

Aprendizado Automático de Sum-Product Networks (SPN)

Projeto de MAC0215 (Atividade Curricular em Pesquisa)
Aluno: Renato Lui Geh (Bacharelado em Ciência da Computação)
Orientador: Denis Deratani Mauá

1 PROPOSTA

O objetivo deste projeto é realizar uma comparação das principais técnicas de aprendizado automático de Sum-Product Network a partir de um conjunto de dados, avaliando o desempenho e os resultados das técnicas utilizadas.

O aluno irá estudar a estrutura e propriedades de uma Sum-Product Network e em seguida aplicar as seguintes técnicas de aprendizado:

- Busca gulosa.[1]
- Aprendizado com clustering de variáveis.[2]
- Aprendizado por meio de SPNs Bayesianas Não-Paramétricas.[5]

Em seguida o aluno irá comparar as técnicas implementadas e apresentar os resultados.

2 ACOMPANHAMENTO

Relatórios semanais serão publicados no seguinte endereço:

<http://www.ime.usp.br/~renatolg/mac0215/doc/reports/>

É também possível acompanhar tanto os relatórios quanto as implementações pelo repositório do projeto:

<https://github.com/RenatoGeh/mac0215/>

3 SUM-PRODUCT NETWORKS

3.1 DEFINIÇÃO

Sum-Product Networks são uma nova classe de modelos probabilísticos cuja inferência é sempre tratável.

Definição. *Uma SPN é, pela definição de Gens e Domingos[4]:*

1. *Uma distribuição univariante tratável.*
2. *Um produto de SPNs cujos escopos são disjuntos.*
3. *Uma soma de SPNs com peso não negativo cujos elementos tem mesmo escopo.*
4. *Nada mais é uma SPN.*

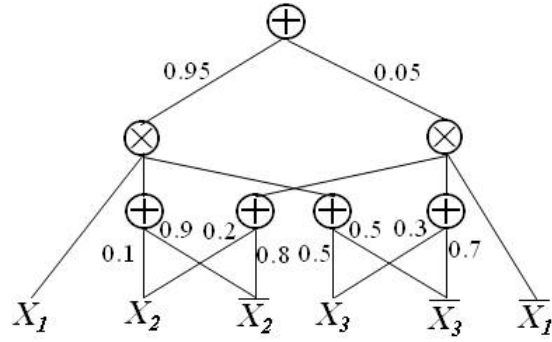


Figura 1: Um exemplo de uma SPN com variáveis Booleanas, onde x_1, \dots, x_d e $\overline{x_1}, \dots, \overline{x_d}$ são folhas e o resto dos nós são somas ou produtos.[6]

Podemos definir graficamente uma SPN como um grafo enraizado, direcionado e acíclico onde as folhas são sempre variáveis (ou distribuições univariantes), seus nós internos são somas ou produtos e para todo vértice conectando um nó soma com um nó filho há um peso não negativo. Também assume-se que toda soma e produto estão em alturas alternantes, ou seja, todo nó pai de um nó interno que é soma é um produto e vice-versa.

3.2 CARACTERÍSTICAS

Há várias vantagens de SPNs sobre outras redes de aprendizado:

1. Estudos mostram que SPNs tem estrutura parecida a Modelos Gráficos Probabilísticos (PGM), mas inferência é mais rápida e precisa em dependências com grande largura.[4]
2. SPNs apresentam inferência mais rápida comparado a Redes Bayesianas, cuja inferência é NP-difícil.
3. Experimentos mostram que aprendizado de arquiteturas SPN tiveram melhores resultados quando comparadas a outras arquiteturas estáticas.[2]

3.3 APLICAÇÕES

SPNs obtiveram resultados impressionantes em muitos conjuntos de dados[3], tais como:

- Reconstrução de imagens.
- Classificação.
- Reconhecimento de atividade.
- Logs click-through.
- Sequências de ácido nucleico.
- Filtragem colaborativa.

4 REFERÊNCIAS

ARTIGOS

- [1] Aaron Dennis e Dan Ventura. “Greedy Structure Search for Sum-Product Networks”. Em: *International Joint Conference on Artificial Intelligence* 24 (2015).
- [2] Aaron Dennis e Dan Ventura. “Learning the Architecture of Sum-Product Networks Using Clustering on Variables”. Em: *Advances in Neural Information Processing Systems* 25 (2012).
- [4] Robert Gens e Pedro Domingos. “Learning the Structure of Sum-Product Networks”. Em: *International Conference on Machine Learning* 30 (2013).
- [5] Sang-Woo Lee, Christopher Watkins e Byoung-Tak Zhang. “Non-Parametric Bayesian Sum-Product Networks”. Em: *Workshop on Learning Tractable Probabilistic Models* (2014).
- [6] Hoifung Poon e Pedro Domingos. “Sum-Product Networks: A New Deep Architecture”. Em: *Uncertainty in Artificial Intelligence* 27 (2011).

WEBSITES

- [3] University of Washington Department of Computer Science. *Sum-Product Networks*. URL: <http://spn.cs.washington.edu/index.shtml>.