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 Networks a partir de um conjunto de dados, avaliando o desempenho e os resultados das técnicas utilizadas.

2 PROJETO

2.1 DEFINIÇÃO

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

Definição. Uma SPN é:

- 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.

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.

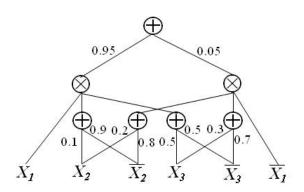


Figura 1: Um exemplo de uma SPN com variáveis Booleanas, onde $x_1,...,x_d$ e $\overline{x_1},...,\overline{x_d}$ são folhas e o resto dos nós são somas ou produtos. [Domingos Poon]

2.2 CARACTERÍSTICAS

Há várias vantagens de SPNs sobre outras redes de aprendizado:

- 1. Estudos mostram que SPNs tem estrutura parecida a Modelos Gráficos Probabilisticos (PGM), mas inferência é mais rápida e precisa em dependências com grande largura. [Gens Domingos]
- 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. [Clustering Dennis Ventura]

2.3 APLICAÇÕES

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

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