



Universidade de Brasília

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

UnBBayes- Um Estudo

Guilherme N. Ramos

Dissertação apresentada como requisito parcial
para conclusão do Mestrado Profissional em Computação Aplicada

Orientador

Prof. Dr. Marcelo Ladeira

Coorientador

Prof. Shou Matsumoto

Brasília
2015

Universidade de Brasília — UnB
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Mestrado Profissional em Computação Aplicada

Coordenadora: Prof.^a Dr.^a Coorde Nadora

Banca examinadora composta por:

Prof. Dr. Marcelo Ladeira (Orientador) — CIC/UnB
Prof.^a Dr.^a Membra da Banca — MEC
Prof. Dr. Membro do Banco — CIC/UnB

CIP — Catalogação Internacional na Publicação

Ramos, Guilherme N..

UnBBayes- Um Estudo / Guilherme N. Ramos. Brasília : UnB, 2015.

19 p. : il. ; 29,5 cm.

Dissertação (Mestrado) — Universidade de Brasília, Brasília, 2015.

1. L^AT_EX, 2. Redes Bayesianas, 3. UnBBayes

CDU 004.4

Endereço: Universidade de Brasília
Campus Universitário Darcy Ribeiro — Asa Norte
CEP 70910-900
Brasília-DF — Brasil



UnBBayes- Um Estudo

Dissertação apresentada como requisito parcial
para conclusão do Mestrado Profissional em Computação Aplicada

Prof.^a Dr.^a Membro da Banca MEC Prof. Dr. Membro do Banco CIC/UnB

Prof.^a Dr.^a Coorde Nadora
Coordenadora do Mestrado Profissional em Computação Aplicada

Brasília, 07 de dezembro de 2015

Dedicatória

Citando o poeta: “Eu dedico essa música a primeira garota que tá sentada ali na fila. Brigado!”

Agradecimentos

Agradeço ao Prof. José Ralha, cujos esforços na versão anterior foram bem “adaptados” a este trabalho.

Resumo

Um parágrafo resumindo todo o trabalho. Classe L^AT_EX para gerar documentos do Departamento de Ciência da Computação da Universidade de Brasília.

Palavras-chave: L^AT_EX, Redes Bayesianas, UnBBayes

Abstract

Um parágrafo *na língua Inglesa* resumindo todo o trabalho. A L^AT_EX class for generating documents for the Departament of Computer Science of the University of Brasília.

Keywords: L^AT_EX, scientific method

Sumário

1	Introdução	1
1.1	Probabilidade e Estatística Bayesiana	2
2	Design Patterns	3
2.1	Metologia Científica	3
2.1.1	Fontes de Informação Digital	3
2.2	L ^A T _E X	3
2.3	Normas?	4
3	Arquitetura do UnBBayes	5
4	Como desenvolver um Plugin	6
5	Redes Bayesianas	7
5.1	Definição Formal	8
	Referências	9

Lista de Figuras

5.1 extraída do artigo do Laecio	7
--	---

Lista de Tabelas

Lista de Abreviaturas e Siglas

BN Bayesian Network.

CPT Conditional Probability Tables.

DAG Directed Acyclic Graph.

Capítulo 1

Introdução

Rede Bayesiana (BN) é um modelo gráfico para relações probabilísticas dado conjunto de variáveis. Nas últimas décadas, redes Bayesianas se tornaram representações populares para codificar conhecimento especialista incerto para sistemas especialistas [2]. Mais recentemente, pesquisadores desenvolveram métodos de aprendizagem de redes Bayesianas a partir de dados. As técnicas desenvolvidas são 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, rule bases, 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 [2] 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.

1.1 Probabilidade e Estatística Bayesiana

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

Resumidamente

Capítulo 2

Design Patterns

Este documento serve de exemplo da utilização da classe `UnB-CIC` para escrever um texto cujo objetivo é apresentar os resultados de um trabalho. A sequência de ideias apresentada deve fluir claramente, de modo que o leitor consiga compreender os principais conceitos e resultados apresentados, bem como encontrar informações sobre conceitos secundários.

2.1 Metodologia Científica

2.1.1 Fontes de Informação Digital

- [Google Acadêmico](#)
- [ACM Digital Library](#)
- [Portal CAPES](#)
- [IEEE Xplore](#)
- [ScienceDirect](#)
- [Springer Link](#)

2.2 L^AT_EX

Eis alguns links interessantes para familiarização com L^AT_EX:

- [L^AT_EX](#)
- [Introdução ao LaTeX](#)
- [Aprendendo L^AT_EX em 5 minutos](#)

- [L^AT_EX Wikibook](#)
- [T_EX - L^AT_EX Stack Exchange](#)

A rede CTAN (<http://www.ctan.org/>), idealizada em [1], oferece acesso milhares de contribuições do sistema.

2.3 Normas?

Política de Publicação de Monografias e Dissertações no Repositório Digital do CICRepositório
do Departamento de Ciência da Computação da UnB
Biblioteca Digital de Monografias de Graduação e Especialização

Capítulo 3

Arquitetura do UnBBayes

Capítulo 4

Como desenvolver um Plugin

Este documento serve de exemplo da utilização da classe `UnB-CIC` para escrever um texto cujo objetivo é apresentar os resultados de um trabalho. A sequência de ideias apresentada deve fluir claramente, de modo que o leitor consiga compreender os principais conceitos e resultados apresentados, bem como encontrar informações sobre conceitos secundários.

<http://www.escritacientifica.com/>

Capítulo 5

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.

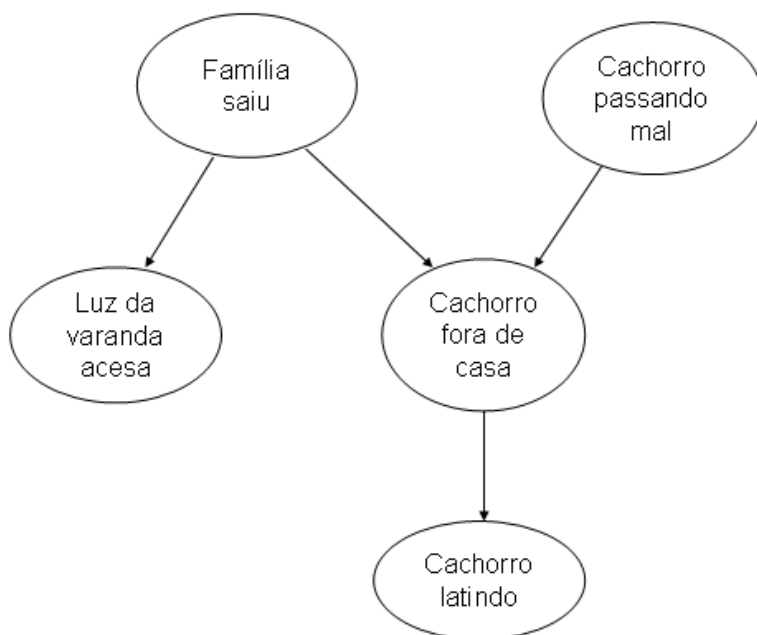


Figura 5.1: Exemplo family-out

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.

5.1 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

Referências

- [1] George D. Greenwade. The Comprehensive Tex Archive Network (CTAN). *TUGBoat*, 14(3):342–351, 1993. 4
- [2] David Heckerman. A tutorial on learning with bayesian networks. 156:33–82, 2008. 1