Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Targeted Knowledge Infusion To Make Conversational AI Explainable and Safe

Manas Gaur

Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County
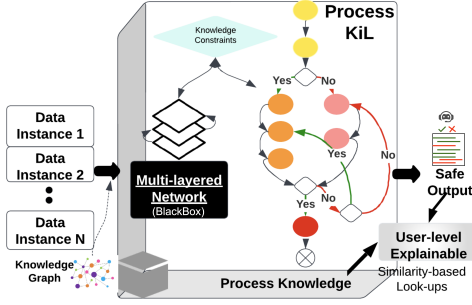1000 Hilltop Cir, Baltimore, MD, USA 21250
manas@umbc.edu

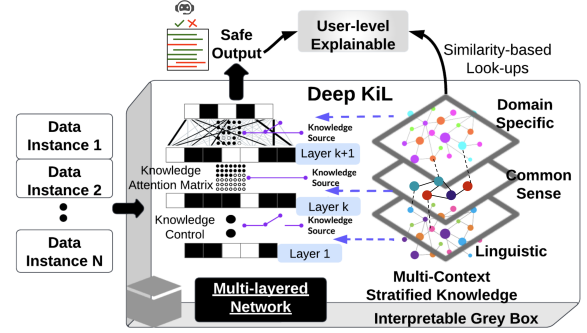Figure 1: Process Knowledge-infused Learning

Figure 2: Deep Knowledge-infused Learning

New Faculty Talk Description

Conversational Systems (CSys) represent practical and tangible outcomes of advances in NLP and AI. CSys see continuous improvements through unsupervised training of large language models (LLMs) on a humongous amount of generic training data. However, when these CSys are suggested for use in domains like Mental Health, they fail to match the acceptable standards of clinical care, such as the clinical process in Patient Health Questionnaire (PHQ-9).

The talk will present, Knowledge-infused Learning (**KiL**), a paradigm within NeuroSymbolic AI that focuses on making machine/deep learning models (i) learn over knowledge-enriched data, (ii) learn to follow guidelines in process-oriented tasks for safe and reasonable generation, and (iii) learn to leverage multiple contexts and stratified knowledge to yield user-level explanations (Gaur et al. 2022a). **KiL** established Knowledge-Intensive Language Understanding, a set of tasks for assessing safety, explainability, and conceptual flow in CSys (Sheth et al. 2021).

KiL is a continuum of three stages: The **first** stage, *Shallow Infusion*, aims to make the *dataset explainable* through semantic annotation and guarantee that the model has drawn on reliable, knowledge-guided data standards. The **second** stage is Process Knowledge-infused Learning (**PKiL**) which enable CSys to generate sentences following domain-specific expert guidelines, such as clinical questionnaires (Figure 1 illustrates the flow in a questionnaire). For instance, if a user says to a CSys: *For the past **several days**, I*

*am facing **trouble in concentrating** while **reading newspaper** or **watching television***, a PKiL-trained agent would generate the following series of questions by following a clinically approved process of diagnosis: (1) Do you feel like you sleep a lot but are still tired? (*Symptoms*) →(2) How many hours of sleep do you get on average each night? (*Cause*) → (3) Are you taking any prescribed medications (*Medications*)? → (4) Have you been diagnosed with any sleep disorder? (*Diagnosis*) (Roy et al. 2023)(Gaur et al. 2022b). Figure 2 is the **third** stage, *Deep Knowledge Infusion*, which supports the stronger and controlled weaving between multiple knowledge sources and different layers in LLMs in CSys.

References

Gaur, M.; Gunaratna, K.; Bhatt, S.; and Sheth, A. 2022a. Knowledge-Infused Learning: A Sweet Spot in Neuro-Symbolic AI. *IEEE Internet Computing*.

Gaur, M.; Gunaratna, K.; Srinivasan, V.; and Jin, H. 2022b. Iseeq: Information seeking question generation using dynamic meta-information retrieval and knowledge graphs. In *Proceedings of the AAAI Conference*.

Roy, K.; Gaur, M.; Rawte, V.; Kalyan, A.; and Sheth, A. 2023. ProKnow: Process Knowledge for Safety Constrained and Explainable Question Generation for Mental Health Diagnostic Assistance. *Frontiers in Big Data*.

Sheth, A.; Gaur, M.; Roy, K.; and Faldu, K. 2021. Knowledge-intensive language understanding for explainable AI. *IEEE Internet Computing*.