

Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Abdullah, Zuhair, Aseem, Polly, Lavanya, Keshav, Shweta, Shad
The Web Conference (WWW) 2025



INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**



Research by



Supported by



Need to Understand Online Mental Health Support Forums

Role of Social Media and Online Platforms:

- Users often express feelings and experiences related to anxiety, depression, and other mental health conditions through posts, comments, or shared images.
- Platforms like Instagram, Reddit, and Facebook serve as rich data sources for understanding mental health patterns.

AI Techniques in Detection:

- **Textual Analysis:** Analyzing user posts to classify behaviors as depressive or non-depressive.
- **Image Analysis:** Identifying visual cues of depression in shared images (e.g., color tone, facial expressions).

Me texting my friends mental health advice when I belong in an asylum



Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Challenges:

- **Understanding Complex Memes:** Current AI models, including VLMs and LLMs, struggle with interpreting the layered meanings in memes, especially those expressing mental health symptoms like depression and anxiety.
- **Datasets:** There is a notable lack of publicly available datasets for classifying mental health memes. Only one dataset, RESTORE, exists for depressive meme classification, which limits research in this domain, especially for under-explored topics like anxiety.
- **Gap in Figurative Understanding:** Existing AI models lack the capability to combine figurative reasoning with commonsense knowledge to effectively decode mental health-related memes.



Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Mental Health Symptom Coverage: Expanding Beyond Depression

- **Current Limitation (RESTORE Dataset):**

Existing Dataset primarily focused on depressive content, with limited scope for other mental health symptoms.

- **Introducing AxiOM (Anxiety-Oriented Memes):**

A new dataset inspired by the **GAD (Generalized Anxiety Disorder) questionnaire**, specifically curated for anxiety-related content.

- **Why AxiOM?**

- **Enhanced Diversity & Coverage:**

Expands the dataset landscape by incorporating anxiety-specific memes, moving beyond depression-focused content.

- **Fine-Grained Classification:**

Enables more precise identification and differentiation of various mental health symptoms.

when you have to talk to
the teacher about retaking
a test to save your grade



Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Proposed Dataset Statistics: Axiom vs. RESTORE

(a) Depression: RESTORE dataset								
	LOI	FD	ED	SD	LSE	CP	SH	Total
Train	441	1555	1714	997	549	348	1259	6863
Val	13	55	35	18	29	20	27	197
Test	35	82	61	48	32	50	28	336
Total	489	1692	1810	1063	610	418	1314	7396

(b) Anxiety: AxiOM dataset							
	NV	LWC	EW	DR	RST	ID	Total
Train	373	331	322	356	405	366	2153
Val	53	47	46	51	58	52	307
Test	106	94	92	102	116	105	615
Total	532	472	460	509	579	523	3582

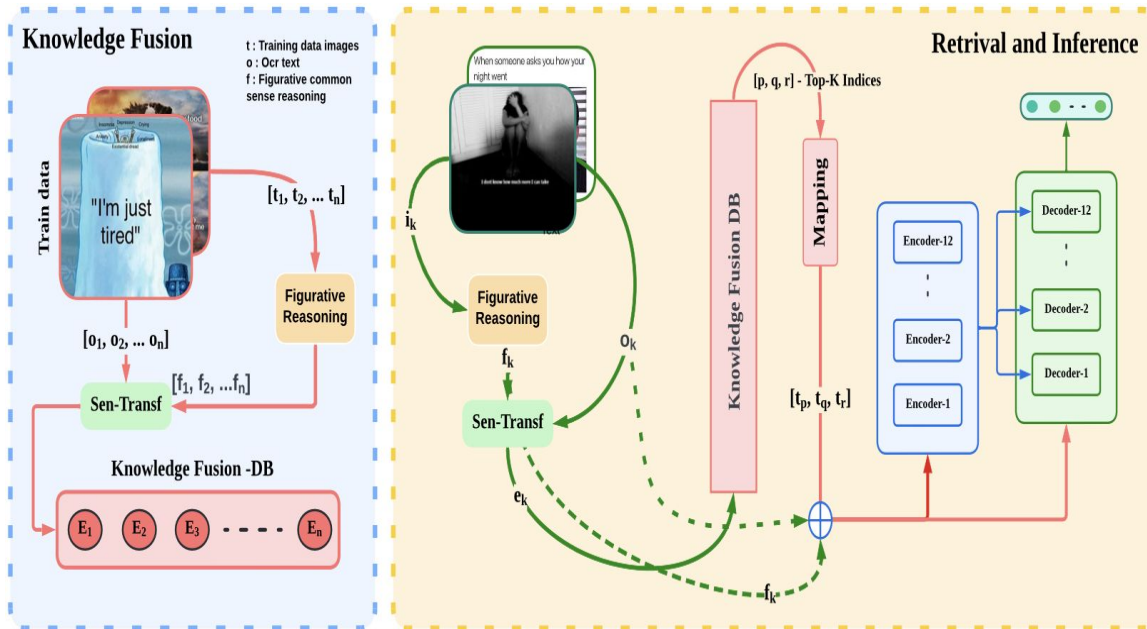
Table 1: Data statistics for the depression and anxiety datasets. (a) Depression Labels: LOI: *Lack of interest*; FD: *Feeling down*; ED: *Eating Disorder*; SD: *Sleeping Disorder*; LSE: *Low Self-Esteem*; CP: *Concentration Problem*; SH: *Self Harm* (b) Anxiety Labels: NV: *Nervousness*; LWC: *Lack of Worry Control*; EW: *Excessive Worry*; DR: *Difficulty Relaxing*; RST: *Restlessness*; ID: *Impending Doom*

Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Proposed Framework

Module 1: Figurative Reasoning

- **Key Attributes:**
 - **Cause-Effect:** Identifying real-world relationships (e.g., anxiety causing stress).
 - **Figurative Understanding:** Decoding metaphors, irony, and humor.
 - **Mental State:** Recognizing emotional or psychological states (e.g., anxiety, depression).
- **LLMs:** Use prompt engineering on large language models (LLMs) to generate enriched figurative reasoning.



Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Proposed Framework.

Module 1: Figurative Reasoning

- **Key Attributes:**
 - **Cause-Effect:** Identifying real-world relationships (e.g., anxiety causing stress).
 - **Figurative Understanding:** Decoding metaphors, irony, and humor.
 - **Mental State:** Recognizing emotional or psychological states (e.g., anxiety, depression).
- **LLMs:** Use prompt engineering on large language models (LLMs) to generate enriched figurative reasoning.

D Prompts and Examples

Analyze the following anxiety meme image to extract common sense reasoning in the form of triples.. These relationships should capture the following elements:

- 1. Cause-Effect: Identify concrete causes or results of the situation depicted in the meme.
- 2. Figurative Understanding: Capture underlying metaphors, analogies, or symbolic meanings that convey the meme's deeper message, including any ironic or humorous undertones.
- 3. Mental State: Capture specific mental or emotional states depicted in the meme.

Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

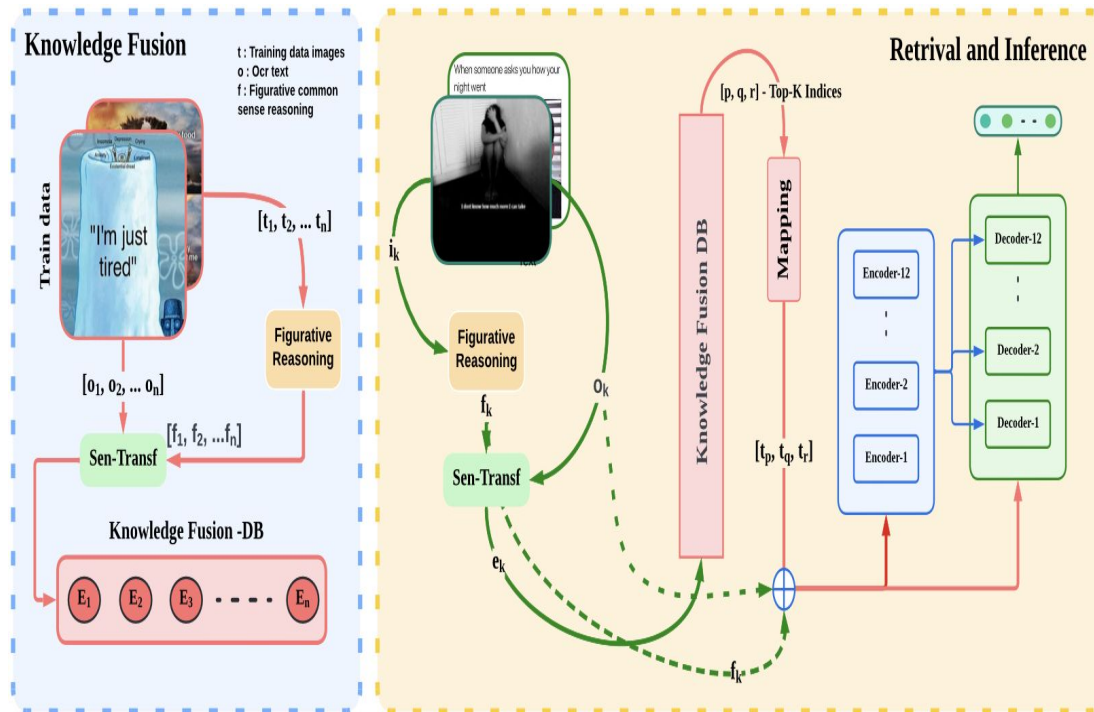
Proposed Framework.

Module 2: Knowledge Fusion (RAG Module)

- **Retrieval-Augmented Generation (RAG):** Combines OCR text and figurative reasoning with dynamic knowledge from a database to fetch relevant, up-to-date information.

Module 3: Classifier

- **Input:** Combined inputs (OCR-text, figurative reasoning, relevant knowledge attributes).
- **Approach:** Encoder-decoder architecture leveraging few-shot learning for improved generalization.



Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Proposed Framework.

Results: Our proposed model, M3H, shows substantial improvements across 20 variations of baselines which highlights the M3H's Superiority in classifying Symptoms.

Models	Depressive Memes		Anxiety Memes	
	Macro-F1	Weigh-F1	Macro-F1	Weigh-F1
OCR-Text:				
BERT	58.16	57.12	62.99	63.06
Deberta-V3	57.99	57.40	62.03	61.86
Mental-BERT	59.14	60.03	63.41	63.40
BART	41.56	38.35	62.82	62.62
Mental-BART	54.19	52.55	64.97	64.88
OCR + COMET (f_k):				
BERT	58.71	58.14	61.00	61.00
Deberta-V3	58.81	58.90	60.52	60.17
Mental-BERT	60.11	60.79	63.00	63.00
BART	62.55	63.28	62.71	62.61
Mental-BART	62.94	63.43	63.97	63.86
OCR + LLAVA (f_k):				
BERT	63.97	66.15	62.43	62.38
Deberta-V3	60.80	60.55	60.70	60.80
Mental-BERT	65.45	63.02	63.00	63.00
BART	67.00	68.55	64.01	64.02
Mental-BART	65.06	66.00	65.03	65.07
OCR + GPT-4o (f_k):				
BERT	58.65	59.05	62.01	62.02
Deberta-V3	59.87	59.61	61.33	62.40
Mental-BERT	60.80	60.29	62.38	62.52
BART	63.32	64.09	64.00	64.05
Mental-BART	63.26	64.82	65.51	65.54
Yadav et al. [41] (SOTA):	63.58	64.59	64.21	65.34
M3H (Proposed):	67.52	68.79	70.00	70.10
– (FCS+RAG)	65.20	66.75	64.28	64.97
– (OCR+RAG)	66.06	66.01	62.54	62.53
– FCS	58.50	58.91	62.82	62.62
$\Delta_{M3H-SOTA}(\%)$	$\uparrow 4.94$	$\uparrow 4.20$	$\uparrow 5.79$	$\uparrow 4.66$

Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

Ablation Study

OCR	OCR-RAG	FCS	FCS-RAG	Depression			Anxiety		
				Macro-F1	Weighted-F1	Micro-F1	Macro-F1	Weighted-F1	Micro-F1
✓	✗	✗	✗	58.50	58.91	62.29	62.82	62.62	62.21
✓	✓	✗	✗	66.01	67.33	68.60	62.54	62.53	62.05
✗	✗	✓	✗	61.18	62.05	66.84	57.00	58.01	56.84
✗	✗	✓	✓	66.42	67.65	69.24	60.45	60.69	60.46
✓	✗	✓	✗	63.32	64.09	66.22	64.28	64.97	64.20
✓	✗	✓	✓	65.20	66.75	69.01	64.94	64.95	64.84
✓	✓	✓	✗	66.06	66.01	67.92	61.82	61.79	62.21
✓	✓	✓	✓	67.52	68.79	70.09	70.00	70.00	69.38

Table 4: Experimental results for various model configurations

Figurative-cum-Commonsense Knowledge Infusion for Multimodal Mental Health Meme Classification

To connect and know more about the work

Contact: abdullahm@iiitd.ac.in

Web: <https://www.linkedin.com/in/abdullah-mazhar-129259164>

Thank You

