



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

Incremental Learning de Redes Bayesianas e o UnBBayes

Pedro da C. Abreu

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientador

Prof. Dr. Marcelo Ladeira

Coorientador

Prof. Dr. Shou Matsumoto

Brasília
2015



Universidade de Brasília

**Instituto de Ciências Exatas
Departamento de Ciência da Computação**

Incremental Learning de Redes Bayesianas e o UnBBayes

Pedro da C. Abreu

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Prof. Dr. Marcelo Ladeira (Orientador)
CIC/UnB

Prof. Dr. Donald Knuth Dr. Leslie Lamport
Stanford University Microsoft Research

Prof.^a Dr.^a Ada Lovelace
Coordenadora do Bacharelado em Ciência da Computação

Brasília, 07 de dezembro de 2015

Dedicatória

Na *dedicatória* o autor presta homenagem a alguma pessoa (ou grupo de pessoas) que têm significado especial na vida pessoal ou profissional. Por exemplo (e citando o poeta):
Eu dedico essa música a primeira garota que tá sentada ali na fila. Brigado!

Agradecimentos

Nos *agradecimentos*, o autor se dirige a pessoas ou instituições que contribuíram para elaboração do trabalho apresentado. Por exemplo: *Agradeço aos gigantes cujos ombros me permitiram enxergar mais longe. E a Google e Wikipédia.*

Resumo

Com o crescente interesse em aprendizado de máquina, frameworks robustos e implementações do estado da arte são uma necessidade do mundo moderno. Pensando nisso levantamos um estudo sobre o estado da arte de aprendizado no contexto de Rede Bayesiana (BN), já que BNs provêm um excelente modelo para lidar com relações causais e probabilidade bayesiana. Para tal utilizaremos o framework UnBBayes e desenvolveremos plugins para este.

Palavras-chave: LaTeX, metodologia científica, trabalho de conclusão de curso

Abstract

With the growing interest in machine learning, good frameworks and state of the art implementation are a need in modern world. With this in mind we provide a study in the state of the art of BN, since BNs provide a great model to deal with causal relations and bayesian probability. For such we shall use UnBBayes framework and implement plugins for it.

Keywords: LaTeX, scientific method, thesis

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Estrutura da Monografia	3
2	Probabilidade	4
2.1	Probabilidade física versus probabilidade Bayesiana	4
2.2	Axiomas da Probabilidade	5
3	Redes Bayesianas	6
3.1	Definição Formal	7
3.2	D-separation	7
3.3	M-separation	7
3.4	Aprendizado	7
4	Aprendizado de Redes Bayesianas	8
4.1	Aprendizado de Parametros	8
4.2	Aprendizado de Estrutura	8
4.3	Aprendizado Incremental	8
5	UnBBayes	9
5.1	Plugins	9
5.1.1	Como desenvolver plugins	9
6	Design Patterns	10
6.1	Padrões de Comportamento	10
6.2	Padrões de Estrutura	10
6.3	Padrões de Construção	10
7	Arquitetura do UnBBayes	11

8	Resumo das Atividades	12
8.1	O que foi feito	12
8.2	O que será feito	12
9	Conclusão	13

Lista de Figuras

3.1	extraída do artigo do Laecio	6
-----	--	---

Lista de Abreviaturas e Siglas

BN Bayesian Network.

CPT Conditional Probability Tables.

DAG Directed Acyclic Graph.

GMU George Mason University.

UnB Universidade de Brasília.

Capítulo 1

Introdução

No cotidiano do ser humano é muito comum raciocinar e tomar decisões sob condições de incerteza. Isto é tão visível e profundo para algumas pessoas que no século XVIII Bishop Butler declarou "probabilidade é o guia da vida".

Como um exemplo do quanto probabilidade é importante tome a medicina. Para um especialista médico determinar a doença de uma pessoa com base nos sintomas observados (evidências) é preciso que ele leve em conta a probabilidade daquele sintoma refletir esta ou aquela doença, pois a doença ocasiona determinado sintoma apenas com alguma probabilidade, mas não com certeza.

Podemos ainda citar vários outros exemplos, como na economia, teoria dos jogos, genética, previsão do tempo.

Pensando nisso é necessário um modelo robusto para lidar com tantas incertezas, para tanto escolhemos as BNs.

1.1 Motivação

BN é um modelo gráfico para relações probabilísticas dado conjunto de variáveis. Nas últimas décadas, redes Bayesianas se tornaram representações populares para codificar conhecimento incerto para sistemas especialistas [?]. Mais recentemente, pesquisadores desenvolveram métodos de aprendizagem de redes Bayesianas a partir de dados. As técnicas desenvolvidas são relativamente novas e ainda em evolução, mas eles têm se mostrado muito eficientes para alguns problemas de análise de dados.

Existem diversas representações possíveis para análise de dados, entre elas, decision trees, e redes neurais artificiais; e outras tantas como estimação de densidade, classificação, regressão e clusterings. Portanto o que métodos de BNs têm a oferecer? Segundo Heckerman [?] podemos oferecer pelo menos quatro respostas, sendo elas:

1. BNs lidam com um conjunto incompleto de dados de maneira natural.

2. BNs permitem aprender sobre as relações causais. Aprender sobre tais relações são importantes por pelo menos duas razões: O processo é útil quando se está tentando entender sobre um dado problema de domínio, como por exemplo, durante uma análise de dados exploratória. E mais, conhecimento de relações causais nos permitem fazer previsões na presença de intervenções. Por exemplo, um analista de mercado pode querer saber se é lucrativo aumentar o investimento em determinada propaganda para aumentar as vendas de seu produto. Para responder esta pergunta o analista pode determinar se esta propaganda é a causa para o aumento de suas vendas, e em caso afirmativo, quanto. O uso de BNs nos ajuda a responder tal pergunta até mesmo quando não há experimentos nos efeitos de tal propaganda.
3. BNs em conjunto com técnicas estatísticas bayesianas facilitam a combinação de conhecimento de domínio e dados. Qualquer um que tenha feito uma análise do mundo real sabe a importância de conhecimento prévio ou de domínio, em especial quando os dados são poucos ou caros. Pelo fato de alguns sistemas comerciais (i.e., sistemas especialistas) podem ser construídos a partir de conhecimentos prévios. BNs possuem uma semântica causal que permitem conhecimentos prévios serem representados de uma forma muito simples e natural. Além disto, BNs encapsulam tais relações causais com suas probabilidades. Consequentemente, conhecimento prévio e dados podem ser combinados com técnicas bem estudadas da estatística Bayesiana.
4. Métodos Bayesianos em conjunto com BNs e outros tipos de modelos oferecem uma forma eficiente para evitar over fitting dos dados. Como veremos, não há necessidade de excluir parte dos dados do treinamento do aprendizado da rede. Usando técnicas Bayesianas, modelos podem ser "suavizados" de tal forma que todo dado disponível pode ser usado para o treinamento.

1.2 Objetivos

O objetivo deste trabalho é apresentar os conhecimentos adquiridos durante este semestre na disciplina de Estudos Em Inteligência Artificial com o professor Ladeira da Universidade de Brasília (UnB), com a colaboração do Doutorando Shou Matsumoto da George Mason University (GMU), através de aulas pelo skype. E também firmar bases sólidas sobre os conhecimentos que serão necessários na concepção da verdadeira monografia, isto é, o trabalho de conclusão de curso.

1.3 Estrutura da Monografia

Este documento é organizado da seguinte maneira: no capítulo 2 explicamos a visão probabilística necessária, isto é, firmamos a base do trabalho que se segue, como todos axiomas e definições necessários. No capítulo 3 formalizamos o conceito de BNs e algumas de suas propriedades. No capítulo 4 discutimos como construí-las a partir de um conjunto de dados. No capítulo 5 explicamos o framework do UnBBayes, como ele funciona e como extendê-lo. O capítulo 5 discutimos alguns dos Design Patterns mais importantes. No capítulo 6 resumimos o trabalho do semestre e oferecemos uma prévia dos trabalhos a serem feitos nos próximos semestres.

Capítulo 2

Probabilidade

2.1 Probabilidade física versus probabilidade Bayesiana

Para entender BNs e as técnicas de aprendizado associadas, é importante entender a diferença entre a Probabilidade e Estatística padrão e a Bayesiana.

A Probabilidade física, também conhecida como probabilidade frequentista, foi advogada pelo matemático John Venn [?] no século XIX, identificando probabilidade com frequências dos eventos no longo prazo. A reclamação mais óbvia que surge neste modelo é que as frequências no curto prazo obviamente não casam com as calculadas, por exemplo, se jogarmos a moeda apenas uma vez, certamente concluiríamos que a probabilidade de cara é ou 1 ou 0.

Uma alternativa ao conceito de probabilidade física é pensar nas probabilidades como nosso grau de crença subjetivo. Esta visão foi expressa por Thomas Bayes [?] e Pierre Simon de Laplace [?] a duzentos anos atrás. Esta é uma visão mais geral da probabilidade pois leva em conta que temos crenças subjetivas em um grande variedade de proposições, muitas das quais não estão claramente ligadas a um processo físico. Por exemplo, a maioria de nós acreditamos na Hipótese de Copérnico de que a terra orbita em torno do sol, mas isto é baseado em evidencia não obviamente da mesma forma que um processo de amostragem. Isto é, ninguém é capaz de gerar sistemas solares repetidamente e observar a frequência com a qual os planetas giram em torno do sol. Seja como for Bayesianistas estão preparados para conversar sobre a probabilidade da tese de Copérnico ser verdadeira e ainda relacionar as relações de evidências a favor e contra ela.

Bayesianismo pode ser visto como uma generalização da probabilidade física. Para isto adotamos o que David Lewis apelidou de Principal Principle [?]: sempre que se aprender uma probabilidade física de uma amostragem é r , atualize sua probabilidade subjetiva

para aquela amostragem r . Basicamente isto é senso comum: pense na probabilidade de um colega raspar a cabeça como 0.01, mas se você aprende que ele faz isso se e somente se um dado justo for lançado e der 2, você revisará sua opinião de acordo.

Tendo explicado como as probabilidades Físicas e Bayesianas são compatíveis podemos seguir para a definição dos axiomas da probabilidade.

2.2 Axiomas da Probabilidade

Seja U o universo de possíveis eventos. Os Axiomas de Kolmogorov [?] são:

Axiom 2.1. $P(U) = 1$

Axiom 2.2. *for all $X \subseteq U$, $P(X) \geq 0$*

Axiom 2.3. *for all $X, Y \subseteq U$, if $X \cap Y = \emptyset$, then $P(X \cup Y) = P(X) + P(Y)$*

Capítulo 3

Redes Bayesianas

Uma BN provê uma representação compacta de distribuições de probabilidades grandes demais para lidar usando especificações tradicionais e provê um método sistemático e localizado para incorporar informação probabilística sobre uma dada situação.

Uma BN é um Directed Acyclic Graph (DAG) que representa uma função de distribuição de probabilidades conjunta de variáveis que modelam certo domínio de conhecimento. Ela é constituída de uma DAG, de variáveis aleatórias (também chamadas de nós da rede), arcos direcionados da variável pai para a variável filha e uma Conditional Probability Tables (CPT) associada a cada variável.

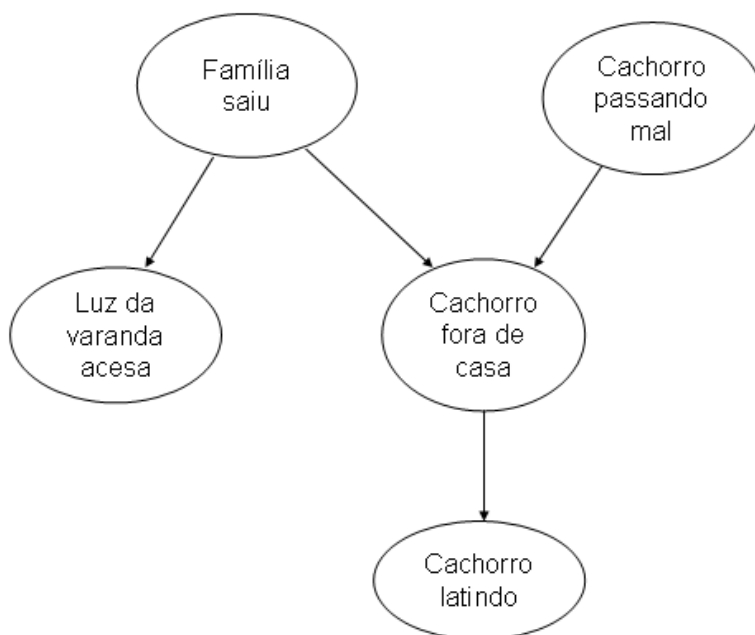


Figura 3.1: Exemplo family-out

Nesse exemplo, suponhamos que se queira determinar se a família está em casa ou se ela saiu. Pelo grafo, podemos perceber que o fato de a luz da varanda estar acesa e de o cachorro estar fora de casa são indícios de que a família tenha saído.

3.1 Definição Formal

Uma BN consiste em uma fatoração de uma distribuição de probabilidade e um DAG correspondente. Tais assertivas de independências condicionais podem ser inferidas diretamente da fatoração correspondente às

3.2 D-separation

3.3 M-separation

3.4 Aprendizado

Capítulo 4

Aprendizado de Redes Bayesianas

4.1 Aprendizado de Parametros

4.2 Aprendizado de Estrutura

4.3 Aprendizado Incremental

Capítulo 5

UnBBayes

5.1 Plugins

5.1.1 Como desenvolver plugins

Capítulo 6

Design Patterns

6.1 Padrões de Comportamento

6.2 Padrões de Estrutura

6.3 Padrões de Construção

Capítulo 7

Arquitetura do UnBBayes

Capítulo 8

Resumo das Atividades

8.1 O que foi feito

8.2 O que será feito

Capítulo 9

Conclusão