ADScAI
Applied Data Science & Artificial Intelligence Symposium

# Commonsense-Driven Symbolic ReAct-NLI Prompting (CSR-NLI) for Causal Analysis of Mental Health Issues in Workspace Communication; Advancements in the CAMS Dataset

1st Dilanka Wickramasinghe
*Department of Computer Science & Engineering*
*University of Moratuwa*
Sri Lanka
dilanka.23@cse.mrt.ac.lk

2nd Thanuja Ambegoda
*Department of Computer Science & Engineering*
*University of Moratuwa*
Sri Lanka
thanuja@cse.mrt.ac.lk

***Keywords*—Neuro-Symbolic AI, Commonsense Reasoning, Natural Language Inference, Mental Health Discourse Analysis, Causal Inference**

## I. Introduction

Mental health issues expressed on workspace communication often involve complex causal relationships that require deeper analysis beyond sentiment detection. Traditional NLP models face challenges in **interpretability and causal inference**, limiting their ability to accurately identify stressors and triggers. To address this, **Commonsense-Driven Symbolic ReAct-NLI (CSR-NLI)** is introduced as a **neuro-symbolic prompting framework** that integrates **commonsense reasoning with Natural Language Inference (NLI)** for structured causal analysis. By leveraging **symbolic validation, iterative refinement, and neural inference**, CSR-NLI ensures that causal explanations align with real-world commonsense knowledge, enhancing the robustness of AI-driven mental health analysis.

Additionally, advancements in the **Causal and Abductive Mental State (CAMS) dataset** [1] improve causal annotations, contributing to better reasoning consistency and classification accuracy. The proposed framework was evaluated using **GPT-3.5-Turbo, GPT-4-Turbo, Allam-2-7B, LLaMA-3.3-70B, and DeepSeek-R1**, demonstrating superior causal reasoning performance compared to existing approaches. CSR-NLI effectively enhances **explainability and reliability in mental health AI applications**, offering a scalable solution for analyzing stressors and triggers in online discourse.

## II. Literature Review

Neuro-Symbolic AI (NSAI) has emerged as a promising approach for enhancing interpretability in mental health detection by integrating deep learning with symbolic reasoning. TAM-SENTICNET utilizes a symbolic layer for sentiment analysis, improving depression detection from social media posts [2]. While Natural Language Inference (NLI) benchmarks like SNLI and MultiNLI have advanced textual entailment modeling [3], [4], domain-specific datasets such as SCITAIL and e-SNLI improve inference robustness in specialized contexts [5], [6]. However, current NLI models struggle with causal reasoning, highlighting the need for more structured inference frameworks such as Graph-Constrained Reasoning (GCR) and Commonsense NLI (ATOMIC, COMET) [7], [8]. These advancements have improved knowledge representation but remain computationally expensive and sensitive to dataset biases.

Commonsense reasoning is vital for AI-driven causal analysis, allowing models to infer implicit cause-effect relationships. Generated Knowledge Prompting (GKP) dynamically retrieves relevant information, outperforming static knowledge bases in inference tasks [9]. Further, Chain-of-Thought (CoT) and Tree-of-Thought (ToT) prompting strategies enhance logical reasoning by structuring multi-step inferences [10]. Despite these improvements, prompting techniques remain highly sensitive to input variations, impacting reasoning consistency. Addressing these limitations, the proposed *CSR-NLI* framework integrates symbolic reasoning with neural inference, enhancing causal reasoning and explainability in workplace mental health assessments.

## III. Methodology: CSR-NLI Framework

The *CSR-NLI* framework integrates commonsense reasoning with symbolic validation to enhance causal inference in mental health discourse. It combines Natural Language Inference (NLI) with structured prompting to ensure interpretability and logical consistency in causal classification (Figure 1).

https://doi.org/10.31705/ADScAI.2025.43

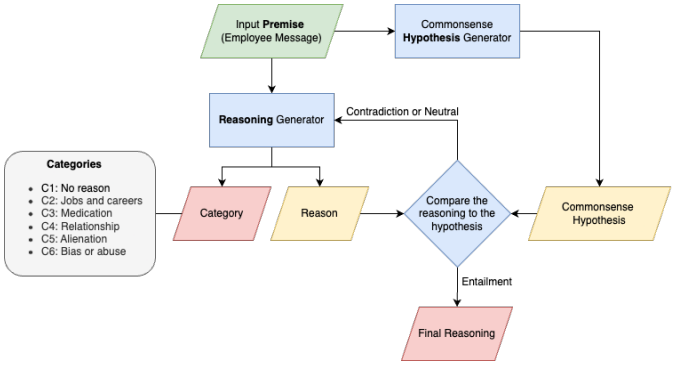


Fig. 1: Commonsense-Driven Symbolic ReAct-NLI Framework

CSR-NLI begins with an input **Premise**, representing an employee message extracted from workspace communication. A commonsense **Hypothesis** is generated using LLM, GPT-3.5 Turbo selected for its cost-effectiveness and strong reasoning capability. The model then generates a **Reasoning Statement** and assigns it to one of six predefined **Causal Category** (*No Reason, Jobs and Careers, Medication, Relationships, Alienation, or Bias/Abuse*). These outputs are compared against the hypothesis to determine logical consistency in terms of *Entailment, Contradiction, or Neutral*. To improve alignment and consistency, the framework employs the **ReAct (Reasoning + Acting)** strategy, which iteratively refines the reasoning until it aligns with the hypothesis. The final output is a validated causal explanation and its corresponding category.

Through this process, CSR-NLI achieves robust causal inference with improved interpretability. Integrating symbolic reasoning and neural inference makes it suitable for real-world applications such as workspace health monitoring.

## IV. Results and Discussion

The CSR-NLI framework was evaluated using the **CAMS dataset** [1], which contains 5,051 Reddit posts with corresponding reasoning statements. To assess CSR-NLI's effectiveness in generating causal reasoning, both the GPT-3.5 Turbo generated reasoning and the baseline CAMS dataset reasoning were evaluated against the original input premise using **GPT-3.5-Turbo, GPT-4-Turbo, Allam-2-7B, LLaMA-3.3-70B, and DeepSeek-R1-Distill-LLaMA-70B**. These models were selected for their strength in reasoning and diversity across proprietary and open-source systems. Each model independently rated reasoning quality using a **0–100 scale**, evaluating **logical coherence, causal alignment, and interpretability**. Statistical analyses, including **paired t-tests, Wilcoxon signed-rank tests, and effect size analysis (Cohen's d)** were conducted to ensure significance and robustness.

Results demonstrate that **CSR-NLI significantly enhances causal reasoning quality** across all models, with **notable gains over the baseline CAMS reasonings** (Table I). Models like GPT-4-Turbo and LLaMA-3.3-70B showed over 50-point improvements. Statistical validation ($\mathbf{p < 0.001}$, **Cohen's d** $>$ **2.5**) confirmed both the significance and the large effect size of the improvements. LLaMA-3.3-70B yielded the lowest failure rate ($\sim$2%), when comparing GPT-3.5-Turbo–generated reasoning with CAMS baselines, indicating high consistency in recognizing reasoning improvements introduced by CSR-NLI. Visual analyses (histograms, scatter plots, and correlation heatmaps) revealed a consistent shift toward higher reasoning quality, confirming that CSR-NLI enhances AI-driven causal inference. The findings highlight CSR-NLI's potential to generate more reliable, interpretable, and actionable AI explanations in mental health analysis.

TABLE I: Summary Statistics of Reasoning Evaluation

| Statistic | GPT-3.5-Turbo | GPT-4-Turbo | Allam-2-7B | LLaMA-3.3-70B | DeepSeek-R1 |
|---|---|---|---|---|---|
| CAMS Mean | 83.37 | 35.66 | 76.05 | 38.68 | 59.18 |
| CSR-NLI Mean | 92.08 | 82.07 | 84.64 | 87.13 | 84.09 |
| CAMS Std. Dev. | 11.08 | 20.68 | 13.91 | 17.41 | 21.31 |
| CSR-NLI Std. Dev. | 5.75 | 12.48 | 9.63 | 8.17 | 10.69 |
| P-Value | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Cohen's d | 0.986 | 2.716 | 0.718 | 3.562 | 1.473 |
| Failure Rates | $\sim$13% | $\sim$6% | $\sim$11% | $\sim$2% | $\sim$13% |

## V. Conclusion & Future Work

This study demonstrates that integrating **symbolic reasoning with neural models** enhances interpretability and logical coherence in causal analysis of workspace mental health discourse. Unlike static knowledge-based approaches such as ATOMIC and COMET [7], [8], the CSR-NLI dynamically generates and refines hypotheses, making it more adaptable to workplace stressors. Compared to pattern-based models like logistic regression and CNN-LSTM, CSR-NLI ensures robust causal inference through iterative hypothesis validation. These advances support early stress detection, personalized employee support, and HR-driven wellness initiatives.

Despite its strengths, CSR-NLI introduces computational overhead due to its iterative refinement process, which may impact scalability. As the CAMS dataset is derived from Reddit posts, it may carry platform-specific cultural and demographic biases, limiting generalizability to broader workplace settings. Ethical concerns related to model transparency and privacy also warrant attention.

Future work should focus to optimize inference speed for real-time use, enhance scalability, diversify datasets to verify generalization of the framework, and explore applications in domains such as healthcare and education. Moreover, integrating privacy-aware mechanisms and ethical prompting strategies will be vital for responsible and adaptable deployment.

## References

[1] M. Garg et al., "CAMS: An Annotated Corpus for Causal Analysis of Mental Health Issues in Social Media Posts," in *Proc. LREC*, 2022.
[2] S. Tam et al., "TAM-SENTICNET: A Neuro-Symbolic Approach to Depression Detection," in *Proc. AAAI Conf. Artif. Intell.*, 2022.
[3] S. Bowman et al., "A Large Annotated Corpus for Learning Natural Language Inference," in *Proc. EMNLP*, 2015.
[4] A. Williams et al., "A Broad-Coverage Challenge Corpus for Natural Language Inference," in *Proc. ACL*, 2018.
[5] T. Khot et al., "SciTail: A Textual Entailment Dataset from Science Question Answering," in *Proc. AAAI*, 2018.
[6] O. Camburu et al., "e-SNLI: Natural Language Inference with Human Explanations," in *NeurIPS*, 2018.
[7] M. Sap et al., "ATOMIC: An Atlas of Machine Commonsense for Everyday Inference," in *Proc. AAAI*, 2019.
[8] A. Bosselut et al., "COMET: Commonsense Transformers for Automatic Knowledge Graph Completion," in *Proc. ACL*, 2019.
[9] H. Wu et al., "Generated Knowledge Prompting for Commonsense AI," *J. AI Res.*, 2023.
[10] J. Wei et al., "Chain of Thought Prompting Elicits Reasoning in Large Language Models," in *Proc. NeurIPS*, 2022.