Università degli Studi di Bari
Dipartimento di Informatica

LACAM Laboratory
Machine Learning

Towards Representation Learning with Tractable Probabilistic Models

Antonio Vergari, Nicola Di Mauro and Floriana Esposito

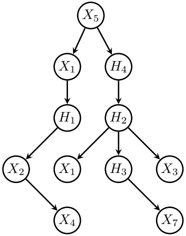
*19th September - **ECML-PKDD** - Riva del Garda, Italy*

Tractable Probabilistic Models (TPMs)

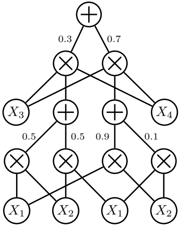
Plenty of Probabilistic Models learned as *density estimators*

Many Machine Learning problems can be reframed as probabilitic inference

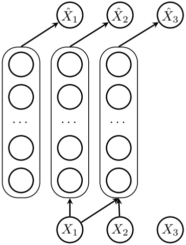
...but ***inference is hard***.



Low treewidth PGMs



Computational Graphs



Autoregressive NNs

$\longrightarrow$ TPMs allow **exact** inference to be computed in **polynomial time**!

Representation learning with TPMs

Given a set of i.i.d samples $\{\mathbf{x}^i\}_{i=1}^m \sim \mathbf{X}$, a TPM θ, we want to generate an embedding $\mathbf{e}^i \in \mathbb{R}^d$ for each sample i such as:

$$\mathbf{e}^i = f_{p,\theta}(\mathbf{x}^i)$$

with f being the transformation by θ encoding the distribution $p(\mathbf{X})$.

Idea: evaluate θ ***several*** times by constructing ***random queries*** (e.g. sample $\mathbf{Q}_j \subseteq \mathbf{X}, j = 1 \ldots, d$), then use the probability value of each query as an embedding component:

$$e_j^i = p_\theta(\mathbf{Q}_j = \mathbf{x}^i_{\mathbf{Q}_j})$$

$\rightarrow$ reuse previously learned models, as *black boxes*
$\rightarrow$ exploit embeddings for clustering, classification,...

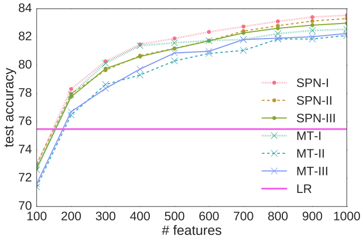
Experimental evaluation

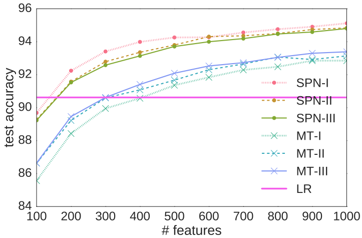
I. Learning SPNs and MTs unsupervisedly on five binary image datasets

II. extract embeddings from 1000 random marginal queries

III. train a supervised linear classifier on them

$\rightarrow$ ***meaningful representations*** *if better accuracy scores*





$\rightarrow$ trade-off between likelihood over $\mathbf{X}$ and accuracy over Y