บทความวิจัย

การเปรียบเทียบ BERTopic และ LDA สำหรับการจำแนกหัวข้อภาวะซึมเศร้าในข้อความจาก Reddit

ธเนศร เขียวบริบูรณ์ สรวิศ เถาว์ชู อริศรา ยกยอคุณ และ นันทพงศ์ เขียนดวงจันทร์*
สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 6402 8462 อีเมล: nantapong.kean@kmutt.ac.th DOI: 10.14416/j.kmutnb.2026.04.001
รับเมื่อ 10 มิถุนายน 2568 แก้ไขเมื่อ 22 กันยายน 2568 ตอบรับเมื่อ 15 มกราคม 2569 เผยแพร่ออนไลน์ 2 เมษายน 2569
© 2026 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

บทคัดย่อ

ในปัจจุบันภาวะซึมเศร้าและปัญหาสุขภาพจิตเป็นประเด็นที่ทวีความรุนแรงและส่งผลกระทบต่อคุณภาพชีวิตของประชาชนโดยเฉพาะในกลุ่มวัยรุ่นและวัยทำงาน ทั้งนี้ผู้คนจำนวนมากมักแสดงออกถึงความรู้สึกและอาการผ่านสื่อสังคมออนไลน์ซึ่งสามารถนำมาใช้เป็นข้อมูลสำหรับการวิเคราะห์เชิงคอมพิวเตอร์ได้ งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมการจัดกลุ่มหัวข้อ (Topic Modeling) ได้แก่ Latent Dirichlet Allocation (LDA) และ BERTopic โดยใช้ชุดข้อมูลจำนวน 6,397 ข้อความจากแพลตฟอร์ม Reddit ที่เกี่ยวข้องกับภาวะซึมเศร้า การประเมินผลดำเนินการโดยใช้ตัวชี้วัด 3 ประการ ได้แก่ Purity Score, Entropy Score และ Rand Index (RI) ผลการศึกษาแสดงให้เห็นว่า BERTopic มีประสิทธิภาพเหนือกว่า LDA โดยให้ค่า Purity Score สูงกว่า (39.06%) ค่า Entropy ต่ำกว่า (1.93%) และค่า RI สูงกว่า (66.84%) เมื่อเปรียบเทียบกับ LDA ที่ได้ค่า 34.38%, 2.11% และ 65.47% ตามลำดับ สะท้อนถึงความสามารถในการสร้างกลุ่มหัวข้อที่แม่นยำและสอดคล้องกับข้อมูลจริงมากกว่า อย่างไรก็ตามงานวิจัยนี้ยังมีข้อจำกัดจากการใช้ชุดข้อมูลทดสอบเพียง 10% ของข้อมูลทั้งหมด ซึ่งอาจส่งผลต่อความครอบคลุมของการประเมินผล ดังนั้นการศึกษาในอนาคตควรเพิ่มปริมาณข้อมูลทดสอบ รวมทั้งพิจารณาบริบทของข้อความภาษาไทย เพื่อขยายขอบเขตการประยุกต์ใช้งานด้านสุขภาพจิตได้อย่างกว้างขวางยิ่งขึ้น

คำสำคัญ: การประมวลผลภาษาธรรมชาติ การวิเคราะห์หัวข้อ ภาวะซึมเศร้า สุขภาพจิต

การอ้างอิงบทความ: ธเนศร เขียวบริบูรณ์, สรวิศ เถาว์ชู, อริศรา ยกยอคุณ และ นันทพงศ์ เขียนดวงจันทร์, "การเปรียบเทียบ BERTopic และ LDA สำหรับการจำแนกหัวข้อภาวะซึมเศร้าในข้อความจาก Reddit," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 36, ฉบับที่ 2, หน้า 1–11, เม.ย.–มิ.ย. 2569, เลขที่บทความ 262-068126, doi: 10.14416/j.kmutnb.2026.04.001.

Research Article

BERTopic vs. LDA: A Comparative Analysis for Identifying Depression Topics in Reddit Messages

Thanesorn Khiewboriboon, Sorawit Taochoo, Arisara Yokyorkhun and Nantapong Keandoungchun*
Department of Information Technology, Faculty of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand

* Corresponding Author, Tel. 08 6402 8462, E-mail: nantapong.kean@kmutt.ac.th
DOI: 10.14416/j.kmutnb.2026.04.001
Received 10 June 2025 ; Revised 22 September 2025 ; Accepted 15 January 2026; Published online: 2 April 2026
© 2026 King Mongkut's University of Technology North Bangkok. All Rights Reserved.

Abstract

Depression and mental health problems have increasingly become critical issues that significantly affect the quality of life, particularly among adolescents and working-age populations. Many individuals often express their emotions and symptoms through social media platforms, which can serve as valuable sources of data for computational analysis. This study aims to compare the performance of two topic modeling algorithms, Latent Dirichlet Allocation (LDA) and BERTopic, using a dataset of 6,397 depression-related posts collected from Reddit. The evaluation employed three metrics: Purity Score, Entropy Score, and Rand Index (RI). The results demonstrate that BERTopic outperformed LDA, achieving a higher Purity Score (39.06%), lower Entropy (1.93%), and higher RI (66.84%) compared to LDA's 34.38%, 2.11%, and 65.47%, respectively. These findings indicate BERTopic's superior capability in producing co-herent and accurate topic clusters that align more closely with the ground truth. Nevertheless, this study is limited by the use of only 10% of the total dataset for testing, which may affect the comprehensiveness of the evaluation. Therefore, future studies should increase the size of the test set and incorporate Thai-language contexts to broaden the scope of practical applications in mental health research.

Keywords: Natural Language Processing (NLP), Topic Modeling, Depression, Mental Health

Please cite this article as: T. Khiewboriboon, S. Taochoo, A. Yokyorkhun, and N. Keandoungchun, "BERTopic vs. LDA: A comparative analysis for identifying depression topics in Reddit messages," *The Journal of KMUTNB*, vol. 36, no. 2, pp. 1–11, Apr.–Jun. 2026 (in Thai), Art. no. 262-068126, doi: 10.14416/j.kmutnb.2026.04.001.

1. บทนำ

ในปัจจุบันภาวะซึมเศร้าและปัญหาสุขภาพจิตเป็นประเด็นสำคัญที่มีแนวโน้มเพิ่มขึ้นอย่างต่อเนื่อง ส่งผลกระทบโดยตรงต่อการดำเนินชีวิตของบุคคล โดยเฉพาะในกลุ่มวัยรุ่นและวัยทำงาน องค์การอนามัยโลก (WHO) รายงานว่า ประมาณ 1 ใน 8 ของประชากรโลกกำลังเผชิญปัญหาสุขภาพจิต [1] รวมถึงภาวะซึมเศร้า และโรควิตกกังวล หลังจากสถานการณ์การแพร่ระบาดของโควิด-19 ส่งผลประชากรที่เผชิญกับปัญหาความวิตกกังวลหรือภาวะซึมเศร้าสูงขึ้นเป็นจำนวนมาก ประชากรจำนวนมากเผชิญกับปัญหาการเข้าถึงหรือเข้ารับการบริการด้านสุขภาพจิตทำให้ไม่สามารถเข้าถึงการดูแลที่เหมาะสมได้ [2] ในยุคดิจิทัลสื่อสังคมออนไลน์ได้กลายเป็นช่องทางที่ผู้คนใช้ในการแสดงออกถึงอารมณ์ความรู้สึก และประสบการณ์ส่วนตัวเกี่ยวกับสุขภาพจิต ข้อมูลที่ได้รับจากโซเชียลมีเดียจึงกลายเป็นแหล่งข้อมูลที่มีคุณค่า ซึ่งสามารถนำมาวิเคราะห์เพื่อทำความเข้าใจแนวโน้มและลักษณะของภาวะซึมเศร้า โดยเฉพาะแพลตฟอร์ม Reddit ซึ่งมีข้อความที่ยาวและมีรายละเอียดเชิงลึกกว่าทวิตเตอร์ ทำให้สะท้อนประสบการณ์จริงของผู้ใช้ได้ชัดเจนกว่า จากการทบทวนงานวิจัยที่เกี่ยวข้องที่ผ่านมาพบว่ามีการนำเทคนิคการจำแนกหัวข้อข้อความ (Topic Modeling) มาใช้ในงานด้านสุขภาพจิตอย่างหลากหลาย เช่น การจำแนกหัวข้อข้อความจากบทสนทนาสุขภาพจิตโดยใช้ การฝังตัวของประโยค (Sentence Embedding) และ BERTopic ซึ่งให้ผลลัพธ์ดีกว่า Latent Dirichlet Allocation (LDA) โดยเฉพาะเมื่อใช้กับข้อมูลระดับ utterance และ segment [3] นอกจากนี้ยังมีงานวิจัยที่ศึกษาการใช้ LDA, NMF (Non-negative Matrix Factorization), LSA (Latent Semantic Analysis), and PAM (Partitioning Around Medoids) Top2Vec และ BERTopic ในการวิเคราะห์รีวิวลูกค้า และพบว่า BERTopic ให้ผลลัพธ์ที่ชัดเจน และตีความได้ง่าย เนื่องจากสามารถรักษาความสัมพันธ์ทางความหมายของคำได้ดี [4] ในทำนองเดียวกันมีงานวิจัยที่แสดงให้เห็นว่า BERTopic และ Top2Vec มีประสิทธิภาพเหนือกว่า LDA ในการจัดกลุ่มข้อมูลจาก Twitter ซึ่งมีลักษณะเป็นข้อความสั้น และไม่มีโครงสร้างแน่นอน [5] และมีงานวิจัยที่พบว่าแม้ LDA จะมีข้อได้เปรียบในด้านคะแนนความสอดคล้องของหัวข้อเมื่อใช้กับข้อมูลจากแบบสอบถามด้านสุขภาพจิต แต่มีข้อจำกัดเมื่อเปรียบเทียบกับโมเดลใหม่ ๆ เช่น BERTopic ที่สามารถวิเคราะห์แนวโน้มของหัวข้อที่เปลี่ยนแปลงตามเวลาได้อย่างแม่นยำ [6] อีกทั้งยังมีงานวิจัยที่ใช้ LDA ในการวิเคราะห์ข้อมูลจาก Reddit โดยแม้จะสามารถจับกลุ่มหัวข้อได้ดี แต่ยังมีข้อจำกัดในเชิงความหมาย [7] และมีงานวิจัยที่ได้พัฒนา BERTopic โดยผสานความสามารถของโมเดล Bidirectional Encoder Representations from Transformers (BERT), Uniform Manifold Approximation and Projection (UMAP) และ Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) เข้าด้วยกัน พร้อมเทคนิค c-TF-IDF ซึ่งช่วยให้สามารถทำ Topic Modeling ได้อย่างยืดหยุ่น และแม่นยำ แม้ในสภาพแวดล้อมที่ข้อมูลมีการเปลี่ยนแปลงต่อเนื่อง [8] และมีงานวิจัยที่ใช้เทคนิค LDA ในการวิเคราะห์โพสต์กว่า 67,000 รายการ จากฟอรัมออนไลน์ (Online Forum) เกี่ยวกับภาวะซึมเศร้า โดยจำแนกออกเป็น 13 หัวข้อหลัก และจัดกลุ่มเป็น 4 ประเภทของอัตลักษณ์ [9] นอกจากนี้งานของ Salmi และคณะ [3] ที่ใช้ Sentence Embedding ร่วมกับ BERTopic วิเคราะห์ข้อความจากสายด่วนสุขภาพจิตพบว่า BERTopic ให้ผลลัพธ์เหนือกว่า LDA โดยเฉพาะในข้อมูลระดับ utterance และ segment

อย่างไรก็ตามแม้ว่าจะมีหลักฐานสนับสนุนว่า BERTopic มักมีประสิทธิภาพเหนือกว่า LDA แต่ยังไม่พบการศึกษาที่มุ่งเน้นการประยุกต์ใช้ในบริบทของภาวะซึมเศร้าจากข้อมูล Reddit อย่างเป็นระบบ โดยเฉพาะเมื่อเปรียบเทียบกับข้อมูลภาษาไทย จึงถือเป็นช่องว่างของงานวิจัยที่ควรได้รับการศึกษาเพิ่มเติม และนอกจากนี้การจำแนกหัวข้อที่แม่นยำยังสามารถสร้างประโยชน์ต่อทางการแพทย์ เช่น การช่วยวินิจฉัยโรค การติดตามอาการ และการจัดยารักษาที่เหมาะสมต่อผู้ป่วย ดังนั้นงานวิจัยจึงมุ่งเน้นการพัฒนา และเปรียบเทียบประสิทธิภาพของโมเดลการจำแนกหัวข้อข้อความจากข้อมูลที่เกี่ยวข้องกับภาวะซึมเศร้า โดยเปรียบเทียบระหว่าง LDA และ BERTopic โดยนำผลลัพธ์มาเปรียบเทียบด้วยเกณฑ์

ด้านความแม่นยำ ความหลากหลาย และความสอดคล้องกับข้อมูลจริง นอกจากนี้เพื่อเพิ่มขีดความสามารถในการตีความผลลัพธ์ของการจำแนกหัวข้อ งานวิจัยนี้ยังนำปัญญาประดิษฐ์ (AI) ได้แก่ ChatGPT, DeepSeek และ Gemini เข้ามามีบทบาทในการวิเคราะห์ความหมายของหัวข้อที่ได้จากโมเดลโดยที่ไม่ใช้มนุษย์ ทั้งนี้มีจุดประสงค์เพื่อลดอคติในการจำแนกของประเภทของโรคซึมเศร้าในมุมมองที่หลากหลาย อย่างไรก็ตามการใช้ AI ไม่ได้มีเจตนาที่จะเข้ามาลดบทบาทของผู้เชี่ยวชาญทางจิตวิทยา แต่เป็นการใช้ AI เพื่อเพิ่มความหลากหลาย และความแม่นยำจากหลายมุมมอง ซึ่งสามารถเป็นเครื่องมือเสริมในการศึกษา และวิเคราะห์โรคซึมเศร้าได้อย่างมีประสิทธิภาพ

2. วิธีการทดลอง

2.1 ชุดข้อมูล

ในการศึกษานี้ได้ใช้ชุดข้อมูลจากแพลตฟอร์ม Huggingface โดยชุดข้อมูลนี้เป็นการรวบรวมโพสต์จาก ชุมชนย่อย “depression” และ “SuicideWatch” ของแพลตฟอร์ม Reddit เพื่อให้ได้ชุดข้อมูลที่เหมาะสมสำหรับการวิเคราะห์ จึงทำการสุ่มชุดข้อมูลออกมาจำนวนทั้งหมด 6,397 ข้อความ แบ่งออกเป็นชุดข้อมูลฝึกสอน (Training Set) จำนวน 5,117 ข้อความ และชุดข้อมูลทดสอบ (Test Set) จำนวน 1,280 ข้อความ เพื่อใช้ในการประเมินผลของโมเดล อย่างไรก็ตาม ชุดข้อมูลทดสอบที่ใช้มีสัดส่วนเพียง 10% ของข้อมูลทั้งหมด จึงถือเป็นข้อจำกัดของงานวิจัยนี้ ซึ่งอาจส่งผลต่อความครอบคลุมและความแม่นยำในการประเมินผลลัพธ์ของโมเดล

2.2 การเตรียมข้อมูล

งานวิจัยนี้ได้นำชุดข้อมูลฝึกสอน (Training Set) มาดำเนินกระบวนการทำความสะอาดข้อความ (Text Cleaning) เพื่อให้เหมาะสมกับการนำไปใช้ในการวิเคราะห์ต่อไป โดยมีขั้นตอนดังนี้

1) การลบอักขระที่ไม่จำเป็น เช่น เครื่องหมายวรรคตอน (Punctuation Marks) แท็ก HTML ลิงก์ URL อักขระ Unicode ที่ไม่ใช่ตัวอักษร และตัวเลขที่ไม่มีความสำคัญต่อความหมายข้อความ

2) การแปลงตัวอักษร (Text Normalization) ทำการแปลงตัวอักษรทั้งหมดเป็นตัวพิมพ์เล็ก (lowercasing) เพื่อป้องกันความซ้ำซ้อน เช่น “Depression” และ “depression” ถูกมองว่าเป็นคำเดียวกัน

3) การตัดคำ (Tokenization) แยกข้อความออกเป็นหน่วยคำ (Tokens) ที่สามารถนำไปวิเคราะห์ได้

4) การกำจัดคำที่ไม่จำเป็น (Stopwords Removal) ลบคำทั่วไปที่ไม่ช่วยเพิ่มความหมายในการวิเคราะห์ เช่น the, is, was, and

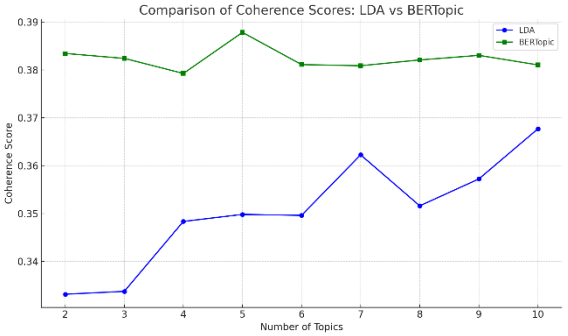
5) การลดรูปคำ (Stemming & Lemmatization)

5.1) Stemming ตัดคำให้เป็นรูปแบบพื้นฐาน เช่น running → run

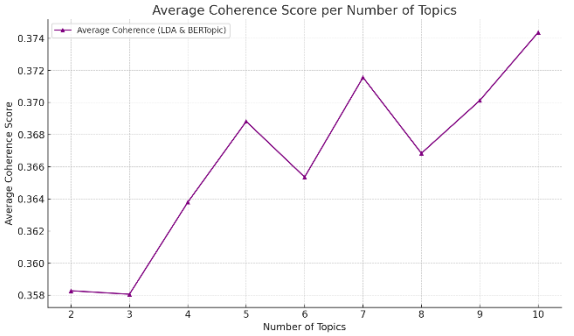
5.2) Lemmatization แปลงคำให้เหลือรากศัพท์ที่ถูกต้องตามไวยากรณ์ เช่น better → good

สำหรับ LDA หลังจากทำการทำความสะอาดและลดรูปคำแล้ว ข้อความจะถูกแปลงให้อยู่ในรูปแบบถุงคำ (Bag of Words; BoW) เพื่อแทนข้อความเป็นเวกเตอร์เชิงสถิติ โดยนับจำนวนการปรากฏของแต่ละคำในเอกสาร การแทนข้อมูลด้วยวิธีนี้เหมาะสมกับ LDA เนื่องจากอาศัยความถี่ของคำในการจำแนกหัวข้อ อย่างไรก็ตาม BoW มีข้อจำกัดตรงที่ไม่สามารถรักษาลำดับและความสัมพันธ์เชิงความหมายของคำได้

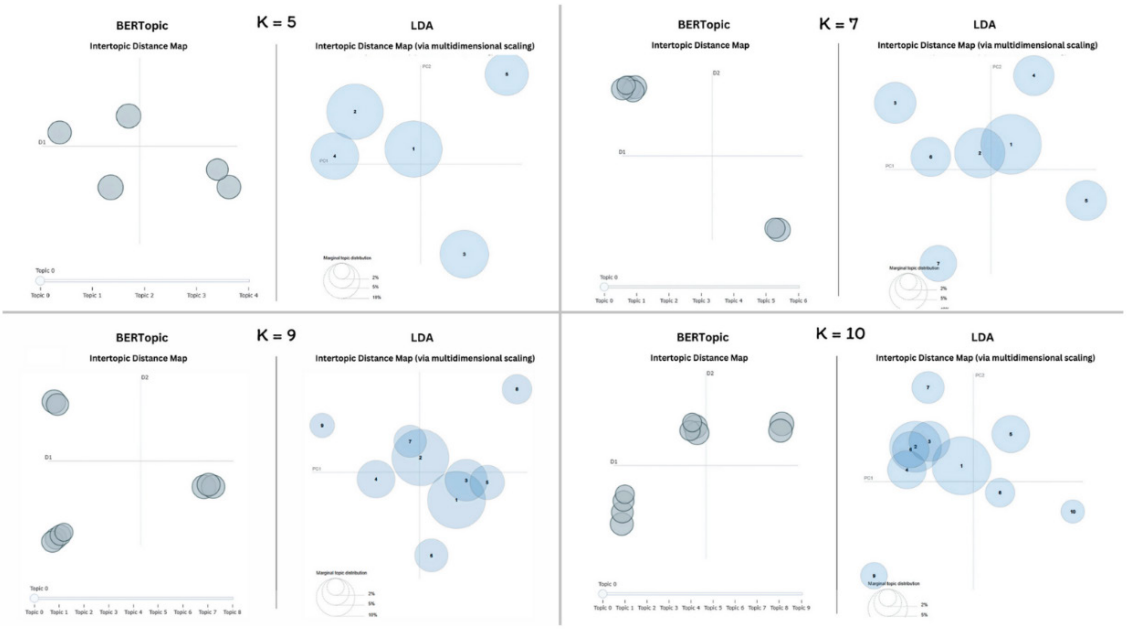
สำหรับ BERTopic ใช้วิธีแนวทางการฝังตัว (Embedding Approach) โดยนำข้อความแต่ละประโยคไปผ่านโมเดลภาษาที่ใช้โครงสร้าง Transformer โดยเฉพาะ Sentence-Transformers (เช่น all-MiniLM-L6-v2) เพื่อสร้างเวกเตอร์เชิงความหมาย (Sentence Embeddings) ที่สามารถสะท้อนความสัมพันธ์เชิงบริบทของคำและประโยคได้ดีกว่าการนับคำแบบ BoW จากนั้นเวกเตอร์ที่ได้จะถูกนำไปลดมิติด้วย UMAP เพื่อลดจำนวนมิติให้เหลือเพียง 5 มิติ พร้อมทั้งคงโครงสร้างเชิงความหมายของข้อมูลไว้ ก่อนจะนำไปจัดกลุ่มด้วย KMeans/HDBSCAN และแปลงคำสำคัญแต่ละหัวข้อด้วย c-TF-IDF เพื่อให้สามารถตีความผลลัพธ์ได้ง่ายขึ้น



รูปที่ 1 การเปรียบเทียบค่า Coherence Score ของแต่ละจำนวนหัวข้อ ระหว่างโมเดล LDA และ BERTopic



รูปที่ 2 ค่าเฉลี่ย Coherence Score ของโมเดล LDA และ BERTopic ตามจำนวนหัวข้อ



รูปที่ 3 แผนภาพระยะห่างระหว่างหัวข้อจากวิธี BERTopic และ LDA

2.3 การสร้าง และปรับแต่งโมเดล

2.3.1 การหาจำนวนหัวข้อที่เหมาะสม

งานวิจัยนี้ได้ทำการเปรียบเทียบค่าความสอดคล้อง (Coherence Score; c_v) ของโมเดลการจัดกลุ่มหัวข้อ ได้แก่ LDA และ BERTopic เพื่อหาจำนวนหัวข้อ (k) ที่เหมาะสมที่สุดในการวิเคราะห์ข้อมูล โดยทำการทดลองในช่วงจำนวนหัวข้อ k = 2 ถึง 10 และนำผลลัพธ์มาแสดงในรูปที่ 1

ผลการทดลองแสดงให้เห็นว่า BERTopic มีค่า Coherence Score สูงกว่า LDA อย่างสม่ำเสมอในทุกค่า k โดยมีค่าสูงสุดที่ k = 5 (≈ 0.39) ในขณะที่ LDA มีค่า Coherence Score ต่ำกว่าอย่างเห็นได้ชัด และมีแนวโน้มเพิ่มขึ้นตามจำนวนหัวข้อ (รูปที่ 1) นอกจากนี้เพื่อให้การเปรียบเทียบมีความชัดเจนมากยิ่งขึ้น จึงได้ทำการคำนวณค่าเฉลี่ยของ Coherence Score ระหว่างทั้งสองโมเดลในแต่ละค่า k ดังแสดงในรูปที่ 2 ซึ่งพบว่าค่าเฉลี่ยสูงสุดอยู่ที่ k = 7 และ k = 10 เพื่อยืนยันผลลัพธ์เชิงโครงสร้างของ

หัวข้อ จึงได้ทำการสร้างแผนภาพระยะห่างระหว่างหัวข้อ (Intertopic Distance Map) ของทั้ง LDA และ BERTopic สำหรับค่า k = 5, 7, 9 และ 10 (รูปที่ 3) ผลการวิเคราะห์พบว่า

- ที่ k = 5 กลุ่มหัวข้อที่ได้จาก BERTopic แสดงการกระจายตัวที่ชัดเจน ไม่ทับซ้อนกัน และสามารถแยกแยะเนื้อหาได้ง่าย ในขณะที่ LDA ยังมีการทับซ้อนบางส่วน
- ที่ k = 7 และ k = 10 แม้ค่าเฉลี่ย Coherence Score จะสูง แต่จากแผนภาพพบว่ามีการทับซ้อนระหว่างหัวข้อในระดับสูงขึ้น ทำให้การตีความเชิงเนื้อหามีความซับซ้อนและไม่สามารถระบุลักษณะเฉพาะของแต่ละกลุ่มได้อย่างชัดเจน
- ที่ k = 9 ผลลัพธ์แสดงการกระจายตัวที่ไม่สม่ำเสมอ โดยบางกลุ่มหัวข้อมีการรวมตัวแน่นหนาเกินไป ขณะที่บางกลุ่มมีขนาดเล็กเกินไป ทำให้เกิดความไม่สมดุลของหัวข้อ

จากการวิเคราะห์ทั้งหมดนี้งานวิจัยจึงเลือก k = 5 เป็นจำนวนหัวข้อที่เหมาะสม เนื่องจากนอกจากจะได้ค่า Coherence Score ที่ค่อนข้างสูงแล้ว ยังมีการกระจายตัวของหัวข้อที่ชัดเจนที่สุดไม่เกิดการทับซ้อนมากเกินไป ส่งผลให้สามารถตีความและอธิบายลักษณะเฉพาะของแต่ละหัวข้อได้อย่างชัดเจน และรองรับการนำผลไปใช้งานเชิงปฏิบัติได้ดีกว่าค่าของ k อื่น ๆ

2.3.2 การสร้างโมเดล

โมเดลที่ 1 Latent Dirichlet Allocation

สำหรับการสร้างและปรับแต่งโมเดล LDA ได้ใช้ฟังก์ชัน LdaModel จากไลบรารี Gensim โดยแปลงเอกสารให้อยู่ในรูปแบบคลังข้อความ (corpus) ซึ่งเป็นการแทนข้อความในรูปแบบ BoW จากนั้นกำหนดจำนวนหัวข้อ (k) ตามผลการเลือกจำนวนหัวข้อที่เหมาะสม (k = 5)

ในการฝึกโมเดล ได้กำหนดพารามิเตอร์ที่สำคัญดังนี้

1) passes กำหนดจำนวนรอบการเรียนรู้เพื่อให้โมเดลสามารถปรับปรุงการกระจายของคำได้ดีขึ้น

2) chunksize กำหนดจำนวนเอกสารที่ใช้ในแต่ละรอบการอัปเดตพารามิเตอร์เพื่อให้การเรียนรู้มีความเสถียร

3) Alpha (α) ใช้ควบคุมการกระจายของหัวข้อในแต่ละเอกสาร หากค่า α สูง เอกสารหนึ่ง ๆ จะมีหลายหัวข้อ แต่ถ้าค่า α ต่ำ เอกสารจะมีหัวข้อหลักเพียงไม่กี่หัวข้อ

4) Beta (β) ใช้ควบคุมการกระจายของคำในแต่ละหัวข้อ หากค่า β สูง แต่ละหัวข้อจะมีคำที่หลากหลาย แต่ถ้าค่า β ต่ำ หัวข้อจะมีคำที่จำกัดและเฉพาะเจาะจง

การเลือกค่าพารามิเตอร์นี้มีความสำคัญ เนื่องจากมีผลโดยตรงต่อการตีความหัวข้อ หากค่าของ α และ β ไม่เหมาะสมอาจทำให้หัวข้อที่ได้ไม่ชัดเจนหรือมีการซ้ำซ้อนสูง

โมเดลที่ 2 BERTopic

สำหรับการสร้างและปรับแต่งโมเดล BERTopic ได้ใช้โครงสร้างที่ประกอบด้วย 3 ขั้นตอนหลัก ได้แก่

1) Document Embedding ใช้โมเดล all-MiniLM-L6-v2 จาก Sentence-Transformers เพื่อแปลงข้อความแต่ละประโยคเป็นเวกเตอร์เชิงความหมายที่สะท้อนความสัมพันธ์เชิงบริบทของคำและข้อความได้อย่างมีประสิทธิภาพ

2) Dimensionality Reduction ใช้เทคนิค UMAP ในการลดมิติของเวกเตอร์จากหลายร้อยมิติให้เหลือ 5 มิติ เพื่อให้การจัดกลุ่มมีประสิทธิภาพ โดยตั้งค่า

n_neighbors = 15 → กำหนดจำนวนเพื่อนบ้านที่ใช้ในการเรียนรู้โครงสร้างข้อมูล (balance ระหว่าง global/local structure)

min_dist = 0.1 → ควบคุมความหนาแน่นของกลุ่มที่ได้ (ค่าเล็กทำให้คลัสเตอร์แน่นขึ้น)

n_components = 5 → กำหนดจำนวนมิติที่ต้องการหลังการลดมิติ

3) Clustering and Topic Representation ใช้อัลกอริทึม KMeans โดยกำหนดจำนวนกลุ่มเท่ากับ k = 5 (ตามผลการวิเคราะห์ Section 2.3.1) ใช้ k-means++ ในการ initialization เพื่อให้การสุ่มจุดเริ่มต้นมีเสถียรภาพ ตั้งค่า max_iter = 300 เพื่อให้การเรียนรู้ค่าศูนย์กลางกลุ่ม (centroids) มีความแม่นยำ และใช้ CountVectorizer ร่วมกับการกำจัด Stopwords เพื่อสร้าง c-TF-IDF (class-based TF-IDF) สำหรับการดึงคำสำคัญที่แทนแต่ละหัวข้อ BERTopic จึงสามารถรักษาความสัมพันธ์เชิงความหมายของข้อความได้ดีกว่า LDA ที่อิงเพียงการนับคำ (BoW) และทำให้ผลลัพธ์การจำแนกหัวข้อมีความสอดคล้องกับเนื้อหาจริงมากกว่า

2.4 วิธีการประเมินผล

หลังจากฝึกโมเดลกับชุดข้อมูลฝึกสอนเรียบร้อยแล้ว งานวิจัยนี้ได้นำโมเดลที่ได้ไปทดสอบกับชุดข้อมูลทดสอบ (Test Set) โดยใช้การสุ่มตัวอย่างแบบ Stratified Sampling จำนวน 128 ข้อความ (คิดเป็น 10% ของชุดข้อมูลทดสอบทั้งหมด) เพื่อให้การสุ่มมีการกระจายตัวอย่างสมดุลในแต่ละประเภทของหัวข้อ ทั้งนี้ได้มีการบันทึกโมเดล LDA และ BERTopic ที่ผ่านการฝึกมาแล้ว เพื่อให้สามารถเรียกใช้งานซ้ำได้ในการทำนายหัวข้อของข้อความในชุดทดสอบ

2.4.1 การกำหนดข้อมูลจริง

การกำหนดข้อมูลจริง (Ground Truth) ในงานวิจัยนี้ได้ใช้ AI 3 ตัว ได้แก่ ChatGPT, DeepSeek และ Gemini มาช่วยในการจำแนกประเภทของข้อความแต่ละชุด ข้อความจะถูกกำหนดหัวข้อโดยใช้หลักการ การโหวตเสียงข้างมาก (Majority Voting) คือ หากมีอย่างน้อย 2 ใน 3 AI ให้ผลตรงกันจะถือว่าผลลัพธ์นั้นเป็นข้อมูลจริง

อย่างไรก็ตามหากไม่มี AI ใดให้ผลตรงกันเลย จึงได้ใช้วิธี Weighting Score ตามสมการที่ (1)

$$Weight_i = \frac{C_i}{\sum_{j=1}^{n} C_j} \tag{1}$$

โดยที่

C_i คือ จำนวนครั้งที่ AI ตัวที่ i ทำนายได้ตรงกับผลที่ยืนยันแล้วจากข้อความอื่น ๆ

n คือ จำนวน AI ทั้งหมด (ในที่นี้คือ 3 ตัว)

ผลลัพธ์สุดท้ายจะเลือกจาก AI ที่มีค่าน้ำหนักสูงที่สุดเพื่อนำมาใช้เป็นข้อมูลจริง

หมายเหตุ การเลือกใช้ AI เพื่อช่วยสร้างข้อมูลจริงมีวัตถุประสงค์เพื่อลดอคติจากการประเมินโดยผู้เชี่ยวชาญเพียงคนเดียว และเปิดโอกาสให้ได้ผลการจำแนกที่สะท้อนจากหลายมุมมอง แม้ว่าจะไม่สามารถทดแทนผู้เชี่ยวชาญทางจิตวิทยาได้โดยตรง แต่ถือเป็นแนวทางเสริมเพื่อให้การวิเคราะห์หัวข้อมีความหลากหลายและยืดหยุ่นมากขึ้น

2.4.2 การวัดประสิทธิภาพของโมเดล

เพื่อนำมาใช้เปรียบเทียบความสามารถของ LDA และ BERTopic งานวิจัยนี้จึงเลือกใช้ตัวชี้วัดจาก Clustering Evaluation Metrics ดังนี้

1) Purity Score ใช้วัดความแม่นยำของการจัดกลุ่ม โดยตรวจว่าข้อมูลแต่ละกลุ่มสอดคล้องกับข้อมูลจริงได้มากน้อยเพียงใด หากค่ามีค่าใกล้ 1 หมายถึงการจัดกลุ่มมีความแม่นยำสูง ดังสมการที่ (2)

$$Purity = \left(\frac{1}{N}\right)\sum_{k=1}^{K}\left|C_k \cap L_j\right| \tag{2}$$

โดยที่

N คือ จำนวนข้อมูลทั้งหมด

K คือ จำนวนคลัสเตอร์

C คือ คลัสเตอร์ที่ k

L_j คือ ข้อมูลในคลาสจริง (Ground Truth) ที่ j

$\left|C_k \cap L_j\right|$ คือ จำนวนข้อมูลในคลัสเตอร์ C_k ที่อยู่ในคลาสจริง L_j ที่มีมากที่สุด

2) Entropy Score ใช้วัดความไม่เป็นเอกภาพของคลัสเตอร์ หากค่า Entropy ต่ำ หมายถึงกลุ่มมีความชัดเจนและไม่ซ้ำซ้อน ดังสมการที่ (3)

$$Entropy = \sum_{k=1}^{K}\left(\frac{|C_k|}{N}\right)E(C_k) \tag{3}$$

โดยที่

K คือ จำนวนคลัสเตอร์ทั้งหมด

C_k คือ คลัสเตอร์ที่ k

$|C_k|$ คือ จำนวนข้อมูลในคลัสเตอร์ C_k

N คือ จำนวนข้อมูลทั้งหมด

$E(C_k)$ คือ ค่า entropy ภายในคลัสเตอร์ C_k

3) Rand Index (RI) ใช้วัด ความสอดคล้องระหว่างผลการจัดกลุ่มของโมเดลกับข้อมูลจริง โดยพิจารณาการจับคู่ของข้อมูลทุกคู่ว่ามีการจัดกลุ่มถูกต้องหรือไม่ ดังสมการที่ (4)

$$RI = \frac{(TP+TN)}{(TP+FP+FN+TN)} \tag{4}$$

โดยที่

TP (True Positive) คือ จำนวนคู่ข้อมูลที่อยู่ในกลุ่มเดียวกันทั้งจากโมเดล และข้อมูลจริง

TN (True Negative) คือ จำนวนคู่ข้อมูลที่อยู่คนละกลุ่มทั้งจากโมเดล และข้อมูลจริง

FP (False Positive) คือ จำนวนคู่ข้อมูลที่โมเดลจัดให้อยู่ในกลุ่มเดียวกัน แต่ข้อมูลจริงอยู่คนละกลุ่ม

FN (False Negative) คือ จำนวนคู่ข้อมูลที่โมเดลจัดให้อยู่คนละกลุ่ม แต่ข้อมูลจริงอยู่ในกลุ่มเดียวกัน

3. ผลการทดลองและอภิปรายผลการทดลอง

3.1 ผลการจำแนกหัวข้อ

งานวิจัยนี้ได้ดำเนินการสร้างและปรับแต่งโมเดลการจำแนกหัวข้อ โดยใช้เทคนิค LDA และ BERTopic เพื่อจำแนกประเภทของโรคซึมเศร้าจากข้อความ ผลลัพธ์ที่ได้จากการฝึกโมเดลถูกนำมาวิเคราะห์เพื่อระบุหัวข้อ พร้อมคำสำคัญที่เป็นตัวแทนของแต่ละหัวข้อ ดังแสดงในตารางที่ 1 และ 2

นอกจากนี้งานวิจัยยังใช้ AI 3 ตัว ได้แก่ ChatGPT, DeepSeek และ Gemini มาช่วยในการตีความความหมายของหัวข้อ และจัดกลุ่มอาการให้สัมพันธ์กับโรคซึมเศร้าที่พบบ่อย 7 ประเภท (MDD, PDD, Bipolar Disorder, PPD, PMDD, SAD, Atypical Depression) โดยพิจารณาคำสำคัญที่ปรากฏในแต่ละหัวข้อ ผลลัพธ์แสดงในตารางที่ 1 และตารางที่ 2

ตารางที่ 1 การจำแนกประเภทของโรคซึมเศร้าโดย AI จากผลลัพธ์ของโมเดล LDA

หัวข้อ (Types)	ประเภทของโรคซึมเศร้า (Types of Depression Disorders)	คำที่เกี่ยวข้อง (Related Terms)
0	ภาวะซึมเศร้าแบบผิดปกติ (Atypical depression)	friend, know, year, go, tell, get, talk, school, try, time
1	โรคซึมเศร้าระดับรุนแรง (Major Depressive Disorder; MDD)	depress, help, know, need, people, time, think, tell, thing, try
2	โรคซึมเศร้าเรื้อรัง (Persistent Depressive Disorder; PDD)	work, job, get, go, year, time, life, anxiety, start, home
3	โรคอารมณ์สองขั้ว หรือโรคไบโพลาร์ (Bipolar Disorder)	suicide, tired, way, hurt, love, time, think, die, kill, hope
4	ภาวะซึมเศร้าหลังคลอด (Postpartum Depression; PPD)	Life, fuck, live, people, care, know, anymore, hate, kill, thing

ตารางที่ 2 การจำแนกประเภทของโรคซึมเศร้าโดย AI จากผลลัพธ์ของโมเดล BERTopic

หัวข้อ (Types)	ประเภทของโรคซึมเศร้า (Types of Depression Disorders)	คำที่เกี่ยวข้อง (Related Terms)
0	ภาวะซึมเศร้าแบบผิดปกติ (Atypical depression)	friends, therapy, taking, suicide, ago, suicidal, started, having, medication, worse
1	โรคซึมเศร้าระดับรุนแรง (Major Depressive Disorder; MDD)	thinking, reason, birthday, happy, say, tried, thing, things, hope, good
2	โรคซึมเศร้าเรื้อรัง (Persistent Depressive Disorder; PDD)	trying, bad, shit, try, point, home, thing, depressed, end, live
3	โรคอารมณ์สองขั้ว หรือโรคไบโพลาร์ (Bipolar Disorder)	say, sad, depressed, having, makes, hope, years, hard, bad, reason
4	ภาวะซึมเศร้าหลังคลอด (Postpartum Depression; PPD)	better, thought, tried, happy, relationship, school, live, away, good, right, best

3.2 ผลการเปรียบเทียบด้วย Clustering Evaluation Metrics

ตารางที่ 3–5 แสดงผลการเปรียบเทียบที่ได้จาก AI แต่ละตัว และตารางที่ 6 แสดงผลสรุปรวม โดยมีข้อสังเกตดังนี้

1) Purity Score

โดยรวม BERTopic มีค่า Purity สูงกว่า LDA (0.3906 vs 0.3438 ตามตารางที่ 6) ซึ่งสะท้อนว่า BERTopic มีความสามารถในการจัดกลุ่มข้อความให้อยู่ในหัวข้อที่ใกล้เคียงกับข้อมูลจริงได้ดีกว่า อย่างไรก็ตามจากตารางที่ 5 (Gemini) พบว่า LDA และ BERTopic ให้ค่า Purity เท่ากันที่ 0.7422 แสดงว่าในบางกรณี LDA สามารถจัดกลุ่มหัวข้อได้ดี โดยเฉพาะเมื่อข้อความมีคำซ้ำที่ชัดเจน

ตารางที่ 3 ผลการเปรียบเทียบระหว่างโมเดล LDA และ BERTopic ของผลลัพธ์ที่ได้จาก Deepseek

โมเดล (Model)	Purity Score	Entropy Score	RI Score
LDA	0.4062	1.9602	0.6479
BERTopic	0.4219	1.7918	0.6572

ตารางที่ 4 ผลการเปรียบเทียบระหว่างโมเดล LDA และ BERTopic ของผลลัพธ์ที่ได้จาก ChatGPT

โมเดล (Model)	Purity Score	Entropy Score	RI Score
LDA	0.2891	2.1842	0.6512
BERTopic	0.4219	1.9397	0.6766

ตารางที่ 5 ผลการเปรียบเทียบระหว่างโมเดล LDA และ BERTopic ของผลลัพธ์ที่ได้จาก Gemini

โมเดล (Model)	Purity Score	Entropy Score	RI Score
LDA	0.7422	1.1925	0.4628
BERTopic	0.7422	1.1412	0.4592

ตารางที่ 6 ผลการเปรียบเทียบระหว่างโมเดล LDA และ BERTopic

โมเดล (Model)	Purity Score	Entropy Score	RI Score
LDA	0.3438	2.1115	0.6547
BERTopic	0.3906	1.9325	0.6684

2) Entropy Score

BERTopic มีค่า Entropy ต่ำกว่า (1.9325 เทียบกับ 2.1115) ซึ่งแปลว่าผลลัพธ์ของ BERTopic มีความเป็นเอกภาพของหัวข้อสูงกว่า และมีความซ้ำซ้อนน้อยกว่า ค่า Entropy ของ LDA สูงกว่า แสดงว่าเนื้อหาที่จัดกลุ่มมักทับซ้อนกัน และมีความหลากหลายภายในกลุ่มมากเกินไป

3) Rand Index (RI)

BERTopic มีค่า RI สูงกว่าเล็กน้อย (0.6684 vs 0.6547) สะท้อนว่าการจัดกลุ่มมีความสอดคล้องกับข้อมูลจริงมากกว่า LDA แม้จะไม่ต่างกันมาก ในผลลัพธ์ของ Gemini (ตารางที่ 5) พบว่า LDA ให้ค่า RI สูงกว่าเล็กน้อย ซึ่งแสดงว่าในบางสถานการณ์ LDA ยังสามารถจับคู่ความสัมพันธ์ของข้อความกับกลุ่มได้ดี

3.3 อภิปรายผล

ผลการทดลองแสดงให้เห็นว่า BERTopic มีประสิทธิภาพโดยรวมสูงกว่า LDA ในการจำแนกหัวข้อของข้อความที่เกี่ยวข้องกับภาวะซึมเศร้า ทั้งในด้านความแม่นยำ (Purity) ความชัดเจนของหัวข้อ (Entropy ต่ำกว่า) และความสอดคล้องกับข้อมูลจริง (RI สูงกว่า) สาเหตุหลักที่ BERTopic ทำงานได้ดีกว่ามาจากความสามารถของโมเดลที่อิงโครงสร้างแบบ Transformer-based Embedding ซึ่งสามารถรักษาความหมายเชิงบริบทของข้อความได้อย่างลึกซึ้ง ในขณะที่ LDA พิจารณาเพียงสถิติความถี่ของคำ (Bag-of-Words) ทำให้คำทั่วไปบางคำถูกระบุว่าเป็นคำสำคัญ แม้จะไม่ได้สะท้อนอาการทางจิตใจที่แท้จริง อย่างไรก็ตาม LDA ยังมีข้อดีบางประการ โดยเฉพาะในกรณีที่ข้อความมีลักษณะตรงไปตรงมาและมีคำซ้ำสูง เช่น ข้อความที่สะท้อนความรู้สึกซ้ำ ๆ อย่าง sad, tired หรือ hopeless ซึ่งในกรณีเช่นนี้ LDA สามารถจัดกลุ่มได้อย่างมีประสิทธิภาพและให้ค่า Purity ใกล้เคียงกับ BERTopic ได้เช่นกัน ผลลัพธ์นี้สอดคล้องกับงานวิจัยที่ผ่านมา [3]–[7] ที่พบว่า BERTopic เหมาะกับข้อความที่ไม่เป็นโครงสร้างและสั้น (เช่น โพสต์บนโซเชียลมีเดีย) ขณะที่ LDA ยังคงมีประโยชน์ในข้อมูลที่มีการซ้ำของคำมาก หรือต้องการการตีความเชิงสถิติ ซึ่งข้อจำกัดของงานวิจัยนี้ยังมีข้อจำกัดที่สำคัญคือ การใช้ข้อมูลทดสอบเพียง 10% ของชุดข้อมูลทั้งหมด ซึ่งอาจทำให้การประเมินประสิทธิภาพของโมเดลยังไม่ครอบคลุมเพียงพอ นอกจากนี้การกำหนดข้อมูลจริง โดยอาศัย AI 3 ตัว (ChatGPT, DeepSeek, Gemini) แม้จะช่วยลดอคติของมนุษย์ผู้ประเมิน แต่ยังคงเป็นแนวทางเชิงทดลองที่อาจต้องมีการตรวจสอบเพิ่มเติมโดยผู้เชี่ยวชาญด้านสุขภาพจิตเพื่อยืนยันความถูกต้อง

ผลลัพธ์ที่ได้จากการจำแนกหัวข้อข้อความเกี่ยวกับภาวะซึมเศร้าสามารถนำไปประยุกต์ใช้ในทางปฏิบัติได้หลายด้าน เช่น การพัฒนาระบบเฝ้าระวัง (Surveillance System) เพื่อใช้ติดตามแนวโน้มของโรคซึมเศร้าผ่านข้อมูลในสื่อสังคมออนไลน์ การสนับสนุนผู้เชี่ยวชาญด้านสุขภาพจิตในการคัดกรองผู้ที่มีความเสี่ยงสูง ตลอดจนการใช้เป็นฐานข้อมูลสำหรับพัฒนา Chatbot หรือผู้ช่วยเสมือน (Virtual Assistant) เพื่อให้คำแนะนำเบื้องต้นด้านสุขภาพจิต นอกจากนี้งานวิจัยในอนาคตควรมุ่งเพิ่มปริมาณและความหลากหลายของชุดข้อมูล รวมถึงข้อมูลจากภาษาไทยเพื่อทดสอบความสามารถของโมเดลในหลายบริบท อีกทั้งอาจพิจารณานำโมเดลใหม่ ๆ เช่น Top2Vec หรือ CombinedTM มาทดลองเปรียบเทียบเพิ่มเติม รวมถึงการพัฒนาวิธี Hybrid Ground Truth ที่ผสมผสานการประเมินจาก AI และผู้เชี่ยวชาญ เพื่อเพิ่มความน่าเชื่อถือของผลลัพธ์ นอกจากนี้ยังควรพิจารณาการใช้ Dynamic Topic Modeling เพื่อวิเคราะห์แนวโน้มการเปลี่ยนแปลงของอาการซึมเศร้าในเชิงเวลา (Temporal Analysis) เพื่อทำความเข้าใจพลวัตของภาวะซึมเศร้าได้อย่างครอบคลุมยิ่งขึ้น

เมื่อเปรียบเทียบผลการทดลองในงานวิจัยนี้กับงานวิจัยอื่น ๆ พบว่ามีแนวโน้มที่สอดคล้องกันเกี่ยวกับศักยภาพของโมเดลแบบ Embedding-based อย่างชัดเจน อาทิ ในการจัดกลุ่มข้อความด้านสุขภาพจิต Ma และคณะ [10] ได้ประเมิน BERTopic และ LDA ด้วยข้อมูลทางการแพทย์จาก PubMed โดยมีการรายงานว่า BERTopic สามารถแยกหัวข้อได้เฉพาะเจาะจง และตีความได้ชัดเจนกว่า LDA โดยเฉพาะเมื่อบูรณาการร่วมกันกับโมเดลภาษาขนาดใหญ่ (LLMs) เพื่อช่วยยกระดับการสรุปและปรับโครงสร้างหัวข้อ ทั้งนี้ยังพบว่าการใช้ Embedding ช่วยลดความซ้อนทับของหัวข้อและเพิ่มคุณภาพของการตีความเชิงคลินิก ในงานของ Qasim และคณะ [11] ซึ่งประเมินความรุนแรงของภาวะซึมเศร้าจากข้อความบนสื่อสังคมออนไลน์พบว่าโมเดลที่ใช้ตัวแทนความหมายเชิงบริบท (Context-based/Embedding-based) สามารถจับสัญญะทางอารมณ์และลักษณะภาษาที่ละเอียดอ่อนได้ดีกว่าโมเดลเชิงสถิติแบบดั้งเดิม ผลลัพธ์ดังกล่าวชี้ให้เห็นว่าเทคนิคที่อิงการแทนข้อมูลเชิงความหมาย (semantic representation) เหมาะสำหรับการประมวลผลข้อมูลสุขภาพจิตที่ไม่เป็นโครงสร้างและเปลี่ยนแปลงตามบริบท เช่น ข้อความจากผู้ใช้บนแพลตฟอร์มโซเชียล

ในทำนองเดียวกันกับงานของ Salmi และคณะ [3] แสดงให้เห็นว่า Embedding-based Topic Modeling สามารถจัดกลุ่มข้อความได้อย่างมีประสิทธิภาพ โดยเฉพาะในบริบทของบทสนทนาสุขภาพจิตที่มีโครงสร้างหลวม และเปลี่ยนแปลงตามบริบทการสนทนา ส่งผลให้หัวข้อมีความสอดคล้องกับความหมายเชิงคลินิกและสะท้อนลักษณะอาการได้อย่างแม่นยำ

สุดท้ายผลการทดลองในงานวิจัยนี้ก็แสดงให้เห็นว่า BERTopic ให้ค่า Purity สูงกว่า ค่า Entropy ต่ำกว่า และค่า RI สูงกว่า LDA จึงมีความสอดคล้องกับผลการทดลองของงานวิจัยทั้งสาม ซึ่งต่างสนับสนุนว่าโมเดลที่อิง Embedding ร่วมกับเทคนิคลดมิติขั้นสูง เช่น UMAP และ HDBSCAN มีศักยภาพในการจัดการข้อมูลที่ไม่เป็นโครงสร้างและซับซ้อน เช่น ข้อความจากโซเชียลมีเดียได้ดีกว่าโมเดลเชิงสถิติอย่างมีนัยสำคัญ ขณะที่ LDA ยังคงมีข้อจำกัดในการตีความบริบทและความหมายที่แฝงอยู่ในข้อความประเภทนี้

4. สรุปผลการทดลอง

จากผลการทดลองสามารถสรุปได้ว่า โมเดล BERTopic มีความสามารถเหนือกว่า LDA ในบริบทของชุดข้อมูลนี้ โดยมีค่าความแม่นยำสูงกว่าเล็กน้อยในด้าน Purity Score และ Rand Index (RI) ซึ่งสะท้อนถึงความสามารถในการจัดกลุ่มที่สอดคล้องกับข้อมูลจริงได้ดีกว่า ขณะเดียวกัน BERTopic ยังมีค่า Entropy Score ต่ำกว่า ซึ่งแสดงถึงความชัดเจนของกลุ่มที่ดีกว่า LDA อย่างไรก็ตามการเลือกใช้งานระหว่างโมเดลทั้งสองควรพิจารณาจากวัตถุประสงค์ของการวิเคราะห์ และลักษณะของชุดข้อมูล หากต้องการเน้นความเข้าใจเชิงบริบท และความหลากหลายของภาษา BERTopic จะเป็นทางเลือกที่เหมาะสมกว่า แต่หากต้องการผลลัพธ์ที่ตีความง่าย และมีโครงสร้างเชิงสถิติที่ชัดเจน LDA ยังคงเป็นโมเดลที่ตอบโจทย์ได้ดี

เอกสารอ้างอิง

[1] World Health Organization. (2022). Mental Disorders. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

[2] World Health Organization. (2022). Mental Health and COVID-19: Early Evidence of the Pandemic's Impact: Scientific Brief. [Online]. Available: https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief-Mental_health-2022.1.

[3] S. Salmi, R. v. d. Mei, S. Mérelle, and S. Bhulai, "Topic modeling for conversations for mental health helplines with utterance embedding," *Journal of Computational Social Science*, vol. 13, 2024, doi: 10.1016/j.teler.2024.100126.

[4] A. Krishnan and P. Kennedyraj, "Exploring the power of topic modeling techniques in analyzing customer reviews: A comparative analysis," *arXiv*, 2023, doi: 10.48550/arXiv.2308.11520.

[5] R. Egger and J. Yu, "A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts," *Frontiers in Sociology*, vol. 7, 2022, doi: 10.3389/fsoc.2022.886498.

[6] A. Rkia, A. Fatima-Azzahrae, A. Mehdi, and L. Lily, "NLP and topic modeling with LDA, LSA, and NMF for monitoring psychosocial well-being in monthly surveys," *Procedia Computer Science*, vol. 251, pp. 398-405, 2024, doi: 10.1016/j.procs.2024.11.126.

[7] A. Khan and R. Ali, "Measuring the effectiveness of LDA-based clustering for social media data," 2022, doi: 10.37394/232025.2022.4.11.

[8] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," *arXiv*, 2022, doi: 10.48550/arXiv.2203.05794.

[9] D. Sik, R. Németh, and E. Katona, "Topic modelling online depression forums: Beyond narratives of self-objectification and self-blaming," *Journal of Affective Disorders Reports*, vol. 32, no. 2, pp. 386–395, 2021, doi: 10.1080/09638237.2021.1979493.

[10] L. Ma, R. Chen, W. Ge, P. Rogers, B. Lyn-Cook, H. Hong, W. Tong, N. Wu, and W. Zou, "AI-powered topic modeling: Comparing LDA and BERTopic in analyzing opioid-related cardiovascular risks in women," *Experimental Biology and Medicine*, vol. 250, 2025, doi: 10.3389/ebm.2025.10389.

[11] A. Qasim, G. Mehak, N. Hussain, A. Gelbukh, and G. Sidorov, "Detection of depression severity in social media text using transformer-based models," *Information*, vol. 16, no. 2, 2025, doi: 10.3390/info16020114.