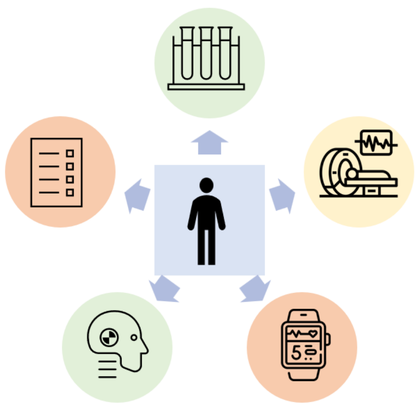
Credibility-aware Multi-Modal Fusion Using Probabilistic Circuits

Sahil Sidheekh*, Pranuthi Tenali*, Saurabh Mathur*, Erik Blasch, Kristian Kersting, Sriraam Natarajan

What?

Motivation

Real world decision-making often requires reasoning about information from multiple heterogeneous sources.



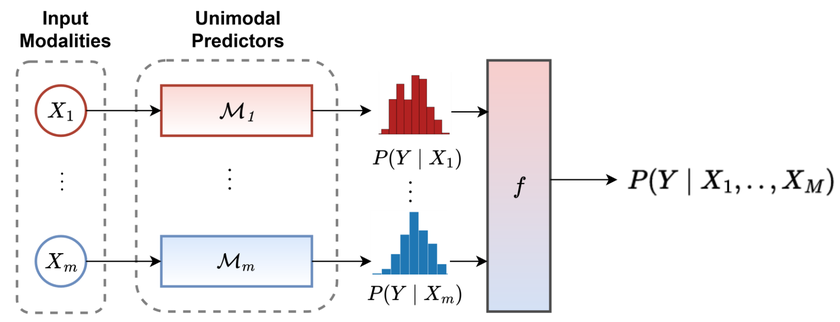
e.g., Medicine

What if sources disagree?

Need a principled way to infer the credibility of information from each source.

Background

Late/decision fusion models make predictions from multi-modal data by combining the predictions of modality-specific models via combination functions.



Probabilistic circuits (PCs) represent the joint distribution over a set of variables using a structured computational graph; the structure enables efficient, exact, and differentiable probabilistic inference.

How?

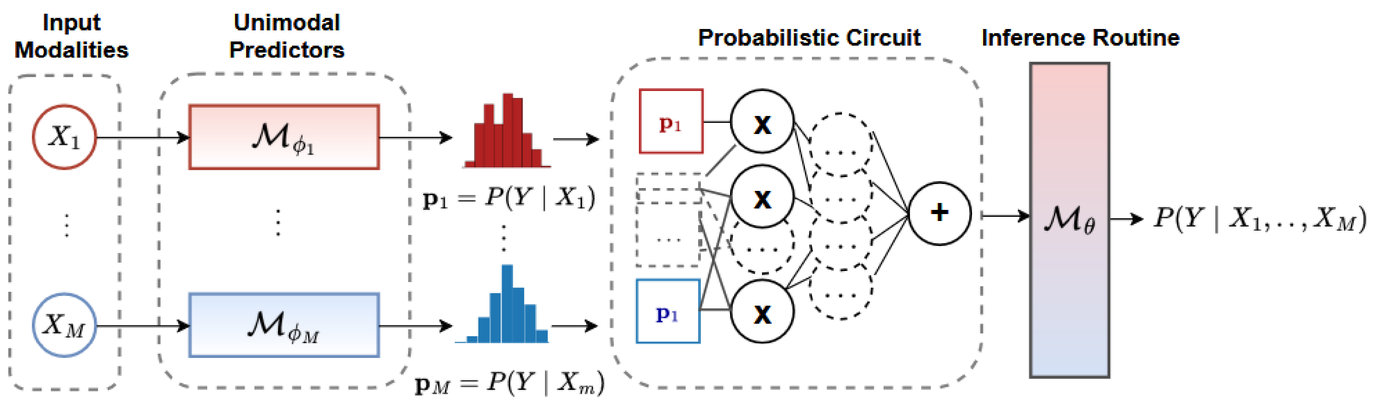
Multimodal-fusion via PCs

Credibility of source $X \in \boldsymbol{X}$

Extent of change in belief about target by observing information from X.

$$C = d_{KL}\big(P(Y \mid \boldsymbol{X}) \,||\, P(Y \mid \boldsymbol{X} \setminus \{X\})\big)$$

Probabilistic Circuits as Combination Functions



Inference. Given multimodal data point $\boldsymbol{x}$, infer distribution over target Y, accounting for source-specific credibility.

Unimodal predictions, $\mathbf{p} = \{\mathrm{p_i} = \phi(x_i)\}_{i=1}^{M}$

Direct-PC (DPC) combination
Combine by directly inferring conditional $P(Y \mid \mathbf{p})$ from PC

Credibility-weighted mean (CWM)
Combine via weighted mean; where each source's weight is its credibility.

Empirical Evaluation

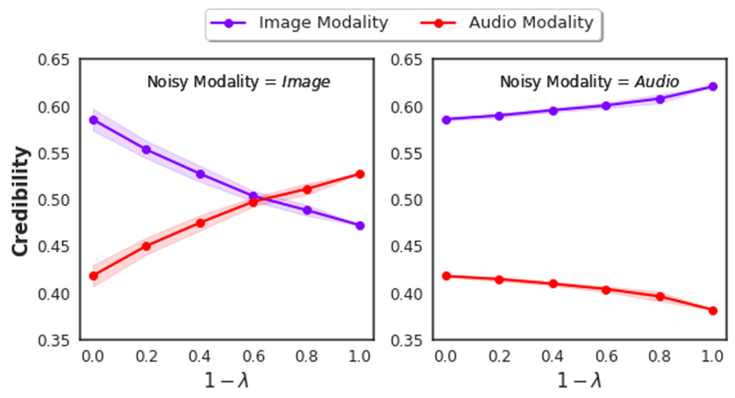
PC-based fusion performs at par with state-of-the art.

	Combination Function	Test Performance				
		Accuracy	Precision	Recall	F1 score	AUROC
Deep	MLP	90.0 ± 1.4	90.4 ± 1.3	89.7 ± 1.4	89.5 ± 1.4	99.4 ± 0.3
Statistical	Weighted Mean	91.3 ± 2.2	91.9 ± 1.7	91.3 ± 2.2	91.4 ± 2.1	99.4 ± 0.3
	Noisy or	90.9 ± 2.7	91.4 ± 2.4	90.8 ± 2.6	90.9 ± 2.6	99.4 ± 0.3
	QMF	90.5 ± 2.4	90.99 ± 2.4	90.5 ± 2.4	90.4 ± 2.4	**99.53 ± 0.4**
Noise-aware	TMC	91.5 ± 3.2	92.1 ± 3.0	91.5 ± 3.2	91.5 ± 3.1	99.4 ± 0.3
	RCML	89.3 ± 5.0	90.0 ± 4.9	89.3 ± 5.01	89.0 ± 5.2	99.34 ± 0.3
	CWM	**92.5 ± 1.4**	**94.0 ± 1.5**	**92.5 ± 1.4**	**92.5 ± 1.0**	99.42 ± 0.2
	Direct-PC	91.7 ± 1.0	92.4 ± 1.1	91.7 ± 1.0	91.6 ± 0.9	99.2 ± 0.4

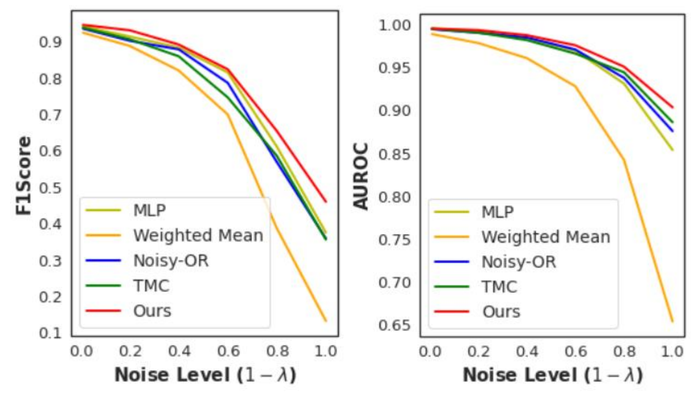
Mean test performance on the CUB dataset across 5 trials

CWM assumes a linear relationships between credibility and the target. So, CWM is better suited to smaller data sets where such linearity holds; while DPC is better for larger data sets where linearity cannot be assumed.

Increasing noise in a modality leads to reduced credibility (AV-MNIST)



Credibility-aware fusion is robust to noise (CUB).



What we get?

Future Work

- Extending to infer credibility across knowledge-graphs.
- Extending to include context-specific credibility.
- Extending to intermediate fusion.

Acknowledgements

We acknowledge support from AFOSR award FA9550-23-1-0239, ARO award W911NF2010224 and DARPA Assured Neuro Symbolic Learning and Reasoning (ANSR) award HR001122S0039.

https://starling.utdallas.edu/

@STARLing_lab

*Equal contributors

Blasch et al. 2013, 2014; Choi et al.2020; Poon and Domingos 2011; Das et al. 2023; Natarajan et al. 2005, 2018; Sidheekh, Kersting, and Natarajan 2023; Tian et al. 2020; Peharz et al. 2020a,b; Khaleghi et al. 2013