ACM web conference 2023

Usha Lokala
Kaushik Roy
Utkarshani Jaimini
Amit Sheth

Follow this and additional works at: https://scholarcommons.sc.edu/aii_fac_pub

Part of the Computer Engineering Commons, and the Electrical and Computer Engineering Commons
Causal AI for web and health care

Usha Lokala\textsuperscript{1} \hspace{1cm} Kaushik Roy\textsuperscript{1} \hspace{1cm} Utkarshani Jaimini\textsuperscript{1}

nlokala@email.sc.edu \hspace{1cm} kaushikr@email.sc.edu \hspace{1cm} ujaimini@email.sc.edu

Amit Sheth\textsuperscript{1}

amit@sc.edu

\textsuperscript{1} Artificial Intelligence Institute South Carolina

Abstract

Improving the performance and explanations of ML algorithms is a priority for adoption by humans in the real world. In critical domains such as healthcare, such technology has significant potential to reduce the burden on humans and considerably reduce manual assessments by providing quality assistance at scale. In today’s data-driven world, artificial intelligence (AI) systems are still experiencing issues with bias, explainability, and human-like reasoning and interpretability. Causal AI is the technique that can reason and make human-like choices making it possible to go beyond narrow Machine learning-based techniques and can be integrated into human decision-making. It also offers intrinsic explainability, new domain adaptability, and bias-free predictions that work with datasets of all sizes. In this tutorial of type lecture style, we detail how a richer representation of causality in AI systems using a knowledge graph (KG) based approach is needed for intervention and counterfactual reasoning (Figure 1), how do we get to model-based and domain explainability, how causal representations helps in web and health care.

1 Tutorial Schedule and Activities

This tutorial has the following four major modules as follows:

(10 mins) The current landscape of AI: We begin the session motivating the need for causal AI in healthcare and when analyzing the web based data. We’ll give instances from actual web examples that highlight the limitations of the available statistical and data-powered AI Systems. The opening session will emphasize the need for causal AI for creating comprehensible explainable AI systems.

(25 mins) Causal Knowledge Graph Development: CausalKG is a step toward symbolic AI for knowledge-infused learning that is causality-based. We demonstrate that systems using CausalKG don’t just learn from correlations; instead, have a causal understanding of the environment around them [3]. In addition to causality, CausalKG uses KG which also represents space, time, and interactions. The AI algorithms used today rely on uniformly distributed, independent data which are unable to deduce hypothetical, interventional scenarios or out-of-distribution situations. We show how to enable interventional and counterfactual reasoning that may be derived utilizing observational data and domain expert knowledge and how CausalKG can be utilized to infuse current KGs with causal knowledge of the domain. We discuss on building a CausalKG which has the benefit of incorporating causality into reasoning and prediction processes, such as the understanding, planning, and diagnosing of medical conditions for example.

(25 mins) Ontology and Knowledge based inference for causal Explanation: We construct an inference system to record explanations based on causal assertions using an ontology in the form of relationships and sub class hierarchy. We begin by ontology construction that enables us to express how
one fact leads to another and how one fact explains another. We provide a formal set of pattern extractions, causal relationships that move from causal claims to explanations. These patterns exhibit the ontological principles considered necessary for reaching judgments on explanatory statements and algorithm predictions. We also present Drug abuse Ontology as an ontological framework that recognizes the patterns in web based data and in the mental healthcare domain.

(25 mins) Applications in web and healthcare: The COVID19 Pandemic has brought attention to the disparity in access to mental health services. Internet users have turned to online forums like Reddit to share their experiences. There are clinically accepted causal interpretations and semantics for diagnosing mental illnesses. In this session, we propose a causal process reasoning infused machine learning techniques called Causal Process Knowledge Infused Reasoning (CPR). CPR uses mental health causal models to guide the learning of neuro-symbolic approximation functions (reasoners) to capture causal model conditional probabilities. We assess the efficacy of CPR using task descriptions, datasets, and baseline techniques for suicidality context identification and intervention. We demonstrate a system that deploys CPR on Reddit’s social media (SocMedia) posts showing CPR’s ability to scale to web data. We further discuss possible improvements to causal theories that may enhance performance on these tasks and have broader mental health care implications.

2 Topic and Relevance

SocMedia emerged as the go-to platform for mental health (MH) support for patients seeking Mental Health Care (MHCare). The transition of patients from clinical settings to SocMedia has been facilitated by peer-support groups and the lack of social stigma. For example, in order to find clues that can demonstrate a correlation or cause between various MH illnesses and a patient’s propensity for suicide, researchers have begun examining SocMedia content. In that case, this tutorial is relevant to develop causal AI systems for a deeper understanding of SocMedia conversations that may result in policy decisions due to the abundance of SocMedia conversations in the areas of health care, mental health, substance use, etc. that are available on the internet whereas surface level deep learning systems cannot achieve such goals.

3 Tutorial Audience and Prerequisites

Researchers from academia, business, and healthcare professionals can participate in this tutorial to discuss the intersection of causal representation, reasoning, semantic linking, NLP, and deep learning. Since the tutorial is taught in a lecture-style format, a fundamental knowledge of causal systems, Knowledge Graphs (KG), Ontology, and natural language processing (NLP) is preferred. This will make it possible for the public to understand the limitations of statistical AI and to monitor the development of AI as it moves toward neuro-symbolic AI. The lesson will include enough examples and use cases of causal AI systems and techniques. The present state of AI/ML systems for web-based analysis of health care will be explained to newcomers interested in these systems. The methodology and datasets provided in the tutorial will be valued by experienced participants as viable solutions to common technical challenges in the social good domains. We welcome additional studies using our datasets and methodology from other researchers in order to deepen our understanding of causal AI systems for the web and health care.

4 Tutorial Societal Impacts

Through the development of domain-specific data sets, cutting-edge frameworks, and computational methods for comprehending user language and asynchronous conversations on a variety of platforms, we investigated a new healthcare dimension of SocMedia, like Mental Health Care and User interactions. Through the use of these methods, we hope to (a) build new
clinical process guidelines for patients and develop task-based Order Sets, (b) create actionable Order Sets that are integrated into the system and aid medical professionals in making choices and avoiding delays and discrepancies in diagnosis and treatment. (c) create process-guided explanation models that are simple for the end-user to comprehend (d) create a causal AI system using a knowledge-graph-based methodology for better explainability, support for intervention, and counterfactuals in social good domains.

5 Presenter’s biographies and Related Papers

Usha Lokala, Kaushik Roy, and Utkarshani Jaimini are Ph.D. students at the Artificial Intelligence Institute South Carolina (AIISC). Amit Sheth is the director at AIISC. The authors have published several related papers at premier conferences such as AAAI [1, 2, 3, 4, 5, 6], IEEE [7, 8, 9]. ECML-PKDD [10], Frontiers in Big Data [11], ACL [12, 13], KGC [14], NeurIPS [15], and ICLR [16]. The authors have also given several related talks and tutorials [17, 18, 19, 20, 21, 22].

6 Expected background and prerequisite of audience

The tutorial would be a mix of lecture-style and hands-on in the python programming language. The audience is expected to have a basic understanding of deep/machine learning, natural language processing, and semantic technologies (e.g., linked open data). We aim to guide attendees through a high-level tour of the most recent approaches proposed by researchers. Also, we expect basic familiarity with social media platforms such as Twitter and Reddit. We expect participants to bring their laptops with all the required tools installed. Details on tools needed and the background material will be provided upon acceptance of the tutorial proposal. We expect that by the end of the tutorial, the attendees will understand the use of knowledge graphs to enhance the performance (quality of results), utility, interpretability, and explainability of deep learning and be prepared to apply knowledge-infused deep learning to real-world applications.

Acknowledgements

This work was supported in part by the National Science Foundation under Grant 2133842, “EAGER: Advancing Neuro-symbolic AI with Deep Knowledge-infused Learning” [23, 24, 25]

References


