สมุนไพร

คำถาม:

รูปแบบและลักษณะ (format) ของผลิตภัณฑ์สุขภาพของท่านเป็นอย่างไร ?

คำตอบ:

2. รูปแบบ (format) ของผลิตภัณฑ์ดังกล่าวเป็นผลิตภัณฑ์สำเร็จรูป (finished product) หมายถึงเป็นผลิตภัณฑ์ที่ผู้บริโภคสามารถใช้ประโยชน์ได้โดยตรง

คำถาม:

ผลิตภัณฑ์ดังกล่าวมีจุดมุ่งหมายใช้กับสิ่งใด ?

คำตอบ:

1. มุ่งหมายใช้กับคน

รูป จ. ภาพ แสดงตัวอย่างของอีเมลแจ้งผลการวินิจฉัยการจัดประเภทผลิตภัณฑ์เบื้องต้น

ผลิตภัณฑ์สุขภาพของท่านชื่อ:

สเปรย์บรรเทาอาการปวดเมื่อยกล้ามเนื้อ

ตัวอย่าง เช่น ขนมหวานแสนอร่อย

ข้อบ่งใช้/สรรพคุณ:

แก้การปวดเมื่อยกล้ามเนื้อ

ตัวอย่าง เช่น ใช้ลดระดับน้ำตาลในเลือด

เริ่มการจัดประเภท !

🧐 สอบถามข้อมูลเกี่ยวกับการขึ้นทะเบียนกับปณณัติรัฐ

ปวดเมื่อยกล้ามเนื้อ สามารถใช้เป็นคำเคลมในยาสมุนไพรได้ไหม

Model's Response:

ได้ แต่มีเงื่อนไขดังนี้

- ต้องขึ้นทะเบียนเป็น "ผลิตภัณฑ์สมุนไพรประเภทยา" กับ อย. และคำสรรพคุณต้องได้รับอนุมัติในทะเบียนตำรับ
- ใช้ถ้อยคำเชิงบรรเทา เช่น "บรรเทาอาการปวดเมื่อยกล้ามเนื้อ" ไม่ใช่คำว่า "รักษา" และต้องไม่พาดพิงโรค/ภาวะรุนแรง
- ควรสอดคล้องกับรูปแบบยา โดยเฉพาะตำรับใช้ภายนอก (เช่น ยาหม่อง น้ำมัน/ยาน้ำระเหย แผ่นแปะ) ที่มีสูตรตำรับหรือบัญชีที่ อย. รับรอง มักอนุญาตคำเคลมนี้ได้
- หากเป็นยารับประทาน ต้องมีหลักฐานสนับสนุนสรรพคุณตามระดับความเสี่ยง (อ้างอิงตำรายา/บัญชียาหลักสมุนไพร/Thai Herbal Pharmacopoeia หรือข้อมูลวิทยาศาสตร์/คลินิกสำหรับสูตรใหม่)
- ห้ามใช้คำเคลมนี้กับผลิตภัณฑ์ที่ขึ้นทะเบียนเป็นอาหารเสริมหรือเครื่องสำอาง เพราะเป็นสรรพคุณยา

คำสรรพคุณที่อนุมัติได้ต้องตรงกับหลักฐานและสอดคล้องกับสูตร-ขนาดใช้-วิธีใช้ที่ยื่นจดทะเบียนเสมอ หากบอกส่วนประกอบ/รูปแบบยาได้ จะประเมินความเป็นไปได้ของคำเคลมได้ชัดเจนขึ้น

✓ ▶

🔄 ส่งคำถามใหม่อีกครั้ง

รูป จ. ภาพแสดงตัวอย่างการใช้งานฟังก์ชันแนะนำการขึ้นทะเบียนเบื้องต้นด้วย generative AI