

น้องเอ็ม: ระบบคัดกรองสุขภาพจิตเบื้องต้นด้วยปัญญาประดิษฐ์และฐานข้อมูลสำหรับคนไทย
NongM: An Artificial Intelligence Based Preliminary Mental Health Screening System
with a Mental Health Dataset for Thais

วรารัตน์ ศิริสวัสดิ์ วิชาญชุต รักษาคอด

ที่ปรึกษา: เอกชัย วัฒนไชย

โรงเรียนวิทยาศาสตร์จุฬารามราชวิทยาลัย บุรีรัมย์ ตำบลสตึก อำเภอสตึก จังหวัดบุรีรัมย์ 31150

*อีเมลล์: worawat.sirisawat0@gmail.com

บทคัดย่อ

ปัจจุบันประเทศไทยเผชิญกับวิกฤตสุขภาพจิตอย่างต่อเนื่อง โดยมีผู้ป่วยโรคซึมเศร้าประมาณ 2.9 ล้านคน ขณะที่อัตราส่วนจิตแพทย์ต่อประชากรมีเพียง 1 ต่อ 50,000 คน ส่งผลให้เกิดช่องว่างในการเข้าถึงบริการ โครงการนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบน้องเอ็ม ซึ่งเป็นระบบปัญญาประดิษฐ์สำหรับคัดกรองสุขภาพจิตเบื้องต้นในรูปแบบเว็บแอปพลิเคชันที่ออกแบบเฉพาะสำหรับภาษาไทย เพื่อลดอุปสรรคการเข้าถึงบริการและสร้างฐานข้อมูลการสนทนาเพื่อการวิจัยในอนาคต การดำเนินงานใช้สถาปัตยกรรมแบบ Microservices ผสานเทคโนโลยี Retrieval Augmented Generation ร่วมกับการทดสอบ Proof of Concept โดยใช้โมเดลภาษาขนาดใหญ่ภาษาไทย typhoon-v2.5-30b-a3b-instruct เพื่อลดปัญหาการคิดไปเองของโมเดล โดยใช้ PyThaiNLP ในการเตรียมข้อความ และใช้ Semantic Similarity ร่วมกับ CrossEncoder Reranking ในการดึงเอกสารทางการแพทย์ ระบบรองรับการคัดกรองตามช่วงอายุของผู้ใช้ผ่านแบบประเมินมาตรฐานของกรมสุขภาพจิต ได้แก่ แบบประเมิน CDI สำหรับเด็ก, PHQ-A สำหรับวัยรุ่นและ PHQ-9 สำหรับผู้ใหญ่ โดยคำนวณคะแนนผ่านฝั่ง backend เพื่อความแม่นยำทางคลินิก นอกจากนี้ยังได้รับการออกแบบตามหลัก Privacy by Design ที่สอดคล้องกับ พ.ร.บ.คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 (PDPA) โดยไม่มีการเก็บข้อมูลที่ระบุตัวตน ผลการประเมินประสิทธิภาพ RAG Pipeline ด้วยชุดทดสอบ 10 queries ใน 8 หมวดหมู่ พบว่าระบบมีค่า Hit Rate อยู่ที่ 0.80 และ MRR อยู่ที่ 0.75 โดยมีค่าความหน่วงเฉลี่ย 908.2 มิลลิวินาทีหมวดหมู่หลัก 5 หมวด depression_basics, assessment_tool, treatment, differential, population มีประสิทธิภาพสูงสุด NDCG@5 = 1.00 และ Hit Rate = 1.00 กลไก CrossEncoder Reranking ช่วยเลื่อนตำแหน่งเอกสารที่เกี่ยวข้องได้ดีขึ้นเฉลี่ย 2.23 ตำแหน่ง อย่างไรก็ตาม มีข้อจำกัดในหมวดวิกฤตและหมวดสนับสนุนซึ่งระบบยังมีความรู้ไม่ครอบคลุมและดึงข้อมูลได้น้อย ซึ่งเป็นจุดที่จะต้องพัฒนาเพิ่มเติมในระยะต่อไป ในอนาคตมีแผนที่จะพัฒนา Safe Messaging เพื่อความปลอดภัย, การทำ LoRA fine-tuning บนโมเดล Qwen2.5 ร่วมกับเทคนิค Quantum-Inspired เพื่อพัฒนาโมเดลเฉพาะทางภาษาไทย, การใช้ BiLSTM ร่วมกับ WangchanBERTa ในการตรวจจับอารมณ์และการเปิดเผยชุดข้อมูล open-source เพื่อสนับสนุนการวิจัยด้านสุขภาพจิตในวงกว้างต่อไป

คำสำคัญ: สุขภาพจิต, ปัญญาประดิษฐ์, การคัดกรองซึมเศร้า, RAG, แบบประเมินซึมเศร้า, LLM, Dataset.