Domain Adaptative Causality Encoder

Farhad Moghimifar, Gholamreza Haffari, Mahsa Baktashmotlagh

The University of Queensland, Australia — Create Change

Monash University

Introduction

Causality is the basis for reasoning and decision making. While human-beings use this psychological tool to choreograph their environment into a mental model to act accordingly, the inability to identify causal relationships is one of the drawbacks of current Artificial Intelligence systems. However, projection of causal relations in natural language enables machines to develop a better understanding of the surrounding context and helps downstream tasks such as question answering, text summarisation, and natural language inference.

Takeaways

- We propose a model which takes into account both syntactic and semantic dependency of words in a sentence.
- We propose to use a gradient reversal approach to minimise the distribution shift between the training and test datasets.
- To fill the gap of the current causality datasets on encompassing different types of causality, we introduce MedCaus, a dataset of 15,000 labelled sentences, retrieved from medical articles.

Graphical Causality Encoder (GCE)

Given $\boldsymbol{h}_i^{(l-1)}$ as the representation of the node i at layer $l-1$, GCN updates the node representation at layer l as follows:

$$\boldsymbol{h}_i^{(l)} = \boldsymbol{f}^{\text{actv}}\Big(\sum_{j=1}^{n} \tilde{A}_{ij}\mathbf{W}^{(l)}\boldsymbol{h}_j^{(l-1)}/d_i + \boldsymbol{b}^{(l)}\Big) \tag{1}$$

After applying the BiLSTM and GCN, each sentence is represented as:

$$\boldsymbol{h}_{\text{GCN}}(\boldsymbol{X}) = \boldsymbol{f}^{\text{pool}}(\text{GCN}(\text{BiLSTM}(\boldsymbol{X}))) \tag{2}$$

$$\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X}) = \text{FFNN}([\boldsymbol{h}_{\text{GCN}}(\boldsymbol{X}); \boldsymbol{h}_{\text{BiLSTM}}(\boldsymbol{X})]) \tag{3}$$

For Task1, Causality Identification, we use this representation to get the probability of output classes,

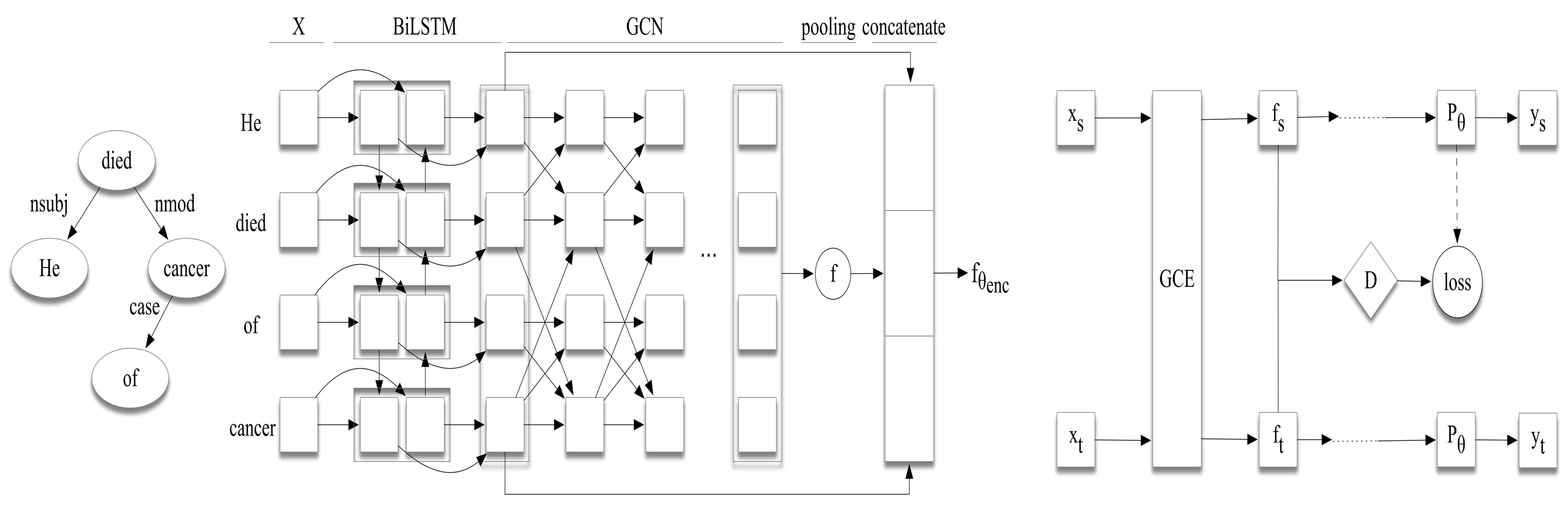
$$P_{\boldsymbol{\theta}_{\text{class}}}(\text{causality}|\boldsymbol{X}) = \sigma(\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X}).\mathbf{W}_{\text{class}} + b_{\text{class}}) \tag{4}$$

For Task2, Causality Localisation, the tags with the respect to each token is identified by:

$$\boldsymbol{s}(\boldsymbol{X}, \boldsymbol{Y}) = \sum_{i=0}^{n} \boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X})_{i,\boldsymbol{y}_i} + \sum_{i=0}^{n} T_{\boldsymbol{y}_i,\boldsymbol{y}_{i+1}} \tag{5}$$

$$P_{\boldsymbol{\theta}_{\text{seq}}}(\boldsymbol{Y}|\boldsymbol{X}) = \boldsymbol{s}(\boldsymbol{X}, \boldsymbol{Y}) - log\Big(\sum_{\boldsymbol{y}\in\boldsymbol{Y}_{\boldsymbol{X}}} e^{\boldsymbol{s}(\boldsymbol{X},\boldsymbol{y})}\Big) \tag{6}$$

Adaptive Causality Encoder (ACE)



In adaptive case, let us consider the following domain classifier,

$$P_{\boldsymbol{\theta}_{\text{dom}}}(\text{source}|\boldsymbol{X}) = \sigma(\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X}).\mathbf{W}_{\text{dom}} + b_{\text{dom}}) \tag{7}$$

Our domain adversarial training objective is defined as,

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}_{\text{enc}}, \boldsymbol{\theta}_{\text{dom}}, \boldsymbol{\theta}_{\text{task}}) := &\sum_{(\boldsymbol{X},Y)\in D_s} \log P_{\boldsymbol{\theta}_{\text{class}}}(Y|\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X})) \\ &- \sum_{\boldsymbol{X}\in D_s} \log P_{\boldsymbol{\theta}_{\text{dom}}}(\text{source}|\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X})) \\ &- \sum_{\boldsymbol{X}\in D_t} \log\big(1 - P_{\boldsymbol{\theta}_{\text{dom}}}(\text{source}|\boldsymbol{f}_{\boldsymbol{\theta}_{\text{enc}}}(\boldsymbol{X}))\big).\end{aligned} \tag{8}$$

Experimental Results on GCE

	MedCaus			FinCausal		
Model	P	R	F1	P	R	F1
P-Wiki	74.4	74.4	74.4	54.0	54.0	54.0
bi-LSTM	84.2	**97.8**	90.5	81.3	77.0	79.1
GCN	90.8	94.6	92.7	85	74.8	79.6
C-GCN	91.2	94.9	93.0	**86.1**	68.7	76.4
GCE	**92.5**	94.0	**93.2**	84.8	**83.3**	**84**

The performance of our model on Causality Identification.

	MedCaus			FinCausal		
Model	P	R	F1	P	R	F1
bi-LSTM-CRF	**77.4**	69.9	73.4	**82.4**	65.0	72.7
GCN-CRF	31.9	46.8	37.9	66.1	55.5	60.3
C-GCN-CRF	72.5	**75.9**	74.1	76.3	68.8	72.3
S-LSTM-CRF	58.6	64.0	61.2	61.5	29.7	40.0
ELMO-CRF	48.5	78.9	60.1	71.8	61.3	66.1
GCE	76.3	73.6	**74.9**	79.2	**69.8**	**74.2**

The performance of our model on Causality Localisation.

Experimental Results

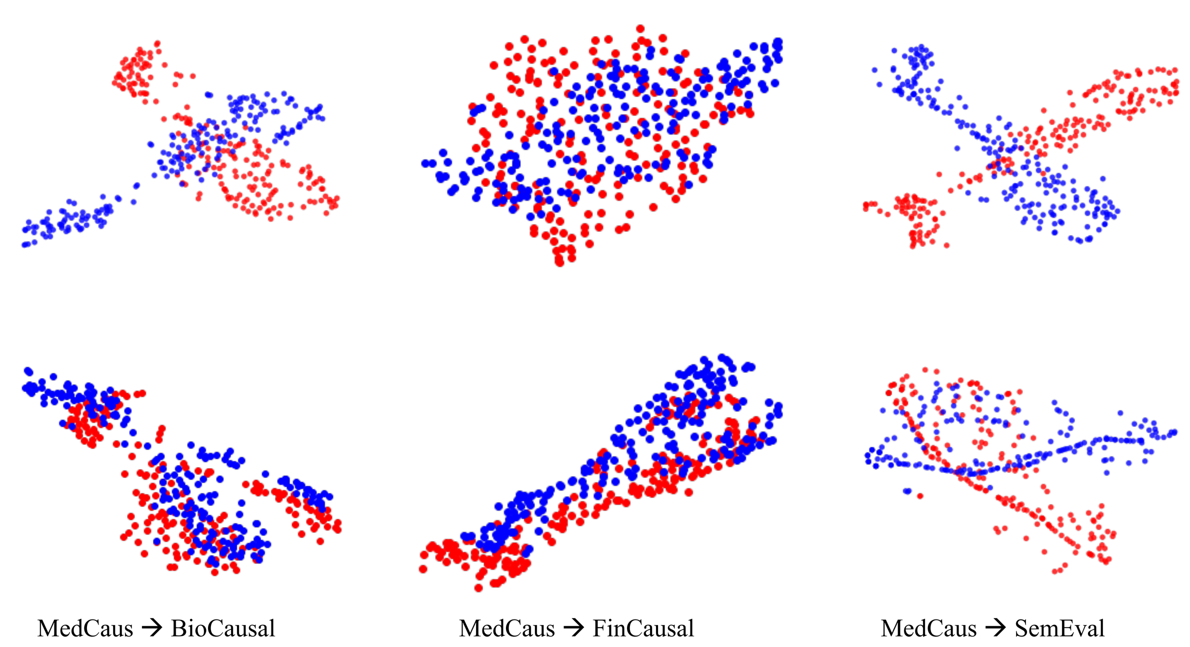
	MedCaus → BioCausal			MedCaus → SemEval			MedCaus → FinCausal		
Models	P	R	F1	P	R	F1	P	R	F1
bi-LSTM	76.1	57.6	66.0	82.5	62.8	71.3	47.6	8.3	14.2
GCN	75.4	51.0	60.8	78.4	67.6	72.6	49.2	53.2	51.1
C-GCN	71.3	42.9	55.3	84.1	70.8	76.9	48.7	52.3	50.4
bi-LSTM+DA	75.6	58.9	66.2	81.6	69.5	75.1	47.9	61.1	53.7
GCN+DA	72.8	70.3	71.5	82.9	66.7	73.9	46.8	57.4	51.6
C-GCN+DA	78.4	55.5	65.1	81.9	71.0	76.1	**49.1**	54.6	52.7
CDAN	**85.5**	50.1	63.8	**84.6**	73.8	78.8	43.6	53.3	48.0
CDAN-E	**83.8**	55.0	66.4	81.2	74.2	77.6	48.3	63.3	54.8
ACE	74.3	**77.1**	**76.7**	**84.4**	**74.2**	**79.0**	47.4	**74.0**	**57.8**

The performance of our model on Causality Identification.

	MedCaus → SemEval			MedCaus → FinCausal		
Models	P	R	F1	P	R	F1
bi-LSTM	16.3	52.2	24.9	**64.1**	16.1	25.8
GCN	8.8	29.6	13.5	41.6	40.9	41.0
C-GCN	18.9	47.6	27.1	63.8	13.0	21.6
bi-LSTM+DA	**51.2**	42.0	46.2	44.9	39.4	41.9
GCN+DA	9.1	25.5	13.4	39.6	45.2	42.2
C-GCN+DA	45.1	42.1	43.6	40.0	**45.5**	42.6
CDAN	40.9	49.7	44.8	36.9	42.8	39.6
CDAN-E	47.3	40.6	43.7	36.8	42.6	39.5
ACE	42.3	**53.6**	**47.3**	42.2	43.2	**42.7**

The performance of our model on Causality Localisation.

Data Visualisation



t-SNE visualisation of the domain adaptation task.