



**USP — UNIVERSIDADE DE SÃO PAULO**

Aprendizagem automática de redes soma-produto

Candidato: Renato Lui Geh

Orientador: Prof. Dr. Denis Deratani Mauá

**São Paulo**

**2017**

# 1 Introdução

Aprendizado de máquina é uma área da Inteligência Artificial cujo objetivo é estatisticamente ajustar os parâmetros de um certo modelo matemático dado um conjunto de dados a fim de que seja possível fazer previsões acuradas do que se está modelando. Para isso, faz-se uso de modelos estatísticos e métodos computacionais. Modelos probabilísticos baseados em grafos (PGM, do inglês *Probabilistic Graphical Model*), são uma classe destes modelos estatísticos em que se usa grafos para representá-los graficamente.

Modelos probabilísticos baseados em grafos representam uma distribuição de probabilidade de forma compacta. Estes modelos representados por grafos facilitam tanto a compreensão humana ao estudá-los, quanto possibilitam que vários problemas já existentes em Teoria dos Grafos sejam utilizados como solução para problemas em PGMs. Extrair conhecimento de PGMs é análogo a extrair a probabilidade de um certo evento ocorrer dado que eventos distintos tenham ocorrido. Tal extração de conhecimento é chamada de inferência. Fazer inferência exata em PGMs clássicas, ou seja, acharmos a probabilidade exata de um certo evento, é intratável. Uma solução para este problema é utilizarmos métodos para inferência aproximada nestes modelos. No entanto, tais algoritmos aproximados são muitas vezes difíceis de analisar. Além disso, como os algoritmos de aprendizado do modelo utilizam inferência como subrotina, por consequência o aprendizado torna-se aproximado.

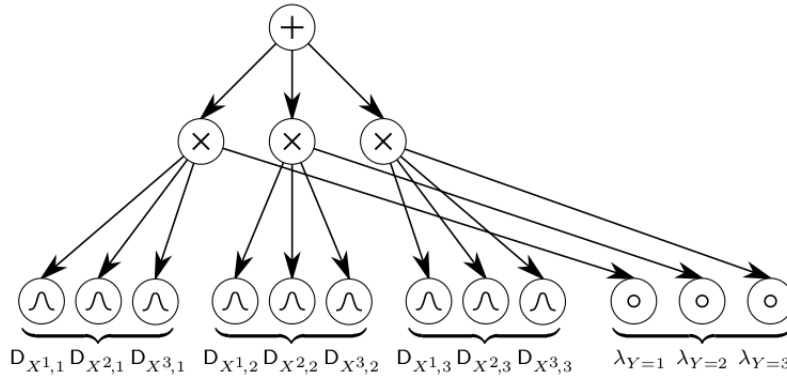


Figura 1

Redes soma-produto (SPN, de *Sum-Product Network*) são PGMs que representam uma distribuição de probabilidade tratável. Proposto em 2011, SPNs computam inferência exata em tempo linear ao número de arestas de seu grafo se sua estrutura obedecer a certas propriedades [9]. A Figura 1 mostra um exemplo de rede soma-produto representando o modelo Naïve Bayes com três atributos. SPNs apresentam uma série de características interessantes, como sua arquitetura profunda que permite representar funções de forma mais eficiente quanto mais profundo seu grafo [3]. Outras interessantes propriedades teóricas incluem uma generalização de SPNs para qualquer semianel em que o produto tenha escopo disjunto [6]. Com relação a aplicações, SPNs tiveram resultados impressionantes em diversas áreas, como enovelamento de proteínas [5], modelagem de sinais [8], classificação e reconstrução de imagens [7, 9, 4], reconhecimento de atividade [1] e linguagem natural [2].

Neste projeto, pretende-se desenvolver uma biblioteca livre e gratuita para inferência e algoritmos de aprendizado estado-da-arte em SPNs, analisando-se experimentalmente tais algoritmos em tarefas reais, como compleição e classificação de imagens.

## 2 Objetivos

Os objetivos esperados neste trabalho são:

1. Domínio de conceitos sobre aprendizagem de SPNs.
2. Desenvolvimento de uma biblioteca livre e gratuita de algoritmos de aprendizado de SPNs estado-da-arte.
3. Análise experimental comparativa em tarefas reais.

## 3 Metodologia

A pesquisa terá seu início com o estudo dos conceitos e do estado-da-arte de SPNs. Uma rede soma-produto pode ser definida como um digrafo acíclico cujos nós podem ser somas ponderadas, produtos, uma distribuição univariável ou variável indicadora. O escopo de um nó de uma SPN é o conjunto de todas variáveis que aparecem em seus descendentes. Adicionalmente, para que uma SPN compute inferência exata, é suficiente que os filhos dos nós somas tenham mesmo escopo e os filhos de nós produtos tenham escopos disjuntos.

Aprendizado de parâmetros envolve uma estrutura fixa com parâmetros dos nós soma variáveis a fim de maximizar a verossimilhança do modelo em relação a um conjunto de dados. Similarmente ao aprendizado paramétrico, o aprendizado estrutural maximiza a verossimilhança descobrindo-se tanto os pesos quanto a estrutura. Nesta pesquisa, estudar-se-á diferentes métodos de aprendizagem de SPNs. Primeiro, será iniciado o estudo e implementação do algoritmo de aprendizagem de parâmetros de Poon e Domingos [9], replicando-se a estrutura da rede para compleição de imagens descrita no artigo. Em seguida, serão estudados os algoritmos de aprendizado estrutural de Ventura e Dennis [4], Gens e Domingos [7], e Vergari e di Mauro [10]. Após o entendimento de cada algoritmo de aprendizagem, será iniciada a implementação do algoritmo.

A implementação dos algoritmos será desenvolvida em uma biblioteca gratuita, em código livre e repositório aberto, buscando documentar o código e elaborar relatórios. Com o desenvolvimento dos diferentes algoritmos de aprendizagem, pretende-se elaborar relatórios contendo uma comparação experimental de todos os algoritmos implementados em conjuntos de dados reais.

Durante a realização do projeto, serão elaborados documentos científicos para possível submissão a ENIAC 2018 e para o acompanhamento geral do desempenho acadêmico do candidato.

O cronograma para o desenvolvimento desta pesquisa segue abaixo.

Atividade/Mês	1º-2º	3º-4º	5º-6º	7º-8º	9º-10º	11º-12º
a	×					
b	×					
c		×				
d		×				
e			×			
f				×		
g	×	×	×	×		
h					×	
i					×	×
j						×

- a. Estudo de conceitos de SPN.
- b. Estudo e implementação do algoritmo de aprendizado de Poon e Domingos (2011) [9].
- c. Estudo e replicação da rede para compleção de imagens descrita em Poon e Domingos (2011).
- d. Estudo e implementação do algoritmo de aprendizado de estrutura de Ventura e Dennis (2012) [4].
- e. Estudo e implementação do algoritmo de aprendizado de estrutura de Gens e Domingos (2013) [7].
- f. Estudo e implementação do algoritmo de aprendizado de estrutura de Vergari e di Mauro (2016) [10].
- g. Avaliação de desempenho de algoritmos em problemas de compleção de imagem.
- h. Elaboração de artigo para ENIAC 2018.
- i. Relatório e participação no SIICUSP.
- j. Disponibilização da biblioteca e documentação do código.

## Referências

- [1] Mohamed R. Amer e Sinisa Todorovic. “Sum-Product Networks for Activity Recognition”. Em: *IEEE Transactions on Pattern Recognition and Machine Intelligence (TPAMI 2015)* (2015).
- [2] Wei-Chen Cheng et al. “Language Modelling with Sum-Product Networks”. Em: *Annual Conference of the International Speech Communication Association 15 (INTER-SPEECH 2014)* (2014).
- [3] Olivier Delalleau e Yoshua Bengio. “Shallow vs. Deep Sum-Product Networks”. Em: *Advances in Neural Information Processing Systems 24 (NIPS 1011)* (2011).
- [4] 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).
- [5] Abram L. Friesen e Pedro Domingos. “Recursive Decomposition for Non-convex Optimization”. Em: *International Joint Conference on Artificial Intelligence 24 (IJCAI 2015)* (2015).
- [6] Abram L. Friesen e Pedro Domingos. “The Sum-Product Theorem: A Foundation for Learning Tractable Models”. Em: *International Conference on Machine Learning 33 (ICML 2016)* (2016).
- [7] Robert Gens e Pedro Domingos. “Learning the Structure of Sum-Product Networks”. Em: *International Conference on Machine Learning 30* (2013).
- [8] Robert Peharz et al. “Modeling Speech with Sum-Product Networks: Application to Bandwidth Extension”. Em: *IEEE International Conference on Acoustics, Speech, and Signal Processing 39 (ICASSP 2014)* (2014).
- [9] Hoifung Poon e Pedro Domingos. “Sum-Product Networks: A New Deep Architecture”. Em: *Uncertainty in Artificial Intelligence 27* (2011).
- [10] Antonio Vergari, Nicola di Mauro e Floriana Esposito. “Simplifying, Regularizing and Strengthening Sum-Product Network Structure Learning”. Em: *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD 2015)* (2015).