

สรุป

งานวิจัยนี้มุ่งเน้นการสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรโดยเฉพาะอย่างยิ่งสมุนไพรไทยจากเอกสารภาษาไทย โดยมีการสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรนี้ออกมาในรูปแบบของEDU ซึ่งเป็นประโยคแบบง่าย ๆ ธรรมดา แต่สกัดออกมาได้ครั้งละหลายๆ EDUs ที่ต่อเนื่องกัน คือได้หลายๆสรรพคุณติดต่อกันต่อพืชสมุนไพรหนึ่งชนิด ในขณะที่งานวิจัยก่อนหน้านี้ จะเป็นการสกัดออกมาได้ครั้งละหนึ่งหรือ หนึ่งถึงสามคุณสมบัติ/สรรพคุณ เท่านั้น ทั้งนี้เพราะงานวิจัยก่อนหน้านี้จะเป็นการสกัดที่อยู่บนพื้นฐานของนามวลีเท่านั้น ในขณะที่งานวิจัยนี้มุ่งเน้นการสกัดความรู้บนพื้นฐานของเหตุการณ์ที่สามารถแทนได้ดีด้วยกริยา คือ ทำให้สามารถเข้าใจถึงเหตุการณ์ได้อย่างต่อเนื่อง เช่น “กระเพราใช้เป็นยาขับลม แก้คลื่นไส้ [แก้]อาเจียน บรรเทาอาการปวดท้อง แก้ท้องเสีย” ฉะนั้นความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรที่สกัดได้นี้ช่วยสนับสนุนการตอบคำถามชนิด “คำถามอะไร/what-Question” ประเภทลิสต์ ได้อย่างชัดเจน

งานวิจัยนี้สามารถสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรในรูปแบบหลายๆ EDUs ได้อย่างมีประสิทธิภาพโดยเฉพาะอย่างยิ่งการสกัดขอบเขตของกลุ่มสรรพคุณทางยาของพืชสมุนไพรด้วยการใช้เทคนิคที่แตกต่างกันสองวิธี คือ วิธีการประมวลผลภาษาธรรมชาติด้วยการประยุกต์ใช้ความรู้ทางภาษาศาสตร์ระดับวลีของประโยคร่วมกับเทคนิคการเรียนรู้ของเครื่องด้วยเทคนิคของ NB และวิธีการประมวลผลภาษาธรรมชาติที่ประยุกต์ใช้ความรู้ทางภาษาศาสตร์ในระดับบทความโดยใช้ CT ซึ่งก็คือศูนย์กลางความสนใจ (Center of Attention) ของแต่ละEDUภายใต้ขอบเขตของสรรพคุณทางยาของพืชสมุนไพร ที่ส่วนใหญ่จะมีผู้กระทำ (Agent) ที่แตกต่างกันน้อย ดังนั้นค่าเฉลี่ยความถูกต้องของการหาขอบเขตของสรรพคุณทางยาของพืชสมุนไพรสำหรับ NB เป็น 91.5% และ CT เป็น 86% กล่าวคือ NB ให้ผลลัพธ์ที่มีประสิทธิภาพกว่า CT ถึงแม้ว่าพฤติกรรมของข้อมูลจะแตกต่างกันก็ตาม ในขณะที่การระบุ EDU ที่เป็นสรรพคุณทางยาของพืชสมุนไพรด้วยการใช้ Lexico Syntactic Pattern (NP1 V_{mp} NP2) ไปพร้อมกับความหมายเชิงความคิด (Concept) ของกริยาและนามวลี ทำให้ได้ค่า Precision และ Recall เฉลี่ยคือ 89% และ 74% ตามลำดับ ทั้งนี้ขึ้นอยู่กับพฤติกรรมของคลังข้อมูล

ระบบการสกัดความรู้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพรจากเอกสารทางวิชาการภาษาไทยนี้สามารถนำไปใช้สนับสนุน ระบบการตอบคำถามชนิด “ถามอะไร/what-Question” ประเภทลิสต์ และประเภทเอนิตี้เกี่ยวกับสรรพคุณทางยาของพืชสมุนไพร ด้วยวิธีการใช้แพทเทิร์นที่ได้จากการสังเกตจากการเรียนรู้ ซึ่งแพทเทิร์นเหล่านี้สามารถใช้สำหรับวิเคราะห์คำถามที่ป้อนเข้ามาโดยไม่มีการใช้ซอฟต์แวร์ Parser ฉะนั้นคำถามที่เข้ามาในระบบจะต้องเป็นคำถามที่ไม่ยาว คือ ไม่ควรรยาวกว่า 10 คำ ผลลัพธ์ของระบบการตอบคำถามของงานวิจัยนี้สามารถตอบคำถามได้ถูกต้อง 72% ทั้งนี้เนื่องมาจากบางคำถามมีความซับซ้อน อย่างไรก็ตามหากระบบมีการพัฒนาต่อไปให้สามารถตอบคำถามได้ถูกต้องเพิ่มขึ้น ก็จะทำให้ระบบการตอบคำถามนี้มีประโยชน์ต่อสุขภาพของสังคมอย่างมาก โดยเฉพาะชาวบ้านที่อยู่ห่างไกลจากตัวเมือง เมื่อมีความจำเป็นต้องใช้ยาและไม่สามารถเดินทางไปหาซื้อได้ ฉะนั้นยาสมุนไพรก็เป็นอีกช่องทางที่สามารถหาใช้ได้ทันทีหากมีการปลูกพืชสมุนไพรที่ได้กล่าวมาข้างต้นไว้ในบ้านบ้าง