

FLÁVIO LUIZ SEIXAS

SISTEMA DE APOIO À DECISÃO APLICADO AO DIAGNÓSTICO DE DEMÊNCIA,
DOENÇA DE ALZHEIMER E TRANSTORNO COGNITIVO LEVE

Tese apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Doutor. Área de Concentração: Computação Visual.

Orientadora: Profa. Dra. AURA CONCI

Coorientadora: Profa. Dra. DÉBORA CHRISTINA MUCHALUAT SAADE

Niterói

2012

Ficha Catalográfica elaborada pela Biblioteca da Escola de Engenharia e Instituto de Computação da UFF

462 Seixas, Flávio Luiz
Sistema de apoio à decisão aplicado ao diagnóstico de demência,
doença de Alzheimer, e transtorno cognitivo leve / Flávio Luiz
Seixas. – Niterói, RJ : [s.n.], 2012.
155 f.

Tese (Doutorado em Computação) - Universidade Federal
Fluminense, 2012.

Orientadores: Aura Conci, Débora Christina Muchaluat Saade.

1. Sistema de apoio à decisão. 2. Doença de Alzheimer. 3.
Demência. 4. Transtorno cognitivo leve. 5. Rede Bayesiana.

I: Título

CDD 658.403

FLÁVIO LUIZ SEIXAS

SISTEMA DE APOIO À DECISÃO APLICADO AO DIAGNÓSTICO DE DEMÊNCIA,
DOENÇA DE ALZHEIMER E TRANSTORNO COGNITIVO LEVE

Tese apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense, como requisito parcial para obtenção do Grau de Doutor. Área de Concentração: Computação Visual.

Aprovada em 19 de Novembro de 2012.

BANCA EXAMINADORA

Profa. Dra. AURA CONCI – Orientadora
Universidade Federal Fluminense

Profa. Dra. DÉBORA CHRISTINA MUCHALUAT SAADE – Orientadora
Universidade Federal Fluminense

Prof. Dr. ARISTÓFANES CORRÊA SILVA
Universidade Federal do Maranhão

Prof. Dr. ANSELMO CARDOSO DE PAIVA
Universidade Federal do Maranhão

Prof. Dr. ARTUR ZIVIANI
Laboratório Nacional de Computação Científica

Prof. Dr. CÉLIO VINÍCIUS NEVES DE ALBUQUERQUE
Universidade Federal Fluminense

Profa. Dra. BIANCA ZADROZNY
Universidade Federal Fluminense

Niterói
2012

aos meus pais e a minha esposa, Elaine.

AGRADECIMENTOS

Meus primeiros agradecimentos são endereçados aos meus pais, Mozart de Moraes Seixas e Maria Helena Zétola Seixas, pelo presente da Vida e pela educação, em uma unidade familiar plena de Amor, cumplicidade e integração.

Agradeço a minha esposa, Elaine Flavio Rangel Seixas, através do amor e apoio incondicional, prestou-me a autoconfiança e determinação, fundamental para a construção desta tese.

Agradeço a Marcos Oliveira Rangel e a Jussara Flavio Rangel, por prover todo o carinho, amizade e conforto necessário para a execução deste trabalho.

Meus agradecimentos a Timothy Wayne Cook e a Profa. Luciana Tricai Cavalini, pela inestimável ajuda na codificação dos conceitos clínicos, explicando com muita paciência, persistência e bom humor, os fundamentos da modelagem multinível das informações clínicas.

Esta tese carrega um desafio adicional devido a suas características multidisciplinares. Desta forma, agradeço ao Dr. Jerson Laks pela parceria na validação dos principais componentes do modelo de decisão proposto neste trabalho, bem como a ajuda na descrição dos conceitos relacionados à área médica. Agradeço também a toda equipe do Centro de Tratamento de Alzheimer da Universidade Federal do Rio de Janeiro, pela paciência e atenção prestada aos meus questionamentos, em especial, a Dra. Andrea Deslandes e a Dra. Cloyra de Paiva Almeida. Agradeço também a Dra. Gerda Fillenbaum por disponibilizar a base de casos clínicos do Centro de Estudo do Envelhecimento e Desenvolvimento Humano da *Duke University*.

Agradeço a Anna Cláudia Cupello de Freitas e a Abdo de Souza Chaim, pela confiança e generosidade, fornecendo todo o apoio e meios necessários para viabilizar este trabalho.

Como qualquer outro trabalho, esta tese é extensão da vida do autor, e produto de inúmeras conversas informais e contribuições voluntárias de amigos e colegas de profissão. Neste sentido, agradeço aos amigos e colegas da Embratel, em especial, Márcia Cristina França Nunes, pela amizade e ideias que guiaram alguns temas aqui apresentados. Agradeço também o apoio dos amigos e colegas da Universidade Estácio de Sá e da Universidade Federal Fluminense.

Agradeço a Profa. Bianca Zadrozny pela inestimável ajuda com os conceitos relacionados aos modelos causais e redes Bayesianas.

Por fim, meus agradecimentos a minhas professoras orientadoras, Profa. Aura Conci e Profa. Débora Christina Muchalut Saade, pela amizade, confiança e incomensurável apoio, através das criteriosas revisões, com uma atitude sempre positiva e acolhedora.

"Não podemos mudar a condição humana, mas podemos mudar as condições sob as quais os humanos trabalham", James Reason, *British Medical Journal*, v. 320, p. 768-770.

RESUMO

A preocupação quanto à qualidade, segurança e eficiência dos serviços de saúde é crescente. Sistemas de apoio à decisão estão sendo implantados com objetivo de melhorar a qualidade e eficiência dos serviços de saúde. A Doença de Alzheimer afeta mais de 12% da população de idosos (maiores que 65 anos de idade) dos Estados Unidos, ocupando a quinta posição das principais causas de óbitos da população, segundo relatório da *Alzheimer's Association* (2012). Esta tese propõe um sistema de apoio à decisão aplicado ao diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, este último, considerado como estágio pré-clínico da Doença de Alzheimer. O modelo de decisão é baseado em rede Bayesiana. Este modelo é adequado quando há necessidade de tratamento de incerteza e representação da causalidade, características presentes no contexto clínico. O processo de modelagem da rede Bayesiana é orientado por dados, sendo o diagrama causal construído com base em uma estrutura Bayesiana genérica e a orientação do especialista do domínio do conhecimento, e os parâmetros da rede estimados por um método supervisionado de aprendizagem. A base de treinamento é formada por um conjunto de pacientes e controle normal, fornecidos pelo Centro Médico da *Duke University* (Washington, Estados Unidos), e pelo Centro de Tratamento da Doença de Alzheimer (Universidade Federal do Rio de Janeiro). Os atributos da base de treinamento contêm fatores de pré-disposição, dados demográficos, pontuações de testes neuropsicológicos, sinais e sintomas de pacientes. Os resultados da aprendizagem são avaliados utilizando métodos quantitativos e análise de sensibilidade, comparando com uma análise do especialista. A ferramenta de autoria, inferência e aprendizagem Bayesiana empregada é GeNIe/SMILE (*Graphical Network Interface/Structural Modeling, Inference, and Learning Engine*), desenvolvida pela Universidade de Pittsburg. Os dados clínicos da rede Bayesiana são especificados utilizando o uma abordagem multinível com o modelo de referência MLHIM (*Multilevel Healthcare Information Modeling*). A plataforma de serviços web foi utilizada para implementação do sistema de apoio à decisão, com os parâmetros de entrada e saída especificados em WSDL (*Web Services Description Language*). Por fim, um protótipo da aplicação é construído, fornecendo uma interface gráfica para entrada das observações clínicas pelo médico.

Palavras-chave: Doença de Alzheimer, Rede Bayesiana, Diagrama de Influência, Sistema de Apoio à Decisão Clínica, Modelo de Referência de Informações de Saúde.

ABSTRACT

Healthcare quality and patient safety have been a growing concern for many countries. Decision support systems have been deployed to improve healthcare quality and reduce the risks of diagnostic errors. Dementia due to Alzheimer's disease affects more than 12% of elder population in the United States (people older than 65 years old), and ranks the fifth leading cause of death in this population, according to Alzheimer's Association (2012). In this thesis, we propose a decision support system used for assisting physicians on decision involving the clinical diagnosis of dementia, Alzheimer's disease and mild cognitive disorder. Mild cognitive disorder can be considered as an early stage of Alzheimer's disease. The decision theoretic model was designed using a Bayesian network, due to its ability to deal with uncertainty and causality, which are characteristics observed in the clinical domain. The Bayesian network was modeled using a data-driven approach, including a Bayesian structure built by a knowledge domain specialist using a generic Bayesian structure, and its parameters were estimated by a supervised learning method. The training base is composed by a real clinical case set of patients and normal controls from the the Duke University Medical Center (Washington, USA) and the Alzheimer's Disease Treatment Center of the Rio de Janeiro Federal University (Rio de Janeiro, Brazil). The training base attributes are composed by predisposal factors, assessment scales, patient demographic data, symptoms and signs. The model is evaluated using quantity methods and sensitivity analysis. The GeNIe/SMILE (Graphical Network Interface / Structural Modeling, Inference, and Learning Engine) software by Pittsburg University was used as a Bayesian authoring, inference and learning tool. The clinical concepts were modeled using a multilevel approach with the reference information model based on MLHIM (Multilevel Healthcare Information Modeling). The implementation is based on web services, and the input and output parameters are defined with WSDL (Web Services Description Language). A prototype application is also proposed, providing a graphical interface for physicians to indicate clinical observations.

Keywords: Alzheimer's Disease, Bayesian Network, Influence Diagram, Clinical Decision Support System, Healthcare Information Reference Model.

LISTA DE ILUSTRAÇÕES

Figura 1.	Componentes funcionais de um sistema de apoio à decisão clínica.....	12
Figura 2.	Grafo com segmentos acíclicos dirigidos.....	28
Figura 3.	Quatro estruturas padrões consideradas na definição de d -separação entre os nodos X e Y	28
Figura 4.	Diferentes tipos de raciocínio com incertezas.....	29
Figura 5.	Inferência Bayesiana baseada em fluxo de mensagens.....	30
Figura 6.	Exemplo genérico de diagrama de influência aplicado a uma decisão clínica. Adaptado de Korb e Nicholson (2004).	33
Figura 7.	Processo de modelagem da rede Bayesiana orientado por dados.....	40
Figura 8.	Processo de diagnóstico de Demência, Alzheimer (DA), Transtorno Cognitivo Leve (TCL).	43
Figura 9.	Representação da relação das caixas de decisão do diagrama de influência com os pontos de decisão representados no processo de diagnóstico.....	44
Figura 10.	Critério de seleção dos casos de pacientes da base de treinamento pelos grupos de interesse.	47
Figura 11.	Quantidade de instâncias por grupo e doença da base CERAD.	48
Figura 12.	Quantidade de instâncias por grupo e doença da base CDA.	49
Figura 13.	Quantidade de instâncias após aplicação do método de balanceamento das bases de treinamento.	50
Figura 14.	Função entropia $ H(X) $ com a variável aleatória X contendo dois estados.	51
Figura 15.	Representação gráfica dos níveis da rede Bayesiana, mostrando a dependência causal entre as variáveis e demais elementos de uma rede de decisão: função utilidade e caixa de decisão.	62
Figura 16.	Curva de aprendizado por <i>fold</i> , utilizando subconjunto de dados para Demência e DA, base CERAD.	64
Figura 17.	Curva de aprendizado por <i>fold</i> , utilizando subconjunto de dados para Demência, DA e TCL, base CDA.	64
Figura 18.	Diagrama de influência modelado para o diagnóstico de Demência, utilizando a base CERAD.	66
Figura 19.	Diagrama de influência modelado para o diagnóstico da Doença de Alzheimer, utilizando a base CERAD.	67

Figura 20. Diagrama de influência modelado para o diagnóstico da Demência, utilizando a base CDA.....	68
Figura 21. Gráfico Boxplot mostrando distribuição dos coeficientes de correlação obtidos da base CERAD e base CDA. As variáveis aleatórias da distribuição estão condicionadas ao nodo da rede Bayesiana que representa a doença sobre análise.	72
Figura 22. Curva ROC e a área sob a curva ROC (AUC).....	80
Figura 23. Rede Bayesiana descoberta por método computacional aplicada ao diagnóstico de Demência, utilizando base CDA.	89
Figura 24. Rede Bayesiana descoberta por método computacional aplicada ao diagnóstico de DA, utilizando base CDA.....	90
Figura 25. Rede Bayesiana descoberta por método computacional aplicado ao diagnóstico de TCL, utilizando base CDA.	91
Figura 26. Descrição do algoritmo meta classificador AdaBoostM1.	93
Figura 27. Componentes do sistema de suporte à decisão.	104
Figura 28. Diagrama de sequência mostrando a implantação do Sistema de Apoio à Decisão em uma arquitetura orientada a serviços.	105
Figura 29. Componentes da modelagem multinível para sistemas de informação da saúde (HIS).	113
Figura 30. Seção de metadados do CCD utilizado para especificação dos instrumentos de avaliação psiconeurológica. Código em XSD e RDF.....	115
Figura 31. Trecho do CCD utilizado para especificação da pontuação do instrumento de avaliação MMSE. Código em XSD.....	117
Figura 32. Trecho do CCD utilizado para especificação da pontuação do instrumento de avaliação CDR. Código em XSD.	118
Figura 33. Exemplo de trecho de código em XML, mostrando uma instância do MMSE...	119
Figura 34. Trecho do arquivo WSDL, com a descrição do serviço de publicação.	121
Figura 35. Trecho do arquivo WSDL, contendo a descrição do serviço de consulta.	122
Figura 36. Trecho do CCD com especificação da resposta do CDSS. Código em XSD.....	122
Figura 37. Resposta do CDSS codificada em XML, modelada em CCD.....	123
Figura 38. Exemplo de evidência da rede Bayesiana, obtida pela associação da instância clínica modelada em CCD, e a variável aleatória representada na rede Bayesiana.	124

Figura 39. Esquema mostrando a sequência de mensagem de requisição e resposta, e os parâmetros encaminhados para o serviço de consulta, hospedado no servidor CDSS.	125
Figura 40. Cópia de tela com o aplicativo cliente do serviço de publicação da rede Bayesiana no servidor CDSS.	126
Figura 41. Exemplo da associação dos controles do formulário dinâmico, com a especificação dos conceitos clínicos no modelo CCD.	127
Figura 42. Cópia da tela do aplicativo cliente do serviço de consulta ao servidor CDSS. ...	128
Figura 43. Cópia das telas do aplicativo cliente do serviço de consulta ao servidor CDSS, mostrando três momentos: (1) seleção do diagnóstico ou rede Bayesiana desejada, (2) preenchimento do formulário com as observações clínicas, (3) visualização da resposta.	129

LISTA DE TABELAS

Tabela 1.	Relação de funções e facilidades dos sistemas de apoio à decisão da área de saúde.	10
Tabela 2.	Relação de sistemas de apoio à decisão clínica por mecanismo de inferência.....	13
Tabela 3.	Exemplo de valores atribuídos a uma função utilidade genérica	26
Tabela 4.	Quatro cenários mostrando problemas de aprendizagem da rede Bayesiana.	35
Tabela 5.	Principais características da Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, conforme critérios publicados pelo Instituto Nacional de Envelhecimento (Estados Unidos).....	42
Tabela 6.	Quantidade de casos de pacientes por grupo e doença da base CERAD.....	48
Tabela 7.	Quantidade de casos de pacientes por grupo e doença da base CDA.....	49
Tabela 8.	Relação de atributos selecionados da base do CERAD como variável aleatória da rede Bayesiana para diagnóstico de Demência.....	55
Tabela 9.	Relação de atributos selecionados da base do CERAD como variável aleatória da rede Bayesiana para diagnóstico da Doença de Alzheimer.	56
Tabela 10.	Relação de atributos selecionados da base do CDA como variável aleatória da rede Bayesiana para diagnóstico da Demência, Doença de Alzheimer e Transtorno Cognitivo Leve.	58
Tabela 11.	Estados ou níveis discretos definidos para cada variável aleatória das redes Bayesianas da base do CDA.	60
Tabela 12.	Parâmetros genéricos utilizados na função utilidade.....	68
Tabela 13.	Correlação dois a dois das variáveis aleatórias da rede Bayesiana para diagnóstico de Demência, utilizando a base CERAD.....	70
Tabela 14.	Correlação dois a dois das variáveis aleatórias da rede Bayesiana para diagnóstico de Demência, utilizando a base CDA.	71
Tabela 15.	Associação cruzada de características presentes nos critérios de diagnóstico de Demência, DA e TCL, com atributos da base CDA.....	74
Tabela 16.	Associação em referência cruzada dos atributos da base CERAD com os atributos da base CDA.....	75
Tabela 17.	Resumo das medidas de desempenho da rede Bayesiana.....	79
Tabela 18.	Matriz confusão.	79
Tabela 19.	Medidas de desempenho e valores da matriz confusão obtidos em cada <i>fold</i> considerando a rede Bayesiana obtida pela base CERAD.....	82

Tabela 20. Principais parâmetros dos classificadores utilizados na comparação.....	83
Tabela 21. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CERAD.....	84
Tabela 22. Medidas de desempenho e valores da matriz confusão obtidos em cada <i>fold</i> considerando a rede Bayesiana obtida pela base CDA.....	85
Tabela 23. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CDA.....	86
Tabela 24. Desempenho apresentado pela rede Bayesiana descoberta por método computacional para o diagnóstico da doença de interesse, utilizando base CDA.	91
Tabela 25. Comparação de resultados de AUC e MSE, considerando o classificador e o mesmo classificador em conjunto com AdaBoostM1.....	94
Tabela 26. Comparação de resultados do AdaboostM1, avaliando a utilização direta da distribuição de pesos D_i pelo classificador fraco, e a distribuição de pesos por reamostragem do conjunto de treinamento.....	96
Tabela 27. Comparação de resultados do AdaboostM1, avaliando a sensibilidade da modificação do número de iterações de 10 para 20 nos resultados do classificador.	96
Tabela 28. Comparação de resultados do AdaboostM1, avaliando a sensibilidade da modificação do ponderação da distribuição de pesos D_i de 100% para 10% nos resultados do classificador.....	96
Tabela 29. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CDA.....	97
Tabela 30. Análise de sensibilidade mostrando as evidências mais relevantes para o diagnóstico (negativo e positivo) de Demência, DA e TCL, utilizando a base CDA.....	99
Tabela 31. Relação de avaliações neuropsicológicas e o grau de relevância para o diagnóstico positivo das doenças de interesse classificado pelo especialista.....	101
Tabela 32. Levantamento de formatos utilizados para representação de redes Bayesianas e/ou diagramas de influência.....	108
Tabela 33. Características desejadas em um modelo de representação do conhecimento	110
Tabela 34. Relação de conceitos associados ao CCD, identificando sua referência terminológica.....	116

LISTA DE ABREVIATURAS E SIGLAS

Sigla	Termo (Original)	Descrição
ADL	<i>Archetype Definition Language</i>	Linguagem de Definição de Arquétipos
AHRQ	<i>Agency of Healthcare Research and Quality</i>	Agência de Pesquisa e Qualidade da Saúde
AMIA	<i>American Medical Informatics Association</i>	Associação Americana de Informática Médica
ANN	<i>Artificial Neural Network</i>	Redes Neurais Artificiais
ANSI	<i>American National Standards Institute</i>	Instituto Nacional Americano de Padronização
APA	<i>American Psyquatric Association</i>	Associação Psiquiátrica Americana
AUC	<i>Area under Receiver Operating Characteristic</i>	Área da Curva ROC
BayPAD	<i>Bayes Pulmonary embolism Assisted Diagnosis</i>	Apoio ao Diagnóstico de Embolia Pulmonar utilizando Bayes
BIC	<i>Bayesian Information Criterion</i>	Critério de Informação Bayesiana
BNIF	<i>Bayesian Networks Interchange Format</i>	Formato de Intercâmbio de Redes Bayesianas
BPEL	<i>Business Process Execution Language</i>	Linguagem de Execução de Processos de Negócio
BPMN	<i>Business Process Modeling Notation</i>	Notação para Modelagem de Processos de Negócios
CAD	<i>Computer-Aided Diagnosis</i>	Diagnóstico Auxiliado por Computador
CADe	<i>Computer-Aided Detection</i>	Detecção Auxiliada por Computador
CADx	<i>Computer-Aided Diagnosis</i>	Diagnóstico Auxiliado por Computador
CAMCOG	<i>Cambridge Cognitive Examination</i>	Exame cognitivo de Cambridge
CCD	<i>Concept Constraint Definition</i>	Definição de Restrição de Conceito
CDA	Centro de tratamento da Doença de Alzheimer	
CDR	<i>Clinical Dementia Rating</i>	Avaliação clínica de demência

CDSS	<i>Clinical Decision Support System</i>	Sistema de Apoio à Decisão
CDT	<i>Clock drawing test scale</i>	Teste do desenho do relógio
CEN	<i>European Committee for Standardization</i>	Comitê Europeu de Padronização
CERAD	<i>Consortium to Establish a Registry for Alzheimer's Disease</i>	Consórcio para estabelecer métodos de avaliação de pessoas com suspeita de Alzheimer
CF	<i>Certain Factor</i>	Fator de Certeza
CGP	<i>Clinical Guideline Practices</i>	Diretrizes de Prática Clínica
CIG	<i>Computer Interpretable Guidelines</i>	Diretrizes Interpretáveis por Computador
CKS	<i>Clinical Knowledge Summaries</i>	Sumários de Conhecimento Clínico
CN	Controle Normal	
CONEP	Comissão Nacional de Ética em Pesquisa	
CPOL	<i>Care Plan On-line</i>	Plano de Tratamento on-line
CPT	<i>Conditional Probability Table</i>	Tabela de Probabilidade Condicional
DA	<i>Doença de Alzheimer</i>	
DAG	<i>Directed Acyclic Graph</i>	Grafo Acíclico Direcionado
DBN	<i>Dynamic Bayesian Networks</i>	Redes Bayesianas Dinâmicas
DM	<i>Domain Model</i>	Modelo de Domínio
DSM-IV-TR	<i>Diagnostic and Statistical Manual of mental disorders – Fourth Edition – Text Revision</i>	Manual Diagnóstico e Estatístico de Transtornos Mentais
EHR	<i>Electronic Health Records</i>	Registros Eletrônicos de Saúde
EM	<i>Expectation-Maximization algorithm</i>	Algoritmo de Expectation-Maximization
ESS	<i>Equivalent Sample Size</i>	Tamanho de Amostra Equivalente
fMRI	<i>functional Magnetic Resonance Images</i>	Ressonância Magnética Funcional
FN	<i>False Negative</i>	Negativos Falsos
FP	<i>False Positive</i>	Positivos Falsos
GLIF	<i>Guideline Interchange Format</i>	Formato de Intercâmbio de Diretrizes Clínicas

GUI	<i>Graphical User Interface</i>	Interface Gráfica do Usuário
HIS	<i>Health Information System</i>	Sistema de Informação Médica
HKCR	<i>Healthcare Knowledge Component Repository</i>	Repositório de Componentes de Conhecimento relacionado à Saúde
HL7	<i>Health Level Seven</i>	
IADL	<i>Instrumental Activity Daily Living</i>	Instrumento para avaliação da atividade de rotina diária
ICD	<i>World Health Organization's International Classification of Disease</i>	Classificação Internacional de Doenças pela Organização Mundial de Saúde
ID	<i>Influence Diagram</i>	Diagrama de Influência
IHTSDO	<i>International Health Terminology Standards Development Organization</i>	Organização Internacional de Desenvolvimento de Padrões Terminológicos para Saúde
IMASC	<i>Intelligent MultiAgent System for Clinical Decision Support</i>	Sistema Multiagente Inteligente para Apoio a Decisão Clínica
IQCode	<i>Informant Questionnaire on cognitive decline in the elderly</i>	Questionário do Informante sobre declínio cognitivo em idosos
ISSO	<i>International Organization for Standardization</i>	Organização Internacional de Padronização
MAP	<i>Maximum a posteriori Probability</i>	Máxima Probabilidade a posteriori
MCMC	<i>Markov Chain Monte Carlo</i>	Cadeia de Markov Monte Carlo
MDL	<i>Minimum Description Length</i>	Descrição de Comprimento Mínimo
MEBN	<i>Multi-Entity Bayesian Networks</i>	Redes Bayesianas Multi-Entidades
MLE	<i>Maximum Likelihood Estimation</i>	Estimativa de Máxima Verossimilhança
MLHIM	<i>Multilevel Health Information Modeling</i>	Modelagem Multinível de Informações de Saúde
MLM	<i>Medical Logic Module</i>	Módulo Lógico Médico
MMSE	<i>Mini-Mental State Examination</i>	Mini-exame de estado mental
MRI	<i>Magnetic Resonance Image</i>	Ressonância Magnética
MSE	<i>Mean Squared Error</i>	Erro Médio Quadrado
MXE	<i>Mean Cross-Entropy</i>	Entropia Média Cruzada
NHS	<i>National Health Service</i>	Serviço Nacional de Saúde - Inglaterra

NIA	<i>National Institute on Aging</i>	Instituto Nacional de Envelhecimento dos Estados Unidos
NINCS-ADRDA	<i>National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association</i>	Instituto Nacional de Distúrbios Neurológicos
NPI	<i>NeuroPsychiatric Inventory</i>	Inventário neuropsiquiátrico
OASIS	<i>Organization for the Advancement of Structured Information Standards</i>	Organização para Padrões Avançados de Informação Estruturada
PCA	<i>Principal Component Analysis</i>	Análise de Componente Principal
PDS	<i>Pediatric Decision Support System</i>	Sistema de Apoio a Decisão Pediátrica
PET	<i>Positrons Emission Tomography</i>	Tomografia de Emissão de Pósitrons
PR-OWL	<i>Probabilistic Ontology Web Language</i>	Representação probabilística de ontologias
QMR	<i>Quick Medical Reference</i>	Referência Médica
RIM	<i>Reference Information Model</i>	Modelo de Referência de Informação
RM	<i>Reference Model</i>	Modelo de Referência
ROC	<i>Receiver Operating Characteristic</i>	
SCHIN	<i>Sowerby Centre for Health Informatics at Newcastle</i>	Centro de Informática de Saúde de Newcastle
SNOMED-CT	<i>Systematized Nomenclature of Medicine - Clinical Terms</i>	Termos Clínicos - Nomenclatura Sistemática de Medicina
SOA	<i>Service-Oriented Architecture</i>	Arquitetura Orientada a Serviços
SPECT	<i>Single-Photon Emission Computed Tomography</i>	Tomografia Computadorizada de Emissão de Fótons Únicos
STROOP	<i>Stroop color word test</i>	Teste de desempenho de cor-palavra
STS	<i>Sit-to-Stand test</i>	
SVM	<i>Support Vector Machine</i>	
TCL	Transtorno Cognitivo Leve	
TMT	<i>Trial making test</i>	Teste de trilhas
TN	<i>True Negative</i>	Negativos Verdadeiros
TP	<i>True Positive</i>	Positivos Verdadeiros

TUG	<i>Timed Up and Go</i>	
UDDI	<i>Universal Description Discovery and Integration</i>	Serviço de diretório para a publicação e busca (descoberta) de serviços web
UMLS	<i>Unified Medical Language System</i>	Sistema de Linguagem Médica Única
UNED	<i>Universidad Nacional de Educación a Distancia</i>	Universidade Nacional de Educação à Distância de Madri
UUID	<i>Unique Universal Identifier</i>	Identificador Universal Único
VFT	<i>Verbal fluency test score</i>	Teste de fluência verbal
W3C	<i>World Wide Web Consortium</i>	Consórcio Internacional voltado para padronização de tecnologias Web
XMI	<i>XML Metadata Interchange</i>	Formato de Intercâmbio de Dados baseado em XML
XML	<i>Extensible Markup Language</i>	Recomendação da W3C para gerar linguagens de marcação para necessidades específicas

SUMÁRIO

Capítulo 1 - Introdução	1
1.1 A Doença de Alzheimer.....	4
1.2 Motivação	6
1.3 Objetivos.....	6
1.4 Estrutura do Texto	8
Capítulo 2 - Sistemas de Apoio à Decisão.....	9
2.1 Componentes de um Sistema de Apoio à Decisão Clínica.....	9
2.2 Sistemas baseados em Regras.....	14
2.3 Sistemas baseados em Diretrizes Clínicas.....	15
2.4 Sistemas baseados em Redes	17
2.5 Discussão	20
Capítulo 3 - Fundamentação Teórica.....	22
3.1 Teoria da Probabilidade Bayesiana e Teoria da Decisão.....	22
3.2 Redes Bayesianas.....	26
3.3 Redes de Decisão	31
3.4 Aprendizagem Bayesiana	33
Capítulo 4 - Modelagem da Rede de Decisão.....	39
4.1 Etapas da Modelagem.....	39
4.2 Critérios de Diagnóstico	41
4.3 Processo de Diagnóstico	43
4.4 Descrição dos Subconjuntos de Treinamento	44
4.5 Preparação das Bases de Casos Clínicos	47
4.5.1 Integração e Reformatação das Bases.....	47
4.5.2 Balanceamento dos Subconjuntos de Treinamento	49
4.5.3 Seleção dos Atributos	50
4.5.4 Discretização dos Atributos Numéricos	53
4.5.5 Descrição dos Atributos Seleccionados.....	55
4.6 Modelagem da Rede Bayesiana.....	61
4.6.1 Definição da Estrutura da Rede Bayesiana.....	61
4.6.2 Estimativa dos Parâmetros.....	62
4.6.3 Teste de Independência das Variáveis Aleatórias.....	69

4.7 Discussão	72
Capítulo 5 - Avaliação dos Resultados	78
5.1 Avaliação do Desempenho por Medidas de Classificação	78
5.2 Avaliação utilizando Descoberta da Estrutura da Rede.....	88
5.3 Comparação do Desempenho com o AdaBoost	92
5.4 Avaliação da Robustez da Rede Bayesiana	98
Capítulo 6 - CDSS para Diagnóstico de Demência, DA e TCL.....	103
6.1 Arquitetura do Sistema	103
6.2 Modelo de Representação de Redes Bayesianas	106
6.3 Modelo de Representação de Registros da Saúde	110
6.4 Descrição do Protótipo do CDSS e dos Módulos de Serviço.....	119
6.5 Descrição dos Aplicativos Cliente.....	124
Capítulo 7 - Conclusão	130
7.1 Contribuições	132
7.2 Trabalhos Futuros	133
Referências.....	136

CAPÍTULO 1 - INTRODUÇÃO

A preocupação quanto à qualidade, segurança e eficiência dos serviços de saúde é crescente. Duas publicações do Instituto de Medicina dos Estados Unidos, “*To Err is Human: Building a Safer Health System*” (2000) e “*Crossing the Quality Chasm: A New Health System for the 21st Century*” (CORRIGAN *et al.*, 2001) confirmam esta preocupação, revelando alguns problemas da área de saúde, tais como: (1) os riscos e danos causados por erros médicos, (2) a grande diferença da qualidade de serviços da saúde praticados por diferentes instituições, e (3) o distanciamento cada vez maior das melhores práticas e rotina clínica, agravado pela enorme quantidade de informações de diversas fontes que o clínico deve processar para a tomada de decisão. Corrigan *et al.* (2001) relacionam os seis principais domínios da qualidade na prestação de serviços à saúde: (1) garantir uma assistência segura, (2) efetiva, (3) centrada no paciente, (4) oportuna, (5) eficiente e (6) equitativa.

Em relação ao diagnóstico clínico, os principais processos cognitivos que o envolvem são resolução de problemas (SCHMIDT *et al.*, 1990) e tomada de decisão (HUNINK *et al.*, 2001). Pesquisas em resolução de problemas têm descrito o raciocínio do médico durante o processo clínico em duas estratégias principais: (1) emprego de um modelo hipotético-dedutivo envolvendo, por exemplo, uma seleção inicial de hipóteses e testes, e (2) o reconhecimento de padrões ou categorização, usualmente utilizando experiência de casos anteriores. Pela teoria da decisão, chegar a um diagnóstico significa atualizar uma opinião com base em informações imperfeitas ou observações parciais (ELSTEIN e SCHWARZ, 2002).

Devido aos altos custos e riscos de um diagnóstico incorreto, melhorar sua precisão ou taxa de diagnósticos verdadeiros tem sido considerado como o próximo desafio a ser vencido na segurança dos pacientes (NEWMAN-TOKER e PRONOVOST, 2009). Erros no diagnóstico são definidos como aqueles em que o diagnóstico clínico ocorre de forma tardia ou ausente, ou incorreta e não intencional, a partir da avaliação de informações disponíveis, suficientes e definitivas do paciente (GRABER *et al.*, 2005). Um diagnóstico incorreto pode levar a um tratamento de uma condição não existente, ou um atraso do tratamento para uma condição existente. Em um estudo coordenado pela Universidade de Linköpings, Olafsdottir *et al.* (2001) demonstraram que apenas 25% dos pacientes receberam corretamente o diagnóstico de demência no primeiro atendimento. Phillips *et al.* (PHILLIPS *et al.*, 2004) analisaram 49.435 ações na justiça por imperícia médica nos Estados Unidos. Um terço destas era motivada por erros de diagnóstico. Chandra *et al.* (CHANDRA *et al.*, 2005) avaliaram que

40% das ações judiciais nos Estados Unidos em 2003 eram motivadas por erros de diagnóstico, representando custos médios de US\$300.000 por ação ao sistema de saúde. Graber *et al.* (2005) estudaram os erros de diagnóstico, classificando-os em quatro categorias, (1) erros sistêmicos, (2) erros cognitivos, (3) erros sistêmicos e cognitivos, (4) outras causas de erros, e concluíram que a maior parte (46%) envolvia erros sistêmicos e cognitivos. Tejerina *et al.* (2012) analisaram os relatórios de autópsia de 866 indivíduos, e compararam com o diagnóstico original: 7,5% dos indivíduos obtiveram erros de diagnóstico do tipo I (falso positivo), e 11,4% dos indivíduos com erros do tipo II (falso negativo). Singh *et al.* (2012) selecionaram 254 casos de pacientes diagnosticados com câncer de pulmão e câncer colorretal. Deste conjunto, constataram erros de diagnóstico em 32,6% dos casos de câncer de pulmão e 33,5% dos casos de câncer colorretal. Schiff *et al.* (2009) avaliaram 583 casos relatados por 310 clínicos de 22 instituições, verificando erros de diagnóstico classificados como importantes em 28% dos casos. Berner e Graber (2008) identificaram uma relação de 10% a 15% de erros de diagnóstico em relação ao conjunto de casos clínicos avaliados.

Embora sejam de fácil prevenção, os erros de diagnóstico podem trazer consequências graves à saúde do paciente (NEWMAN-TOKER e PRONOVOST, 2009). Alguns sistemas computacionais tais como, sistemas de informações médicas e sistemas de apoio à decisão clínica, podem ser utilizados para reduzir a probabilidade de erros no diagnóstico clínico, melhorando a segurança do paciente e qualidade dos serviços da área de saúde (MILLER, R.A., 2009; TEICH *et al.*, 2005).

Sistemas de apoio à decisão clínica, conhecidos pela sigla CDSS (*Clinical Decision Support System*), são sistemas computacionais para ajudar os médicos na tomada de decisão no momento do atendimento médico (BERNER, ETA S., 2007).

Alguns trabalhos mostram uma redução da quantidade de erros médicos após a implantação de CDSS (HAYNES e WILCZYNSKI, 2010). Garg *et al.* (2005) levantaram o resultado de 97 CDSSs, comprovando uma melhora do atendimento em 64% dos casos. Kawamoto *et al.* (2005) acompanharam a implantação de 70 CDSSs no ambiente clínico, mostrando uma melhoria significativa da prática clínica em 68% dos casos. Kaushal *et al.* (2003) levantaram 12 casos de implantação de CDSSs, comprovando uma redução significativa da taxa de erros de prescrição médica em todos os casos.

Algumas organizações vinculadas ao Departamento de Serviços de Saúde dos Estados Unidos (*U.S. Department of Health and Human Services*) foram criadas com o objetivo de coordenar iniciativas, promover padronizações e garantir a segurança e qualidade de sistemas de informações médicas e CDSSs, tais como: (1) *American Health Information Community*,

(2) *Commission for Certification of Healthcare Information Technology*, (3) *Office of the National Coordinator for Health Information Technology*, e (4) *Agency for Healthcare Research and Quality*, este último conhecido pela sigla AHRQ (OSHEROFF *et al.*, 2007). No Brasil, até o presente momento, não há registro de nenhuma iniciativa governamental visando estabelecer princípios e padrões para o desenvolvimento e implantação de CDSSs.

Em 2005, Associação Americana de Informática Médica, ou AMIA (*American Medical Informatics Association*), propôs um plano tático para direcionar iniciativas governamentais e do setor privado no que tange ao desenvolvimento e implantação de CDSS. O plano relacionava alguns benefícios diretos e indiretos quanto ao uso de CDSS. Os principais benefícios genéricos são: (1) disponibilização do melhor conhecimento ou melhores práticas sempre que necessário; (2) utilização efetiva e disseminada da medicina baseada em evidências; e (3) melhoria contínua dos métodos e conhecimentos envolvidos no desenvolvimento do CDSS. O plano descreve também uma infraestrutura necessária ao desenvolvimento e implantação de CDSS, incluindo a adoção de um modelo interoperável de representação do conhecimento clínico, em um formato legível às pessoas e sistemas computacionais. A AHRQ também cita, entre os fatores que podem facilitar a implantação de CDSS, a adoção e uso disseminado de modelos interoperáveis de registros eletrônicos de saúde, conhecidos pela sigla EHR (*Electronic Health Records*).

Em relação aos modelos de registros eletrônicos de saúde, algumas especificações de padronização estão sendo publicadas e mantidas por instituições governamentais e do setor privado, tais como, a ISO (*International Organization for Standardization*), CEN (*European Committee for Standardization*), e padrão openEHR (BEALE, T e HEARD, 2007; CHAUDHRY *et al.*, 2006), especificado e mantido por uma comunidade de pesquisa, conhecida por Fundação openEHR, sem fins lucrativos, com colaboradores oriundos de diversas Universidades e nacionalidades. No Brasil, a Portaria 2.073, de 31 de Agosto de 2011, regulamenta o uso de padrões de interoperabilidade e informação em saúde para sistemas de informação em saúde, no âmbito do Sistema Único de Saúde (SUS), nos níveis Municipal, Distrital, Estadual e Federal, e para os sistemas privados e do setor de saúde suplementar. A Portaria estabelece o openEHR como padrão de referência de informações em saúde. Um workshop realizado em Brasília em maio de 2012 teve como objetivos principais: (1) estabelecer um plano de ação, (2) definir padrões de arquitetura e princípios gerais, e (3) definir uma estrutura de governança. O projeto foi denominado Registro Eletrônico de Saúde Nacional, ou RES-SUS-Nacional.

1.1 A DOENÇA DE ALZHEIMER

A Doença de Alzheimer (DA) foi caracterizada primeiramente pelo neuropatologista alemão Alois Alzheimer, em 1907. A DA é caracterizada por uma afecção neurodegenerativa progressiva e irreversível, causando lesões no encéfalo (JALBERT *et al.*, 2008). A DA é considerada um tipo de Demência, e seus estágios iniciais envolvem dificuldades de memória relacionadas à recuperação ou retenção de novas informações, seguidas de outras deficiências cognitivas, como a cognição espacial, reconhecimento de faces e objetos, alexia (dificuldade de leitura), funções executivas, incluindo o julgamento e solução de problemas mais complexos. Ao longo do curso da doença, DA é caracterizada por um início súbito, seguido de um declínio gradual das funções cognitivas, levando a Demência nos estágios mais avançados (MORRIS, J. C. *et al.*, 1989; WENK, 2003). A DA é considerada o principal tipo de Demência (JALBERT *et al.*, 2008). Outros tipos de Demência são: (1) demência vascular, (2) demência frontotemporal, (3) afasia progressiva primária e (4) demência com corpos de Lewy (DUBOIS *et al.*, 2007). Conforme definição do Manual Diagnóstico e Estatístico de Transtornos Mentais da Associação Psiquiátrica Americana (APA), Demência é uma síndrome caracterizada pela deterioração de múltiplas funções cognitivas, suficientemente severas para comprometer funções sociais e ocupacionais do indivíduo (APA, 2000).

O Transtorno Cognitivo Leve (TCL), em inglês *Mild Cognitive Impairment* (MCI), ou *Mild Cognitive Disorder*, é caracterizado por déficits em um ou mais domínios cognitivos, acima do esperado para a idade ou grau de escolaridade (PETERSEN, 2003). Déficits cognitivos incluem aqueles de memória episódica (habilidade de recordar eventos específicos de um tempo e lugar), funções de atenção e memória semântica. Levantamentos indicam que 10% a 30% de idosos diagnosticados com TCL convertem a DA anualmente, enquanto que a taxa de conversão de idosos cognitivamente normais é de 1% a 2% (CELSIS, 2000). Um diagnóstico de TCL aparece frequentemente associado a um estágio pré-clínico de DA.

Segundo relatório publicado pela *Alzheimer's Association* (2012), mais de 12% da população idosa dos Estados Unidos têm DA, representando a quinta posição de causas de óbitos entre a população de indivíduos maiores que 65 anos de idade. Enquanto a proporção da causa de óbitos por doenças do coração, acidente cerebrovascular e HIV recuaram 13%, 20% e 29%, respectivamente, medidos em 2008 com referência a 2000, a DA avançou 66%. Mais de 15 milhões de norteamericanos estão envolvidos de forma não remunerada com cuidados a pessoas acometidas pela DA e outras Demências, representando em 2012 um custo estimado de US\$ 200 bilhões ao sistema de saúde. Projeta-se para 2030 mais de 7,7 milhões

de idosos nos Estados Unidos com DA. A cada 68 segundos, alguém nos Estados Unidos é diagnosticado com DA. Em 2050, projeta-se a redução deste tempo para 33 segundos.

O princípio de neurodegeneração observada em casos de DA é provocado pela presença de plaquetas e emaranhados neurofibrilares em algumas regiões do encéfalo (KHACHATURIAN, 1985). Fatores genéticos, ambientais e comportamentais (modo de vida) podem predispor o indivíduo a desencadear o processo da doença. A ocorrência mais comum de DA é em idosos (indivíduos acima de 60 anos de idade), representando aproximadamente 95% dos casos da doença (WERNER *et al.*, 2009). Dado o envelhecimento da população, promovido pelos avanços da medicina e melhores condições sociais, ambientais e de higiene, DA representa um problema potencial ao sistema de saúde. Conforme relatório publicado em *Alzheimer's Association* (2012), norteamericanos maiores que 85 anos com diagnóstico de DA representam 48% da população acometida pela doença. Projeções indicam que em 2050 a população de idosos com mais de 85 anos será quatro vezes maior em relação a 2009 (MCDONALD *et al.*, 2009).

No Brasil, Herrera *et al.* (2002) realizaram um levantamento estatístico em uma área urbana de Catanduva, no Estado de São Paulo, mostrando uma taxa de prevalência de 7,1% dos casos de Demência na população de idosos, conforme critérios descritos no Manual Diagnóstico e Estatístico de Transtornos Mentais, da Associação Psiquiátrica Americana. Destes, 55% das pessoas foram diagnosticadas com Demência do tipo Alzheimer.

Em outro levantamento mais recente, Bottino *et al.* (2008) analisaram uma comunidade da cidade de São Paulo, obtendo uma taxa de prevalência de Demência de 6,9%, e DA apareceu novamente como seu tipo principal, representando 59,8% dos casos. Rodriguez *et al.* (2008) mostraram uma prevalência de Demência variando de 4,6% a 9,7%, conforme o critério de diagnóstico adotado, em áreas urbanas da América Latina. Nitrini *et al.* (2009) mostraram que este valor de prevalência é muito próximo ao de países desenvolvidos.

Um estudo longitudinal visa avaliar as variações nas características dos elementos amostrais ao longo de um determinado período. Nitrini *et al.* (2004) conduziram um levantamento longitudinal em um grupo de idosos na cidade de São Paulo, obtendo uma incidência anual média para Demência de 7,7 para um grupo de 1000 idosos. Chaves *et al.* (2009) realizaram o mesmo levantamento nas principais cidades do Rio Grande do Sul, obtendo uma incidência anual média para Demência de 14,8 para um grupo de 1000 idosos. O problema se agrava com a transição demográfica, ou envelhecimento da população. Projeções indicam que, em 2020, o Brasil será o sexto país do mundo em número de idosos, com 30 milhões de pessoas nesta faixa etária (DE CARVALHO e GARCIA, 2003).

1.2 MOTIVAÇÃO

A diferença da qualidade e eficiência dos serviços de saúde observada em diferentes instituições, a imensa quantidade de informações consideradas pelo médico durante um processo decisório, e a crescente preocupação quanto à segurança dos pacientes, conforme descrito anteriormente, são motivações desta tese para a proposta de um sistema de apoio à decisão.

Considerando as projeções sobre envelhecimento da população mundial, e em especial no Brasil, um sistema de apoio à decisão voltado para doenças neurodegenerativas é de grande valia, pois pode facilitar o diagnóstico de tais doenças e antecipar seu tratamento, proporcionando uma melhora na qualidade de vida dos pacientes e uma diminuição nos custos do sistema de saúde.

Nesta tese, propomos um sistema de apoio à decisão aplicado ao diagnóstico de Demência, Doença de Alzheimer (DA) e Transtorno Cognitivo Leve (TCL), com especial interesse no diagnóstico de TCL, pois este pode representar um estado pré-clínico de DA, e sua identificação pode antecipar ações terapêuticas.

1.3 OBJETIVOS

O principal objetivo desta tese é a proposta de um sistema de apoio à decisão (CDSS) aplicado ao diagnóstico de Demência, Doença de Alzheimer (DA) e Transtorno Cognitivo Leve (TCL). Para o projeto deste CDSS, é necessária a proposta de um modelo de decisão que reúna características apropriadas para aplicações computacionais na área de saúde, como o tratamento de incertezas, o mapeamento da causalidade entre os eventos, e a indicação dos fatores mais relevantes que levaram a distribuição de probabilidades apresentadas, assim como itens importantes ainda não observados no paciente que ajudariam na indicação do diagnóstico. As redes Bayesianas (PEARL, 1988) possuem tais características e, por esta razão, foram escolhidas para a modelagem do módulo de decisão do CDSS a ser proposto. Para a modelagem de Redes Bayesianas relacionadas ao diagnóstico das doenças de interesse do presente trabalho (Demência, DA e TCL), é necessário o estudo dos critérios de diagnóstico de cada doença, que devem estar representados no modelo proposto.

Os principais objetivos desta tese são:

- (1) Modelagem dos principais itens clínicos envolvendo o diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, considerando na modelagem formatos interoperáveis de registros eletrônicos de saúde;
- (2) Proposta de modelos de decisão para o diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, considerando os critérios clínicos de diagnóstico mais recentes e aceitos pela comunidade científica para estas doenças. O modelo de decisão deverá representar o conhecimento médico e possuir capacidade de trabalhar com observações parciais, incertezas e causalidade entre eventos, características presentes no domínio clínico, o que justifica o uso de Redes Bayesianas em sua proposta.
- (3) Validação do modelo de decisão proposto com especialistas médicos. O modelo de decisão deverá indicar para o especialista quais as informações obtidas do paciente foram mais relevantes para a predição do diagnóstico, e ainda, quais informações ainda não observadas ajudariam a identificação do diagnóstico. O modelo de decisão proposto deve ser validado utilizando bases de dados reais de pacientes e controles normais das doenças de interesse do trabalho.
- (4) Projeto da arquitetura de um CDSS aplicado a Demência, Doença de Alzheimer e Transtorno Cognitivo Leve.
- (5) Desenvolvimento de um protótipo de CDSS e validação da modelagem de decisão, aplicando uma abordagem orientada a dados.

Utilizaremos duas bases de casos clínicos de pacientes reais e controles normais para validar o modelo de decisão proposto. A primeira base foi fornecida por um consórcio envolvendo laboratórios e clínicas médicas de pesquisa de diagnóstico de DA, denominado CERAD (*Consortium to Establish a Registry for Alzheimer's Disease*). A segunda base é do Centro de Tratamento de Doença de Alzheimer (CDA), do Instituto de Psiquiatria da Universidade Federal do Rio de Janeiro (UFRJ), com apoio da Comissão Nacional de Ética em Pesquisa (CONEP) conforme projeto de pesquisa número 284/2010, aprovado pelos Comitês de Ética em Pesquisa da Universidade Federal Fluminense (UFF) e da UFRJ.

1.4 ESTRUTURA DO TEXTO

Os próximos capítulos estão organizados da seguinte forma. O Capítulo 2 define as principais características dos sistemas de apoio à decisão, aplicados à área de saúde, e apresenta os trabalhos relacionados.

O Capítulo 3 apresenta a fundamentação teórica dos principais componentes do modelo de decisão. O modelo proposto é baseado em redes Bayesianas, portanto, o capítulo aborda o Teorema de Bayes, redes Bayesianas, seus mecanismos de inferência e aprendizagem.

O Capítulo 4 descreve a modelagem Bayesiana e de decisão proposta, incluindo sua integração com o processo de diagnóstico das doenças de interesse deste trabalho.

O Capítulo 5 apresenta a validação dos resultados de desempenho e robustez da rede Bayesiana, considerando as duas bases de dados utilizadas, bem como sua comparação com outros classificadores conhecidos na literatura.

A especificação da arquitetura do CDSS proposto e a implementação do protótipo do sistema serão descritas no Capítulo 6.

Por fim, o Capítulo 7 resume as principais contribuições da tese e sugere alguns trabalhos futuros.

CAPÍTULO 2 - SISTEMAS DE APOIO À DECISÃO

Conforme mencionado anteriormente, instituições, como AHRQ, foram criadas com o objetivo de assegurar a qualificação e certificação de desenvolvedores de sistemas de apoio à decisão à área de saúde. Os benefícios da utilização destes sistemas na rotina clínica incluem a melhora da eficiência do diagnóstico, e o fortalecimento da necessidade de adoção de modelos de dados clínicos interoperáveis e registros eletrônicos de saúde (EHR) (GARG *et al.*, 2005; LINDGREN, 2007).

Este capítulo tem como objetivo fazer um levantamento bibliográfico sobre sistemas de apoio à decisão, comentando os principais trabalhos relacionados. A Seção 2.1 descreve a taxonomia e os principais componentes funcionais dos sistemas de apoio à decisão. As seções restantes relacionam diversos CDSSs, organizados pelo mecanismo de inferência. Os CDSS baseados em regras são relacionados na Seção 2.2. Os CDSS baseados em diretrizes clínicas são descritos na Seção 2.3 e a Seção 2.4 mostra os CDSS baseados em redes, foco desta tese. A Seção 2.5 apresenta os comentários finais do capítulo.

2.1 COMPONENTES DE UM SISTEMA DE APOIO À DECISÃO CLÍNICA

Visando organizar e categorizar os diversos sistemas de apoio à decisão por suas semelhanças funcionais e características diferenciais, diferentes autores propuseram algumas taxonomias (BERLIN *et al.*, 2006; GARG *et al.*, 2005; WANG, J. K. *et al.*, 2002). Wright *et al.* (2011) integraram diversas taxonomias aplicadas a sistemas de apoio à decisão da área de saúde. Estas funções ou facilidades aparecem relacionadas na Tabela 1. A classificação principal do CDSS especificado nesta tese é sistema especialista, conforme realçado em negrito na Tabela 1, embora ele reúna outras funções, como apresentação de dados relevantes ao contexto do paciente. Nos próximos parágrafos, iremos restringir o detalhamento das características e componentes de sistemas de apoio à decisão com aplicações ao diagnóstico clínico, ou CDSS.

Com objetivo de implantação de um CDSS na rotina clínica, Sim *et al.* (2001) relacionaram alguns requisitos importantes, como, por exemplo, o uso de registros eletrônicos de saúde e a confiabilidade da base de conhecimento. Acrescentaram ainda, além da procedência, que a confiabilidade da base de conhecimento pode estar associada à capacidade

do CDSS justificar a decisão de diagnóstico, ou seja, exibir quais foram os itens mais relevantes para o diagnóstico apresentado.

Tabela 1. Relação de funções e facilidades dos sistemas de apoio à decisão da área de saúde.

Função	Descrição das facilidades
Sistema de apoio à dosagem de medicamentos	Ajustes automáticos da dosagem médica, verificação/alerta das composições das medicações, verificação/alerta das dosagens diárias máximas, verificação/alerta dos prazos de vencimentos, indicação de listas padrão.
Sistemas de apoio à elaboração da prescrição médica	Apoio à construção de sentenças e medicações utilizadas em prescrições médicas, prescrições baseadas em indicações, prescrições conforme serviço, condição ou procedimento, protocolos de tratamento, transferência de conjunto de prescrições.
Sistemas de alertas e lembretes no local e momento do atendimento	Verificação/alerta da condição do paciente e a administração de medicamentos, verificação/alerta da administração de diferentes medicamentos, verificação/alerta de alergias apresentadas pelo paciente, planos de alertas ou lembretes, verificação/alerta de itens críticos laboratoriais, verificação/alerta de administração dupla de medicamentos, gerenciamento de problemas, suporte a recomendação do exame ao radiologista, monitoramento de estados de alto risco.
Sistemas de divulgação de informações relevantes	Recuperação de informações sensíveis ao contexto, apresentação de dados relevantes ao contexto do paciente, apresentação de resultados e testes de medicamentos, interface do usuário sensível ao contexto.
Sistemas especialistas	Apoio a prescrição de antibiótico, apoio ao diagnóstico clínico ¹ , ferramentas de gerenciamento de riscos, ferramentas de prognóstico médico, apoio a transfusão, ferramentas de suporte nutricionais, interpretação de testes laboratoriais, planejamento do tratamento, ferramentas de triagem.
Sistemas de apoio a processos	Monitoramento e rastreamento de prescrições médicas, funções de registro, cadastro e reconciliação de dados/medicamentos, aprovação de prescrições, apoio a documentação.

1 Função principal onde o sistema de apoio à decisão (CDSS) proposto nesta tese se enquadra.

Berner (2007) descreve um conjunto de recomendações ou melhores práticas genéricas para os CDSS, conforme relacionadas a seguir:

1. O vocabulário médico utilizado para o sistema de registro dos itens clínicos deve ser compatível com o do CDSS.
2. As faixas de valores considerados normais com base nos resultados de testes laboratoriais devem ser compatíveis com as recomendações clínicas reconhecidas pela medicina e comunidade científica.
3. É importante assegurar que os usuários (médicos/clínicos) do CDSS estejam devidamente treinados. Os usuários de um sistema devem reconhecer as limitações do CDSS. Dependendo da qualidade e integridade das informações registradas, o CDSS pode apresentar conclusões e recomendações não apropriadas (HSIEH *et al.*, 2004). O médico necessita ter o conhecimento suficiente para interpretar as informações

divulgadas pelo sistema e julgar se aceita ou não as recomendações do sistema de apoio à decisão (BERNER, E. S. *et al.*, 2003).

4. É recomendável um acompanhamento da utilização do CDSS após sua implantação. Galanter *et al.* (2002) ilustraram uma situação em que alertas e notificações disparados por um CDSS estavam sendo ignorados por médicos e clínicos pela falta de atualização da base de conhecimento. O envolvimento de médicos e clínicos no teste do sistema e o monitoramento do seu uso são essenciais para a operação efetiva do CDSS.
5. Assegurar que a base de conhecimento seja mantida atualizada. Antes da implantação de um CDSS, um item muito importante envolve a definição das responsabilidades pela atualização da base de conhecimento, dadas as novas doenças, novas evidências, ou novos conhecimentos publicados e agregados à medicina.

A relação de componentes funcionais varia conforme a especialidade do CDSS. Os principais componentes do CDSS proposto nesta tese estão mostrados na Figura 1. Estes são: base do conhecimento (1), mecanismo de inferência (2) e interface de comunicação (3). A Figura 1 mostra também os elementos externos e atores de um CDSS, dispostos em duas colunas. Os elementos da coluna da esquerda são aqueles envolvidos na modelagem ou preparação, são eles: módulo de aquisição do conhecimento (4), especialistas do domínio do conhecimento (5), base de casos clínicos (6), analista de sistemas (7). Os elementos da coluna da direita são aqueles envolvidos na rotina clínica, são eles: repositório de registros eletrônicos de saúde do paciente (8) e o médico responsável pelo atendimento ao paciente (9).

As características clínicas utilizadas no modelo de decisão de diagnóstico das doenças são levantadas avaliando os critérios reconhecidos pela comunidade médica-científica, por especialistas nas doenças (5) e o conjunto de metadados ou atributos que compõe a base de dados clínicos (6). A base de dados clínicos será utilizada na aprendizagem supervisionada dos parâmetros do modelo de decisão. O analista de sistemas (7) valida o modelo de decisão aplicando medidas de desempenho. As características clínicas são representadas e armazenadas na base do conhecimento (1). O módulo de aquisição do conhecimento (4) contém o modelo de representação da base do conhecimento. O objetivo do módulo de aquisição do conhecimento é possibilitar a interação do especialista com os critérios de diagnóstico registrados na base do conhecimento, eventualmente atualizando, refinando ou adicionando novas características clínicas, utilizando para isto um modelo de representação do conhecimento.

No momento do atendimento, o médico (9) resgata os dados clínicos do paciente, consultando o repositório de registros eletrônicos de saúde (EHR) (8). Os registros eletrônicos de saúde são formados por unidades atômicas de informação, contendo dados demográficos, clínicos, resultados de avaliações, tratamentos planejados e ações executadas, centradas no paciente, de forma persistente, abrangendo o histórico longitudinal de longo prazo do paciente (KALRA, DIPAK e INGRAM, 2006). Uma estrutura de dados persistente é uma estrutura de dados que sempre preserva a versão anterior da própria quando é modificada (FIAT e KAPLAN, 2001). O mecanismo de inferência (2), ao receber os registros eletrônicos de saúde, traduzidos por evidências no modelo de decisão, calcula a indicação de diagnóstico de uma doença, associada a uma incerteza. Além da indicação do diagnóstico, o mecanismo de inferência avalia quais informações mais relevantes necessitam ser coletadas do paciente, visando confirmar ou refutar uma hipótese inicial do diagnóstico. Estas informações, oriundas do mecanismo de inferência, são disponibilizadas para o clínico através de uma interface de comunicação (3).

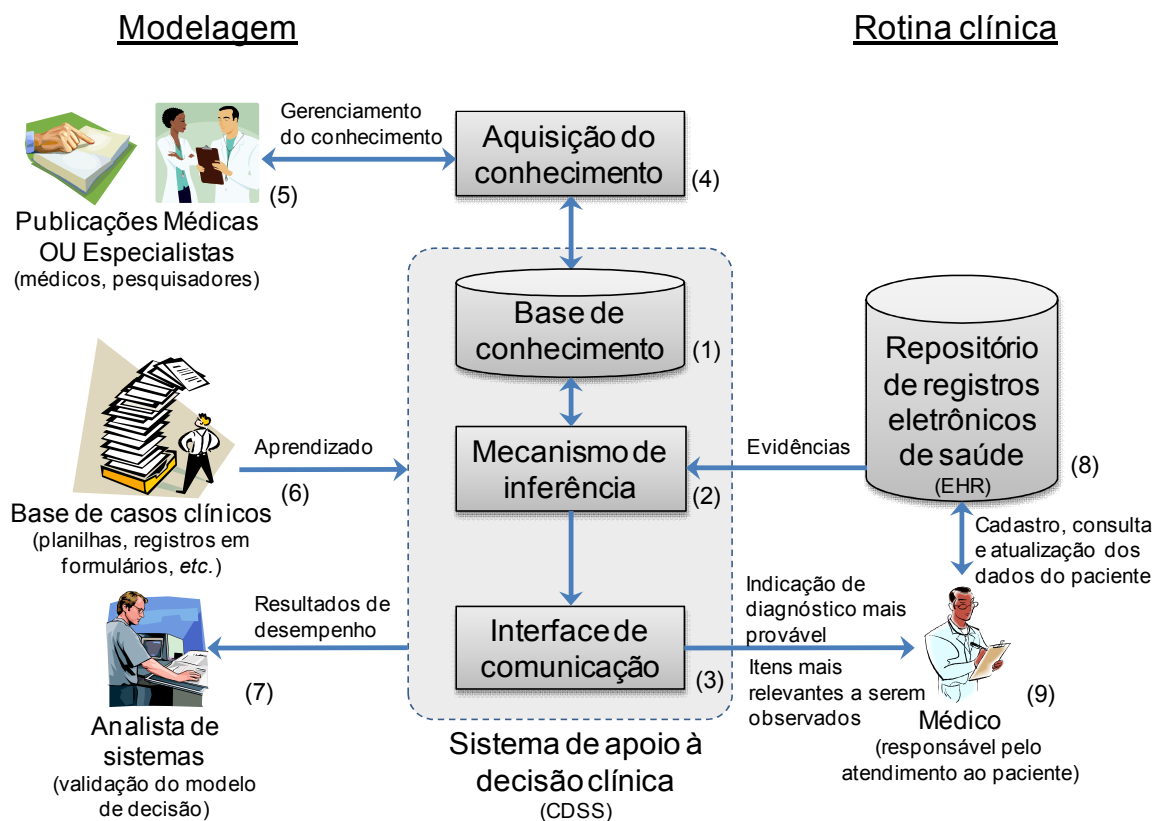


Figura 1. Componentes funcionais de um sistema de apoio à decisão clínica.

Os CDSS podem ser organizados pelo princípio do mecanismo de inferência, podendo ser: (1) sistemas baseados em regras, (2) baseados em diretrizes ou orientações de práticas

clínicas, conhecidas pelos acrônimos CGP (*Clinical Guideline Practices*) ou CIG (*Computer Interpretable Guidelines*), e (3) baseados em grafos ou redes semânticas (KONG *et al.*, 2008). A Tabela 2 relaciona os principais CDSS encontrados na literatura, agrupando-os pelo mecanismo de inferência. As próximas subseções discutem esses trabalhos.

Tabela 2. Relação de sistemas de apoio à decisão clínica por mecanismo de inferência.

Mecanismo de inferência principal	Nome	Doença(s) ou domínio de atuação	Referência
Sistemas baseados em regras	CASNET	Glaucoma	Weiss <i>et al.</i> (1976)
	Internist-I	572 doenças e 4000 sintomas	Miller <i>et al.</i> (1982)
	Mycin	Bactérias responsáveis por infecções (meningite)	Buchanan e Shortliffe (1984)
	sem nome	Diagnóstico de 257 doenças psiquiátricas	Amaral <i>et al.</i> (1995)
	CPOL	Diabetes, doenças cardíacas, cuidados da terceira idade, doenças do pulmão	Beliakov e Warren (2001)
	sem nome	Doença de Alzheimer utilizando classificador baseado em SVM (<i>Support Vector Machine</i>) e dados de exames SPECT	Salas-Gonzalez <i>et al.</i> (2010)
	sem nome	Doença de Alzheimer utilizando Componente Principal e dados de exames fMRI.	Tripoliti <i>et al.</i> (2011)
Sistemas baseados em diretrizes clínicas	EON	Hipertensão	Musen <i>et al.</i> (1996)
	Asbru	Diabetes, icterícia, câncer de mama	Shahar <i>et al.</i> (1998)
	GLIF	Depressão, hipercalemia e diabetes	Peleg <i>et al.</i> (2000)
	Prodigy	Doenças crônicas: asma e hipertensão	Johnson <i>et al.</i> (2000)
	Proforma	-	Fox <i>et al.</i> (2006)
	KON3	Oncologia	Ceccarelli <i>et al.</i> (2008)
Sistemas baseados em redes	QMR / Internist-I	572 doenças e 4000 sintomas	Miller <i>et al.</i> (1986)
	Mentor	Predição de retardo mental em recém-nascidos	Mani <i>et al.</i> (1997)
	Diaval	Doenças do coração	Diez <i>et al.</i> (1997)
	sem nome	Transtorno cognitivo leve, utilizando dados de exames de MRI e avaliações neuropsicológicas	Yan <i>et al.</i> (2007)
	Promedas	2000 doenças e 1000 sintomas	Wemmenhove <i>et al.</i> (2007)
	BayPAD	Embolia pulmonar	Luciani <i>et al.</i> (2007)
	PDS	Doenças relacionadas no ICD-10	Pyper <i>et al.</i> (2008)
	IMASC	Doenças do coração	Czibula <i>et al.</i> (2008)
	TiMeDDx	Diarreia infecciosa e não infecciosa	Denekamp e Peleg (2010)

2.2 SISTEMAS BASEADOS EM REGRAS

Regras são sentenças na forma “SE <condição> ENTÃO <ação>”. As regras são construídas com base em dois tipos de conhecimentos: (1) conhecimento declarativo e (2) conhecimento dos procedimentos. No conhecimento declarativo (1), o conhecimento é descrito utilizando conceitos, objetos e fatos, e a sintaxe é formada por declarações que podem assumir apenas dois estados: Falso e Verdadeiro. Estas declarações podem ser atômicas, representando a unidade de informação, ou compostas por mais de uma declaração atômica, conectadas entre si por operadores lógicos, tais como, “E”, “OU” e “NÃO”. Os procedimentos descrevem como o problema é resolvido, e seu conhecimento é representado por regras, estratégias e procedimentos propriamente ditos. Para esses sistemas, o conhecimento pode ser representado utilizando regras de produção, com mecanismo de inferência derivado de lógica de primeira ordem (BRACHMAN e LEVESQUE, 2004).

Os primeiros CDSS desenvolvidos foram sistemas baseados em regras. O Internist-I, desenvolvido pela Universidade de Pittsburgh em 1970, utiliza regras simples que associavam sintomas, sinais e avaliações clínicas às doenças (MILLER, R.A. *et al.*, 1982). Resultados negativos da avaliação desse sistema em uma rotina clínica, levaram a construção do QMR (*Quick Medical Reference*)/Internist-I (MILLER, R. A. *et al.*, 1986). O QMR foi um dos primeiros sistemas a considerar um modelo probabilístico para representação de incertezas. Em 1986, o QMR já relacionava 572 doenças a 4000 sintomas, sinais e testes de avaliação clínica. Outros exemplos de CDSS baseados em regras são: Casnet (WEISS *et al.*, 1976), desenvolvido pela Universidade de Rutgers, aplicado para o diagnóstico de glaucoma, e Mycin (BUCHANAN e SHORTLIFFE, 1984), desenvolvido na Universidade de Stanford, utilizado na identificação de algumas bactérias responsáveis por infecções, como a meningite.

Amaral *et al.* (1995) apresentaram um CDSS para área de psiquiatria e avaliação dos resultados. O sistema é baseado em regras associadas a incertezas, e inclui 30 grupos de diagnóstico de doenças psiquiátricas, classificando os pacientes em 290 categorias do DSM-IV-TR e 319 do ICD-10 (*World Health Organization's International Classification of Disease*), totalizando 1508 regras, 208 sintomas, sinais e resultados de testes, e 257 doenças. A incerteza é representada associando cada regra a um fator de certeza, ou CF (*Certainty Factor*).

O sistema CPOL (*Care Plan On-line*), desenvolvido pela Universidade do Sul da Austrália, disponibiliza de forma on-line informações sobre pacientes, e ainda fornece um plano de cuidados a saúde, dado o contexto clínico, conforme diretrizes clínicas cadastradas

no sistema denominado *SA HealthPlus*, mantido pelo *South Australian Health Commission* (BELIAKOV e WARREN, 2001). Os pacientes foram classificados em dez grupos de controle, em especial, diabetes, doenças cardíacas, cuidados da terceira idade, doenças do pulmão.

Em relação ao diagnóstico de DA, publicações científicas apontam a relevância de utilizar dados de exames baseados em imagens médicas para o diagnóstico de DA. Diversas pesquisas associam a atrofia na região hipocampal com um estado pré-clínico de DA do indivíduo (DEN HEIJER *et al.*, 2010; DEVANAND *et al.*, 2007; DUBOIS *et al.*, 2007). Alguns CDSS baseados em redes utilizam os resultados de métodos de análise dessas imagens médica para fornecer o diagnóstico. Salaz-Gonzalez *et al.* (2010) aplicaram regras em conjunto com métodos de reconhecimento de padrões em imagens SPECT (*Single Positron Emission Computer Tomography*) para distinguir pacientes com DA, de indivíduos normais. Tripoliti *et al.* (2011) aplicaram um conjunto de regras e dados obtidos de imagens fMRI (*functional Magnetic Resonance Images*) com o objetivo de classificar pacientes com Demência de indivíduos normais. Ao contrário de aquisições em SPECT, imagens fMRI são consideradas técnicas de imagem não invasivas devido a ausência de contraste radioativo, e ainda, possuem alta resolução espacial, e velocidade de resposta. Apesar dos resultados promissores apontados pelas técnicas de análise de imagens médicas e outros sinais biológicos, nesta tese, não utilizamos dados destas imagens, pois estes dados não estavam disponíveis para a grande maioria dos pacientes nas duas bases de casos clínicos utilizadas no desenvolvimento deste trabalho.

2.3 SISTEMAS BASEADOS EM DIRETRIZES CLÍNICAS

No domínio clínico, fluxos de trabalho (*workflows*) são utilizados para definir critérios e etapas em atividades de diagnóstico e tratamento de pacientes. A sequência de tarefas simples é denominada de percurso clínico (*clinical pathway*). O conjunto de percursos no contexto clínico, que seguem padrões de diagnóstico e tratamento reconhecidos, é denominado de diretriz clínica (*clinical guideline*). Diretrizes clínicas são geralmente representadas em forma de texto livre, enumerado e com ordenação temporal, visando auxiliar médicos, agentes de saúde e pacientes, sobre quais decisões, estratégias e procedimentos são mais adequados, dado o contexto clínico. A construção das diretrizes clínicas pode ser baseada em evidências clínicas, ou formulada pelo consenso de opiniões de especialistas sobre uma doença (GRIMSHAW e RUSSELL, 1993).

Em geral, um CDSS baseado em diretrizes clínicas possui três componentes: (1) um modelo de representação apropriado a diretrizes clínicas, (2) um ambiente de autoria para criação, gerenciamento e visualização das diretrizes clínicas, e (3) um método de controle da execução, considerando uma instância de diretriz clínica (PELEG, MOR *et al.*, 2003).

A especificação de diretrizes clínicas no EON, desenvolvido pela Universidade de Stanford (MUSEN *et al.*, 1996), no âmbito do projeto Sage, acontece pela combinação de tipos de mecanismos de tomada de decisão, fluxos de controle, ações, atividades e abstrações. O CDSS está implantado e operando no Departamento de Cuidados aos Militares Veteranos dos Estados Unidos (*U.S. Department of Veteran Affairs*), aplicado ao diagnóstico e controle de hipertensão.

Asbru é uma codificação para diretrizes clínicas, desenvolvida em um projeto denominado Asgaard (SHAHAR *et al.*, 1998). Asbru permite a representação de ações no tempo, e é utilizado em aplicações para diagnóstico de diabetes, icterícia, câncer de mamas, e outras doenças em desenvolvimento pelo Instituto Tecnológico de Software e Sistemas Interativos, da Universidade Tecnológica de Viena, e a Universidade Ben-Gurion em Negev, Israel.

GLIF (*Guideline Interchange Format*), desenvolvido pela *InterMed Collaboratory* (um grupo de pesquisa envolvendo diversas Instituições de pesquisa, incluindo a Universidade de Columbia, Harvard e Stanford), é um modelo de representação, cujo objetivo é o compartilhamento de diretrizes clínicas com outras Instituições (PELEG, M. *et al.*, 2000). A validação foi realizada com base na codificação de diretrizes clínicas aplicadas ao diagnóstico de depressão (CHOI *et al.*, 2007), hipercalemia (grande quantidade de potássio no sangue que gera sintomas como dificuldade de respirar e desmaio) (WANG, D., 2004), e diabetes (WANG, D. e SHORTLIFFE, 2002). O formato de codificação utilizado pelo GLIF é baseado na sintaxe Arden. A sintaxe Arden foi desenvolvida com objetivo compartilhar o conhecimento médico, codificando em forma de módulos lógicos médicos individuais, ou MLM (*Medical Logic Module*), onde cada módulo representa uma decisão médica. A sintaxe Arden foi incorporada ao HL7, Inc. (*Health Level Seven*), Organização acreditada pelo ANSI (*American National Standards Institute*) para o desenvolvimento e gerenciamento de padronizações relacionadas à área de saúde. Algumas ferramentas de autoria foram desenvolvidas para documentar MLMs, por exemplo, o software ArdenML (JUNG *et al.*, 2011).

Prodigy é um CDSS baseado em diretrizes clínicas aplicado ao diagnóstico de doenças crônicas, como, por exemplo, asma e hipertensão (JOHNSON *et al.*, 2000). O Prodigy foi

concebido primeiramente como uma extensão do sistema de registro eletrônico de pacientes, de propriedade do NHS (*National Health Service*) da Inglaterra. As diretrizes clínicas são representadas em forma de um grafo dirigido, onde cada nó é associado a três classes, representando cenário, ações e subdiretrizes. As diretrizes clínicas codificadas no Prodigy, denominadas de CKS (*Clinical Knowledge Summaries*), são mantidas por uma Organização denominada SCHIN (*Sowerby Centre for Health Informatics at Newcastle*) e, em 2010, foram aprovadas no processo de credenciamento do NHS.

Proforma é uma ontologia orientada a tarefa e um modelo de representação em grafos para expressar CGPs (*Clinical Guideline Practices – Diretrizes de Prática Clínica*), onde os nodos representam tarefas, incluindo ações clínicas, decisões, consultas, ou tarefas complexas, podendo estas ser decompostas em tarefas menores (FOX *et al.*, 2006). Algumas ferramentas comerciais de autoria e controle de execução baseadas na formalização do Proforma são Tallis[®] and Arezzo[®] (SUTTON e FOX, 2003).

Ceccarelli *et al.* (2008) apresentaram um CDSS baseado em diretrizes clínicas, denominado KON3, aplicado a câncer de mama. O escopo do sistema abrange a definição de uma ontologia para representar as características da doença, as diretrizes clínicas codificadas em GLIF e um modelo de serviço para acesso aos dados.

Alguns CDSSs representam as diretrizes clínicas utilizando padrões de mercado, como, por exemplo, o padrão de representação de *workflow* BPMN (*Business Process Modeling Notation*), desenvolvido pela WfMC (*Workflow Management Coalition*), e mantido pela *Object Management Group* (ROJO *et al.*, 2008), ou BPEL (*Business Process Execution Language*), mantido pela OASIS (*Organization for the Advancement of Structured Information Standards*) (MORRISON *et al.*, 2006). Yao e Kumar (2012) mostraram um estudo de caso onde as diretrizes clínicas aplicadas ao diagnóstico de insuficiência cardíaca são codificadas no padrão BPMN. Outros CDSS baseados em diretrizes clínicas podem ser obtidos pelo site OpenClinical¹, mantido pelo Laboratório Avançado de Computação, da Fundação de Pesquisa do Câncer da Inglaterra (*Cancer Research United Kingdom*).

2.4 SISTEMAS BASEADOS EM REDES

Uma característica importante da representação do conhecimento no domínio clínico é a incerteza. Em termos de representação e inferência, nos sistemas especialistas a incerteza é usualmente tratada através de duas abordagens: (1) pela lógica difusa (JOHN, R. I. e

¹ <http://www.openclinical.org>, visitado em Nov/2012

INNOCENT, 2005), e (2) teoria da probabilidade (MILLER, R. A. *et al.*, 1986). QMR/Internist-1 (SHWE *et al.*, 1991) foi um dos primeiros CDSS a considerar a teoria da probabilidade na representação de incertezas, modelo de representação do conhecimento mais apropriado ao domínio clínico. O sistema aplica uma estrutura genérica de rede Bayesiana, organizada inicialmente em dois níveis. No primeiro nível estão as variáveis aleatórias não condicionadas, representando as doenças de interesse. No segundo nível, as variáveis aleatórias condicionadas ao primeiro, representando os sintomas, sinais e resultados de testes. As variáveis aleatórias são binomiais, ou seja, expressam dois estados: ausência ou presença. Pradhan *et al.* (1994) propuseram uma extensão deste padrão, incluindo mais um nível para os fatores de pré-disposição, condicionando cada doença de interesse, mapeadas, agora, em um segundo nível desta estrutura genérica. Conseqüentemente, no terceiro nível ficaram as variáveis aleatórias dos sintomas, sinais e resultados de testes. Utilizando esta estrutura Bayesiana, eles mapearam 448 nós, 74 fatores de pré-disposição e 908 conexões causais. Os principais problemas deste modelo genérico foram: (1) limitação dos algoritmos de inferência Bayesiana, e (2) complexidade de construção das tabelas de probabilidades condicionais por levantamento do conhecimento dos especialistas. Para simplificar a construção das tabelas de probabilidade e facilitar o levantamento do conhecimento pelo especialista, o sistema adotou o modelo de estrutura local denominado *Noisy-OR* (DIEZ, F.J., 1993). A principal limitação do *Noisy-OR* é que as variáveis aleatórias da rede Bayesiana ficam restritas à distribuição binomial.

Mentor é um sistema baseado em rede Bayesiana e aplicado à predição de retardo mental em recém-nascidos, avaliando resultados de testes cognitivos (MANI *et al.*, 1997). O sistema foi construído pela Universidade da Califórnia em Berkeley e a Fundação Kaiser, envolvendo aproximadamente 6000 crianças com idade de 5 a 9 anos e 3000 mães. O sistema adota uma modelagem híbrida, onde a rede Bayesiana é descoberta a partir dos dados utilizando o algoritmo Bayesiano proposto por Singh e Valtorta (1995), e podada (*pruned*) por especialistas.

Diaval, desenvolvido pelo Hospital de La Princesa, em Madri, junto com o Departamento de Ciências da Computação da UNED (Universidade Nacional de Educação a Distância, Madri), é aplicado ao diagnóstico de doenças do coração (DIEZ, F. J. *et al.*, 1997). O sistema utiliza vários tipos de dados clínicos, incluindo, resultados de ecocardiogramas. O modelo de decisão foi baseado em uma estrutura Bayesiana, contendo 324 nodos e 335 relacionamentos. Dos 324 nós, 3 representavam dados pessoais (demográficos), 86 doenças

relacionadas ao coração e 235 dados clínicos (sintomas, sinais, resultados clínicos e testes de avaliação).

Promedas², desenvolvido pela Universidade Radboud em Nijmegen, Holanda, utiliza um modelo em rede Bayesiana para diagnosticar 2000 doenças e 1000 sintomas, sinais e resultados de testes, somando 8600 relacionamentos entre doenças e sintomas (WEMMENHOVE *et al.*, 2007). O diagnóstico da doença com maior probabilidade marginal é apresentado, juntamente com o grau de certeza, representado pela probabilidade. A princípio, foram aplicadas técnicas de inferência exata. Para os casos onde a inferência exata não era possível, foi aplicado um método de inferência aproximada, denominado *Loop Corrected Belief Propagation* (MONTANARI e RIZZO, 2005). Para os casos onde o erro excedia um determinado valor de limiar, foi utilizado um método de correção denominado LC (*Loop Corrected*).

BayPAD (*Bayes Pulmonary embolism Assisted Diagnosis*), desenvolvido pelo Centro de Pesquisas Clínicas Farmacológicas e Doenças Raras, em Bergamo na Itália, utiliza um modelo de decisão de rede Bayesiana para o diagnóstico de embolia pulmonar. O sistema foi aplicado em 750 pacientes em um estudo prospectivo, confirmando, ou não, o diagnóstico através de angiografia pulmonar. Para melhorar a acurácia (taxa de diagnósticos verdadeiros) do sistema, foi utilizada uma calibração do modelo, avaliada utilizando CRCM (*Cox Regression-Calibration Model*) (SPIEGELMAN *et al.*, 1997). Os resultados apontaram uma acurácia de 88,6%, e 97,2%, antes e depois da aplicação do método de calibração.

PDS (*Pediatric Decision Support System*) é um sistema desenvolvido e implantado em uma unidade de atendimento do CHEO (*Children's Hospital of Eastern Ontario*), com o objetivo de relacionar os possíveis diagnósticos, dados sintomas, sinais e resultados de testes apresentados pelo paciente (PYPER *et al.*, 2008). O modelo de classificação é baseado em rede Bayesiana, e os diagnósticos são conforme classificação internacional de doenças (ICD-10). O sistema também apresenta uma proposta de GUI (*Graphical User Interface*) para entrada dos dados do paciente.

IMASC (*Intelligent MultiAgent System for Clinical Decision Support*) é um CDSS aplicado ao diagnóstico de doenças do coração (CZIBULA *et al.*, 2008). O modelo se baseia na concepção de agente, trocando mensagens com módulos do sistema, e um classificador baseado em redes neurais artificiais. Os parâmetros da rede neural foram obtidos através de treinamento supervisionado, considerando 303 instâncias para treinamento e 123 para teste.

² <http://www.promedas.nl/online/resources/decision-model>, visitado em Ago/2012

TiMeDDx, desenvolvido pela Faculdade de Medicina e Instituto de Tecnologia em Israel, utiliza uma rede Bayesiana para apoiar os médicos no diagnóstico, e orientar quais dados devem ser coletados do paciente para confirmar ou não o diagnóstico inicial (DENEKAMP e PELEG, 2010). Este sistema foi baseado no processo de raciocínio hipotético-dedutivo aplicado no diagnóstico diferencial. Neste processo, o diagnóstico inicialmente faz a identificação da manifestação clínica principal, ou do grupo diagnóstico. Em seguida, o clínico identifica os sintomas, sinais, resultados de testes mais apropriados a esse grupo diagnóstico. O processo segue através do refinamento e identificação dos sintomas, partindo de um diagnóstico mais abstrato em direção a um diagnóstico mais específico, até que seja atingido um fator de certeza, ou de plausibilidade, apropriado. O estudo de caso adotado foi o diagnóstico de diarreia infecciosa e diarreia não infecciosa.

Yan *et al.* (2007) utilizaram um modelo baseado em rede Bayesiana para classificação de pacientes com TCL. Eles utilizaram variáveis aleatórias discretas representando intervalos numéricos de medidas de volume de estruturas do encéfalo (ex., hipocampo, tálamo), pontuações de avaliações neuropsicológicas e dados demográficos de pacientes e indivíduos normais. Os volumes foram segmentados manualmente a partir de imagens em MRI (*Magnetic Resonance Image*). A estrutura da rede Bayesiana foi descoberta utilizando um algoritmo de busca gulosa.

Uma relação mais completa de CDSSs baseados em redes pode ser obtida em Korb e Nicholson (2004), mostrando um levantamento de 25 aplicações médicas para redes Bayesianas atualizadas até 2010. As aplicações são divididas em duas categorias: aplicações no nível clínico e aplicações no nível epidemiológico. Algumas focam uma determinada condição clínica, enquanto outras são aplicadas no sensoriamento de sinais biológicos, normalmente, em pacientes internados em unidades de terapia intensiva. Algumas aplicações são autônomas (*stand-alone*), e outras estão integradas (*embedded*) em sistemas de informações médicas.

2.5 DISCUSSÃO

Este capítulo apresentou os principais componentes de um CDSS e comentou uma série de trabalhos relacionados encontrados na literatura, classificando-os em sistemas baseados em regras, sistemas baseados em diretrizes clínicas e sistemas baseados em redes.

O sistema de apoio à decisão proposto nesta tese é um sistema baseado em redes, mais especificamente em redes Bayesianas, assim como os sistemas Mentor, Diaval, Promedas,

BayPAD, e TiMeDDx, relacionados na Tabela 2 (DENEKAMP e PELEG, 2010; DIEZ, F. J. *et al.*, 1997; MANI *et al.*, 1997; SPIEGELMAN *et al.*, 1997; WEMMENHOVE *et al.*, 2007). A rede Bayesiana é bastante adequada ao contexto clínico, devido a sua habilidade de representação dos critérios e características consideradas em um processo decisório em termos de um diagrama causal. A rede Bayesiana é apropriada para modelos de decisão envolvendo tratamento para incertezas e observações parciais da realidade, características presentes no domínio clínico, pois utiliza um mecanismo de inferência probabilístico. Além disso, um diagrama causal é facilmente representado graficamente, o que facilita sua interpretação pelo especialista do domínio do conhecimento.

Em uma rede Bayesiana, pela análise da distribuição de probabilidades das variáveis aleatórias após a entrada das evidências, também denominada de probabilidade *a posteriori*, é possível verificar as evidências mais relevantes para o resultado indicado. Esta informação pode ser utilizada por especialistas para validar a robustez do modelo, ou justificar a escolha no processo de tomada de decisão. Além das informações mais relevantes, o modelo Bayesiano pode apresentar quais informações ainda não observadas seriam importantes na identificação do diagnóstico. A rede Bayesiana pode apresentar a distribuição de probabilidades de outras doenças relacionadas à investigação atual, seja pelo processo de diagnóstico clínico (ex., a doença pode ser considerada uma causa, ou subtipo, ou um estágio futuro da doença investigada), seja pela coincidência das evidências registradas na rede Bayesiana (ex., conjunto de sintomas, sinais ou resultados de testes comuns a outras doenças). Estas características são diferenciais do modelo Bayesiano em relação a outros modelos de decisão e mecanismos de inferência relacionados neste capítulo, justificando a escolha de redes Bayesianas para a proposta desta tese.

A comparação do modelo de decisão baseado em redes Bayesianas com outros modelos de decisão e classificadores será estendida no Capítulo 4.

O próximo capítulo apresenta a fundamentação teórica e formalização matemática dos principais componentes do modelo de decisão adotado nesta tese, incluindo o Teorema de Bayes, rede Bayesiana e seus mecanismos de inferência e aprendizagem.

CAPÍTULO 3 - FUNDAMENTAÇÃO TEÓRICA

A Teoria da Probabilidade está relacionada a duas abordagens. A abordagem frequentista da probabilidade preocupa-se em determinar a probabilidade de um evento ocorrer repetido n vezes em um dado experimento aleatório. A obra “*Ensaio Filosófico sobre a Probabilidade*” de Laplace (1812) e o Teorema de Bayes (1864) introduziram uma abordagem analítica à Teoria da Probabilidade, passando a representar também a formalização do julgamento intuitivo, sobre condições de incertezas.

Este capítulo faz uma breve exposição sobre a fundamentação teórica utilizada no modelo de decisão proposto e está organizado da seguinte forma. Na Seção 3.1, apresentamos os principais aspectos da Teoria da Probabilidade e inferência Bayesiana, incluindo a teoria da decisão Bayesiana. Na Seção 3.2, apresentamos os fundamentos da rede Bayesiana. A modelagem de diagramas de influência como modelos de decisão está descrita na Seção 3.3. Na Seção 3.4, descrevemos a aprendizagem Bayesiana.

3.1 TEORIA DA PROBABILIDADE BAYESIANA E TEORIA DA DECISÃO

Antes de descrever a teoria da probabilidade Bayesiana, vamos enumerar as principais leis que compõem a Teoria Geral da Probabilidade, e algumas definições:

1. Lei da convexidade: a probabilidade de um evento qualquer E , condicionado a H , é um número real no intervalo $[0;1]$.

$$0 \leq P(E | H) \leq 1 \quad (1)$$

Por convenção, temos que $P(H | H) = 1$ e $(H' | H) = 0$, o que estabelece os limites para os eventos certo e impossível, respectivamente. H' representa o complemento do conjunto H , onde $P(H) = 1 - P(H')$.

2. Lei da adição: se E_1 e E_2 são eventos exclusivos sob H , então a probabilidade da união lógica $E_1 + E_2$ é igual à soma aritmética das suas probabilidades individuais. Um conjunto de eventos E_i para $i = 1, \dots, n$ é exclusivo sob H , se $H \rightarrow E_i' + E_j'$, para qualquer $i \neq j$.

$$P(E_1 + E_2 | H) = P(E_1 | H) + P(E_2 | H) \quad (2)$$

3. Lei do produto: se E_1 e E_2 são eventos quaisquer, então a probabilidade do produto lógico $E_1 \cdot E_2$ condicionado a H , é o produto da probabilidade de E_1 condicionado a H multiplicada pela probabilidade de E_2 condicionado a E_1, H .

$$P(E_1, E_2 | H) = P(E_1 | H) \cdot P(E_2 | E_1, H) \quad (3)$$

No caso de eventos independentes, a lei do produto pode ser reescrita como:

$$P(E_1, E_2 | H) = P(E_1 | H) \cdot P(E_2 | H) \quad (4)$$

Um evento E é definido como independente de outro evento F , se a probabilidade $P(E | H)$ fica inalterada com a inclusão de informação sobre F .

Um modelo estatístico é definido como paramétrico quando na observação de uma variável aleatória x , com função distribuição de probabilidade $f(x|\theta)$, onde somente os parâmetros θ são desconhecidos e pertencem a um espaço vetor Θ de dimensão finita. O modelo é representado por:

$$x \propto f(x|\theta) = \prod_{i=1}^n f(x_i|\theta) \quad (5)$$

onde $x = (x_1, \dots, x_n)$ é o vetor de observações e θ é o conjunto de parâmetros, $\theta = (\theta_1, \dots, \theta_m)$.

Um dos objetivos da análise estatística é a inversão, ou seja, recuperar as causas a partir dos efeitos através de um conjunto de observações. Em outras palavras, ao se observar um fenômeno aleatório, métodos estatísticos permitem inferir a partir destas observações sobre θ . A modelagem estatística permite prever futuras observações, se estas forem condicionadas a θ . A partir da inversão estatística, podemos descrever a função de verossimilhança como função de θ (a princípio, desconhecida) dependendo dos valores observados em x , conforme mostrado a seguir:

$$L(\theta|x) = f(x|\theta) \quad (6)$$

Uma descrição genérica da inversão de probabilidades é dada pelo Teorema de Bayes, proposto por Thomas Bayes, em obra publicada após sua morte (1964): seja A e E eventos tal que $P(E) \neq 0$, $P(A|E)$ e $P(E|A)$ podem ser relacionados por:

$$P(A|E) = \frac{P(E|A) \cdot P(A)}{P(E|A) \cdot P(A) + P(E|A') \cdot P(A')} = \frac{P(E|A) \cdot P(A)}{P(E)} \quad (7)$$

Ou seja, o Teorema de Bayes relaciona a função verossimilhança $P(A|E)$ descrita a partir da distribuição de probabilidades $P(A)$, observado E . O Teorema de Bayes pode ser descrito na forma contínua, conforme mostrado abaixo:

$$g(y|x) = \frac{f(x|y) \cdot g(y)}{\int f(x|y) \cdot g(y) dy} \quad (8)$$

Seja a distribuição densidade probabilidade $p(\theta)$, denominada distribuição *a priori*. A inferência Bayesiana é baseada na distribuição de θ condicionada às observações em x , ou $p(\theta | x)$, denominada distribuição *a posteriori*, definida por:

$$p(\theta | x) = \frac{f(x | \theta) \cdot p(\theta)}{\int f(x | \theta) \cdot p(\theta) d\theta} = \frac{f(x | \theta) \cdot p(\theta)}{m(x)} \quad (9)$$

onde $m(x)$ representa a distribuição marginal de x , ou coeficiente de normalização. Um modelo Bayesiano é produzido de um modelo estatístico paramétrico, $f(x|\theta)$, e uma distribuição *a priori* dos parâmetros, $p(\theta)$. O conjunto de valores de θ que maximizam a função densidade probabilidade de x , $f(x|\theta)$, ou seja, a probabilidade de observar um dado valor de x , é denominado estimativa de máxima verossimilhança, ou MLE (*Maximum Likelihood Estimation*) dado pela Equação abaixo:

$$MLE = \hat{\theta} = \max_{\theta} L(\theta | x) \quad (10)$$

Pelo Teorema de Bayes, a distribuição *a posteriori* e a distribuição *a priori* são denominadas distribuições conjugadas, e devem pertencer a mesma família. A distribuição *a priori* é também denominada conjugada *a priori* da função verossimilhança (WINKLER, 1967). Reescrevendo a Equação 9 em outros termos, temos: distribuição *a posteriori* \propto função verossimilhança \times distribuição *a priori*. Se considerarmos uma distribuição discreta (ou multinomial) para os valores observados em X (uma distribuição probabilidade que descreve um evento aleatório com k possíveis resultados), a distribuição conjugada *a priori* mais adequada é a Distribuição de Dirichlet (GELMAN *et al.*, 2004). Assim, seja $X = (x_1, \dots, x_p)^n$ o número de ocorrências em cada uma de p categorias em n ensaios independentes, e $\Theta = (\theta_1, \dots, \theta_p)$ as probabilidades associadas, deseja-se fazer inferência sobre estes p parâmetros. Temos as seguintes restrições: $\sum_{i=1}^p \theta_i = 1$ e $\sum_{i=1}^p X_i = n$. Assumindo a distribuição multinomial de X com parâmetros n e θ e função de probabilidade conjunta das p contagens, então:

$$p(x | \theta) = \frac{n!}{\prod_{i=1}^p x_i!} \cdot \prod_{i=1}^p \theta_i^{x_i} \quad (11)$$

Considerando a Equação 6, a função de verossimilhança l para θ é dada por:

$$L(\theta; x) \propto \prod_{i=1}^p \theta_i^{x_i} \quad (12)$$

Assumindo uma distribuição Dirichlet com hiperparâmetros inteiros $\alpha_1, \dots, \alpha_p$, a distribuição *a posteriori* é dada por:

$$p(\theta | x) \propto \prod_{i=1}^p \theta_i^{x_i} \cdot \prod_{i=1}^p \theta_i^{\alpha_i - 1} \quad (13)$$

Ou seja, a distribuição *a posteriori* é calculada pelo produto dos parâmetros da distribuição *a priori* pela função de verossimilhança, o que simplifica a implementação computacional do mecanismo de inferência.

Pela Teoria da Decisão, a inferência Bayesiana pode ser aplicada, com descrições para os seguintes espaços:

1. Espaço dos possíveis resultados de um experimento X ;
2. Espaço dos parâmetros Θ ;
3. Espaço de possíveis decisões D .

Uma regra de decisão δ é uma função definida em X , que assume valores em D , ou seja, $\delta: X \rightarrow D$. A cada decisão δ e a cada possível valor do parâmetro θ podemos associar uma perda $L(\delta; \theta)$, assumindo valores reais, também denominada função perda. Assim, sabendo que a função utilidade $U(\delta; \theta) = -L(\delta; \theta)$, temos que a decisão ótima será aquela que maximizar a função utilidade esperada descrita na Equação 14. Esta regra é denominada regra de decisão de Bayes.

$$\max E[U(\delta)] = \sum_i P(\theta_i) \cdot U(\delta, \theta_i) \quad (14)$$

Por exemplo, um clínico deve decidir pelo diagnóstico ou não de uma determinada doença. É possível associar um parâmetro θ aos estados abaixo:

$$\theta = \begin{cases} \theta = 1 \rightarrow \text{paciente tem a doença} \\ \theta = 0 \rightarrow \text{paciente não tem a doença} \end{cases} \quad (15)$$

E as possíveis ações:

$$\delta = \begin{cases} \delta = 1 \rightarrow \text{diagnóstico positivo da doença} \\ \delta = 0 \rightarrow \text{diagnóstico negativo da doença} \end{cases} \quad (16)$$

Os valores atribuídos à função perda são subjetivos, atribuídos pelo especialista no momento da modelagem do sistema através da avaliação dos riscos e gravidade de possíveis erros cometidos. Dependendo da natureza do problema, muitos especialistas optam por atribuir valores a uma função utilidade, em vez da função perda, sabendo que $U(\delta; \theta) = -L(\delta; \theta)$. A Tabela 3 mostra valores teóricos atribuídos à função utilidade deste problema.

Tabela 3. Exemplo de valores atribuídos a uma função utilidade genérica

	O paciente tem a doença	O paciente não tem a doença
Diagnóstico positivo	700	-100
Diagnóstico negativo	-400	300

Em uma primeira leitura da Tabela 3, é possível verificar que as consequências de um diagnóstico negativo para um paciente que tem a doença trazem mais riscos que um diagnóstico positivo para o paciente que não tem a doença (-400 versus -100). A melhor situação em um processo decisório é prestar um diagnóstico positivo para o paciente que de fato tem a doença. Agora suponha que a incerteza sobre os estados é descrita por $P(\theta = 1) = \pi$, $0 < \pi < 1$, avaliada na distribuição *a posteriori* de θ .

Desta forma, obtemos:

$$\begin{aligned} R(\delta = 1) &= E[D(1; \theta)] = E[-U(1; \theta)] = 800 \cdot \pi + (1 - \pi) \cdot (-100) = 900 \cdot \pi - 100 \\ R(\delta = 0) &= E[D(0; \theta)] = E[-U(0; \theta)] = \pi \cdot (-400) + (1 - \pi) \cdot (300) = -700\pi + 300 \end{aligned} \quad (17)$$

Assim, verifica-se que $R(\delta=1) = R(\delta=0)$ para $\pi=0,4$. Ou seja, para $\pi < 0,4$, temos que $E[U(\delta=1)] > E[U(\delta=0)]$, portanto, a melhor decisão (decisão com menor risco ou maior utilidade) é pelo diagnóstico negativo.

3.2 REDES BAYESIANAS

As redes bayesianas são representadas por grafos dirigidos acíclicos, cujos nodos são variáveis aleatórias conectados por segmentos indicando uma relação de causalidade direta ou dependência entre as variáveis. A dependência entre dois nodos é quantificada por uma tabela de probabilidades condicionais. Russell e Norvig (2009) definem rede bayesiana como um grafo acíclico dirigido (DAG – *Directed Acyclic Graph*) constituído por nós, segmentos direcionados e funções de distribuição de probabilidade. Os nodos representam variáveis aleatórias, podendo ser booleanos (ex.: sim e não, verdadeiro e falso), multinomiais (ex.: estados categóricos, ou intervalos numéricos pré-definidos), ou funções contínuas (ex.: distribuição Gaussiana). Os segmentos dirigidos conectam pares de nodos, indicando relação de causalidade entre as variáveis. Cada variável aleatória tem uma função de distribuição de probabilidade que quantifica os efeitos da causalidade.

Seja $\mathbf{V} = \langle V_1, V_2, \dots, V_N \rangle$ um conjunto de N variáveis aleatórias, e $D(\mathbf{V})$ um conjunto de dados de N_r registros, contendo instâncias representadas pelo vetor $\mathbf{V} : D = \langle \mathbf{V} = \mathbf{v}_1, \dots, \mathbf{V} = \mathbf{v}_{N_r} \rangle$.

Nesta notação, em geral, letras em negrito representam um conjunto de objetos, letras em maiúsculo são variáveis aleatórias, letras em minúsculo são valores assumidos pelas variáveis aleatórias. Ainda, a notação $(X \perp Y | \mathbf{Z})$ expressa a independência da variável aleatória X de Y , dado um conjunto \mathbf{Z} . Nesta tese, assumimos apenas variáveis aleatórias discretas, ou multinomiais, ou seja, cada nó X_i é uma variável aleatória discreta, contendo r_i possíveis estados ou níveis discretos. Seja v_i^k o k -ésimo estado da variável aleatória V_i . Assim, uma rede Bayesiana M em um conjunto de variáveis aleatórias $\mathbf{V} = \langle V_1, V_2, \dots, V_N \rangle$ é um par $\langle G, \theta \rangle$, onde $G(V)$ representa uma estrutura em grafo acíclico dirigido (DAG) em V e θ são o conjunto de probabilidades condicionais $\theta = \langle \theta_{ijk} : \forall ijk \rangle$, tal que $\theta_{ijk} = P(X_i = x_i^k | P_i = p_i^j)$, onde \mathbf{P}_i^G representam o conjunto de nodos-pais de V_i em G , e p_i^j expressa o j -ésima configuração de pais de V_i . A notação (ijk) é geralmente utilizada para identificar o k -ésimo estado e a j -ésima coluna do i -ésimo nó da rede.

Uma segunda premissa assumida nesta tese é que a distribuição *a priori* é dada por uma distribuição de Dirichlet. Assim, seja N_{ijk} representa o número de vezes na base de dados que um nó X_i assume o estado k quando \mathbf{P}_i esta na j -ésima configuração, e seja α_{ijk} os hiperparâmetros da distribuição de Dirichlet correspondentes aos parâmetros da rede θ_{ijk} . Definimos um valor inicial $\alpha_{ijk} = 1$ assumindo o critério de não informação do algoritmo K2, descrito em Cooper e Herskovits (1992).

Uma terceira premissa assumida é independência das variáveis aleatórias. Seja a estrutura de rede Bayesiana S , cada distribuição de probabilidade θ_{ij} é independente de outra distribuição $\theta_{i'j'}$, ou:

$$P(\theta | S) = \prod_{i=0}^N \prod_{j=1}^{q_i} P(\theta_{ij} | S) \quad (18)$$

Dado a premissa da independência dos parâmetros e distribuição a priori Dirichlet, uma rede de parâmetros $\hat{\theta}$ é dada por:

$$\hat{\theta}_{ijk} = \frac{\alpha_{ijk} + N_{ijk}}{\alpha_{ij} + N_{ij}} \quad (19)$$

Os parâmetros nesta configuração representam a condição de máxima probabilidade *a posteriori* (ou MAP).

A Figura 2 mostra um exemplo de estrutura Bayesiana. Utilizando a variável aleatória representada pelo nó C como referência, os nodos A e B são definidos como nós-pais ou pai de C, e os nodos D e E são definidos como nós-descendentes ou descendentes de C.

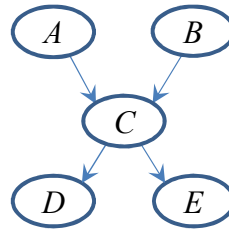


Figura 2. Grafo com segmentos acíclicos dirigidos.

A rede Bayesiana deve atender ao critério local de Markov (RUSSELL e NORVIG, 2009). O critério local de Markov é satisfeito quando cada variável aleatória é condicionalmente independente de seus não-descendentes, dados seus pais, ou:

$$X_v \perp X_{de(v)} \mid X_{P(v)} \quad \forall v \in V \quad (20)$$

onde $X_{de(v)}$ representa os descendentes de X_v e $X_{P(v)}$ os pais de X_v .

Markov blanket de um nó é formado pelo conjunto dos nodos vizinhos, incluindo os pais, os descendentes, e os pais dos seus descendentes do nó sobre análise. Cada variável aleatória deve ser condicionalmente independente de todas as outras variáveis aleatórias que pertencerem ao respectivo *Markov blanket* (GEIGER *et al.*, 1990). Trata-se de outra restrição importante de independência dos parâmetros, também denominada de *d*-separação. Assim, dois nodos X e Y são *d*-separados por um terceiro nó M , se, e somente se, para todo caminho entre X e Y , houver uma variável aleatória M “interpondo” este caminho, tal qual representado nas quatro estruturas padrões da Figura 3.

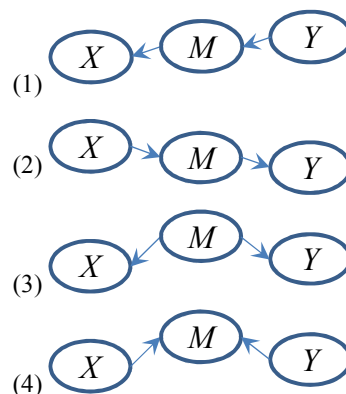


Figura 3. Quatro estruturas padrões consideradas na definição de *d*-separação entre os nodos X e Y .

A inferência Bayesiana possibilita quatro tipos básicos de raciocínio: diagnóstico (dos efeitos para as causas), causal (das causas para os efeitos), intercausal (discriminação entre as causas de um efeito comum), e misto (combinação de dois ou mais tipos acima) (RUSSELL e NORVIG, 2009), conforme ilustrado na Figura 4.

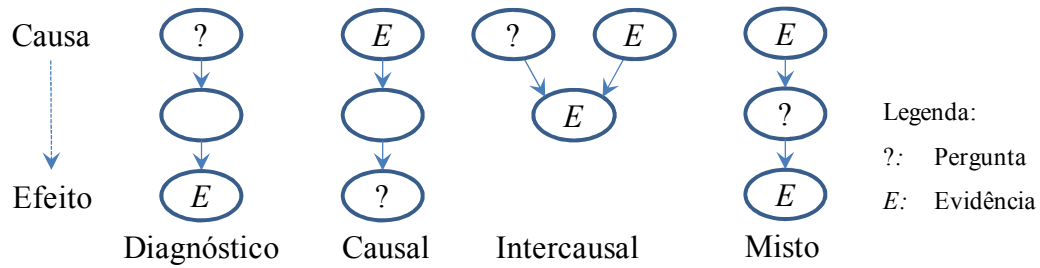


Figura 4. Diferentes tipos de raciocínio com incertezas

O objetivo da inferência Bayesiana aplicada a redes Bayesianas é estimar a distribuição *a posteriori* das variáveis aleatórias não evidenciais. A distribuição *a posteriori* é anunciada na Equação 9. Assim, seja X a variável aleatória consultada ou não evidencial (*query node*), e um conjunto de evidências E . Seja a função $Crença(X) = P(X | E)$. A tarefa da inferência Bayesiana é atualizar a função $Crença(X)$ estimando ou calculando $P(X | E)$.

Os algoritmos de inferência podem ser exatos ou aproximados. Os principais algoritmos de inferência exata são:

1. Algoritmo *polytree* ou algoritmo de passagem de mensagens: algoritmo aplicado em redes Bayesianas conectadas isoladamente, ou seja, com no máximo um caminho possível entre dois pares de nodos (variáveis aleatórias) (NEAPOLITAN, 2004; PEARL, 1988). O diagrama mostrado na Figura 5 mostra o funcionamento do algoritmo *polytree*, onde X é o nó consultado, U_i os nodos pais, Y_j os nodos filhos, Z_{ij} os pais dos filhos, e as mensagens trocadas entre os nodos com a atualização das evidências. A ideia é, a cada iteração do algoritmo, atualizar o valor de $Crença(X)$ utilizando três parâmetros: $\lambda(X)$, $\pi(X)$ e as tabelas de distribuição de probabilidades de X (CPT) com probabilidade marginal descrita por $P(X | U_1, \dots, U_m)$. Seja E^+ a parte das evidências acessíveis através dos pais de X (evidências causais) transmitidas nas mensagens π . Analogamente, E^- as evidências trazidas pelos filhos de X (evidências diagnósticas) transmitidas nas mensagens λ . Então, com $\pi(X) = P(V | E^+)$, $\lambda(X) = P(E^- | V)$, obtemos $Crença(X) = P(X | E) = P(X | E^+, E^-) = \alpha \cdot \pi(X) \cdot \lambda(X)$, sendo α um coeficiente de normalização.
2. Algoritmo de agrupamento (*clustering algorithm*): algoritmo aplicado em redes Bayesianas multiconectadas, isto significa que pelos menos dois nodos estão conectados por mais de um caminho no grafo não dirigido (LAURITZEN e SPIEGELHALTER, 1988). O algoritmo de agrupamento atua em duas etapas sequenciais: (a) compilação do grafo dirigido em outra estrutura de dados equivalente

denominada árvore de junção e, (b) atualização da função $Crença(X)$ utilizando o algoritmo *polytree* na estrutura árvore de junção. Maiores detalhes do algoritmo de transformação de uma rede Bayesiana multiconectada para uma estrutura em árvore de junção com nodos conectados isoladamente podem ser obtidos em Jensen e Nielsen (2007).

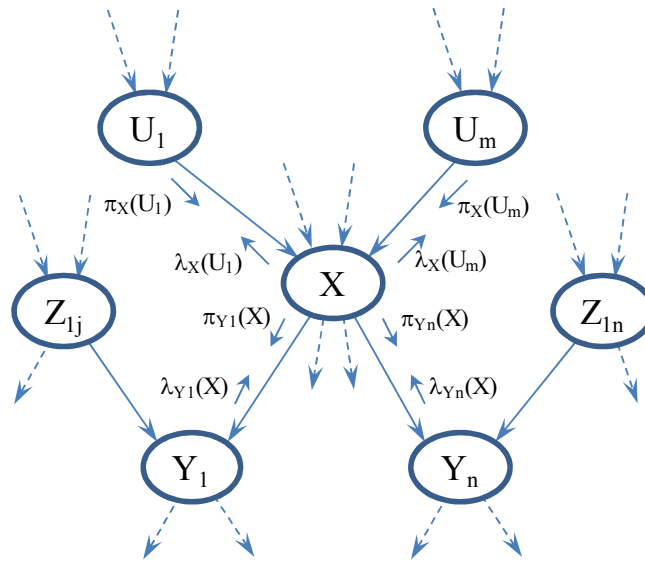


Figura 5. Inferência Bayesiana baseada em fluxo de mensagens

Os algoritmos de inferência exata são de complexidade computacional NP - Difícil (COOPER, 1990). Assim, dependendo do tamanho e complexidade da rede Bayesiana, tempo de resposta e recursos computacionais disponíveis ao mecanismo de inferência, é comum a aplicação de algoritmos de inferência aproximada utilizando simulação estocástica. Os principais são:

1. O algoritmo de amostragem lógica *Logic Sampling*, descrito em Henrion (1988), são atribuídos valores aleatórios para cada nó da rede Bayesiana, gerando uma instância do nó. Cada valor dos nodos é ponderado por sua probabilidade de ocorrência, conforme distribuição *a priori*. Visitados todos os nós, calcula-se a proporção da quantidade de nodos onde X e E são verdadeiros, pela quantidade de nodos onde apenas E é verdadeiro e estima-se a $Crença'(X)$, conforme mostrado na Equação 21. Este algoritmo é ineficiente quando a evidência E é pouco provável.

$$Crença'(X) = P'(X = x_i | E = e) = \frac{Quantidade(x_i; E = e)}{Quantidade(E = e)} \quad (21)$$

2. Fung e Chang (1990) propôs um algoritmo algumas modificações no algoritmo *Logic Sampling*. A primeira é que são geradas instâncias apenas de nodos não evidenciais. A segunda, em vez de considerar a unidade na contagem de nós, soma-se uma fração do valor de verossimilhança de cada nó. Este algoritmo foi denominado de amostragem de verossimilhança, ou *Likelihood Weighting*.

Todos os algoritmos de inferências relacionados acima são apropriados para redes Bayesianas discretas. Nesta tese, iremos utilizar o mecanismo de inferência exato, mais especificamente, o algoritmo de agrupamento. Os motivos são: (1) a estrutura da nossa rede Bayesiana é relativamente simples, não necessitando de modelos de aproximação, (2) facilidades de implementação, e (3) desempenho computacional adequado.

Há outros modelos considerados extensões ou adaptações à representação gráfica probabilística da rede Bayesiana, tais como os enumerados a seguir:

1. Modelos dinâmicos, ou DBN (*Dynamic Bayesian Networks*) (DEAN e KANAZAWA, 1989): possibilitam a representação gráfica de variáveis distribuídas em intervalos fixos de tempo, utilizando extensões de conceitos como cadeias de Markov e modelos escondidos de Markov (BAUM e PETRIE, 1966).
2. Diagramas de influência, ou ID (*Influence Diagram*) (HOWARD e MATHESON, 2005): possibilitam a modelagem de componentes da teoria da decisão, utilizando a mesma representação gráfica, de forma integrada às redes Bayesianas.

Na próxima seção, iremos detalhar as redes de decisão utilizando diagramas de influência.

3.3 REDES DE DECISÃO

A análise da decisão é um conjunto de teorias axiomáticas baseadas na escolha racional que visam modelar um processo decisório (COWELL *et al.*, 2007). Trata-se de um conjunto de ferramentas empíricas e quantitativas, utilizado para modelar problemas envolvendo tomada de decisão, incorporando preferências do paciente e do especialista com base na literatura disponível, permitindo quantificar incertezas, descobrir fatores críticos e principais informações para reduzir a incerteza. A análise da decisão, junto com a teoria da probabilidade, pode fornecer uma visão da decisão, principalmente em situações clínicas mais complexas (NEASE e OWENS, 1997).

A generalização de redes Bayesianas para representar e resolver problemas de decisão, em condições de incerteza, é denominada de diagrama de influência (HOWARD e

MATHESON, 2005). O diagrama de influência fornece uma representação explícita das ações tomadas em dado momento, e uma função utilidade para medir os possíveis resultados decorrentes desta ação. Pela teoria da decisão, a função utilidade é um elemento integrante e essencial aos problemas de decisão, e fornece uma medida da preferência através do mapeamento dos possíveis resultados de um processo de decisão sobre um conjunto de números reais. A definição da função utilidade é subjetiva, usualmente definida pelo especialista no momento da modelagem do diagrama de influência.

O nó utilidade é representado no diagrama de influência por um losango, e geralmente possui associação direta com os nodos de decisão, as variáveis aleatórias ou nodos da rede Bayesiana, e uma matriz de números reais, quantificando a utilidade ou preferência da combinação da decisão com os estados da variável aleatória associada (estrutura similar à da Tabela 3). Conforme exposto na Seção 3.1, pela teoria da decisão Bayesiana, a decisão ótima será aquela que maximizar a função utilidade esperada (Equação 14). Os nodos de decisão são representados por retângulos. Eles são variáveis de controle do tomador de decisão e apresentam todas as possíveis decisões. Um nó de decisão pode ter como nó-pai outro nó de decisão, representando uma sequência de decisões.

A Figura 6 apresenta um exemplo genérico de diagrama de influência aplicado em decisões clínicas. Os segmentos direcionados representam a relação de causalidade, e as tabelas mostram os valores de probabilidades designados para cada nó de probabilidade. O diagrama representa um processo genérico de tomada de decisão entre tomar ou não aspirina em uma situação de febre provocada por um resfriado. A ação é representada pelo nó “Tomar aspirina” (*TA*). A febre (*F*) é condicionada ao fato de o paciente estar com resfriado (*Resfr*). Os sintomas da febre podem ser observados ou identificados através de um termômetro (*Term*). A febre pode diminuir naturalmente ou se manter prolongada (*FP*). A ação representada por *TA* (Tomar aspirina) deve influenciar a febre prolongada e a reação do organismo à medicação (*R*). A função utilidade irá mapear a utilidade da febre prolongada e reação do organismo. Pela leitura da tabela associada à função utilidade, é possível concluir que o especialista designou a melhor decisão ou preferência àquela que representa o estado “Falso” (*F*) à febre prolongada e o estado “Não” para reação à medicação, conforme realçado em negrito na figura.

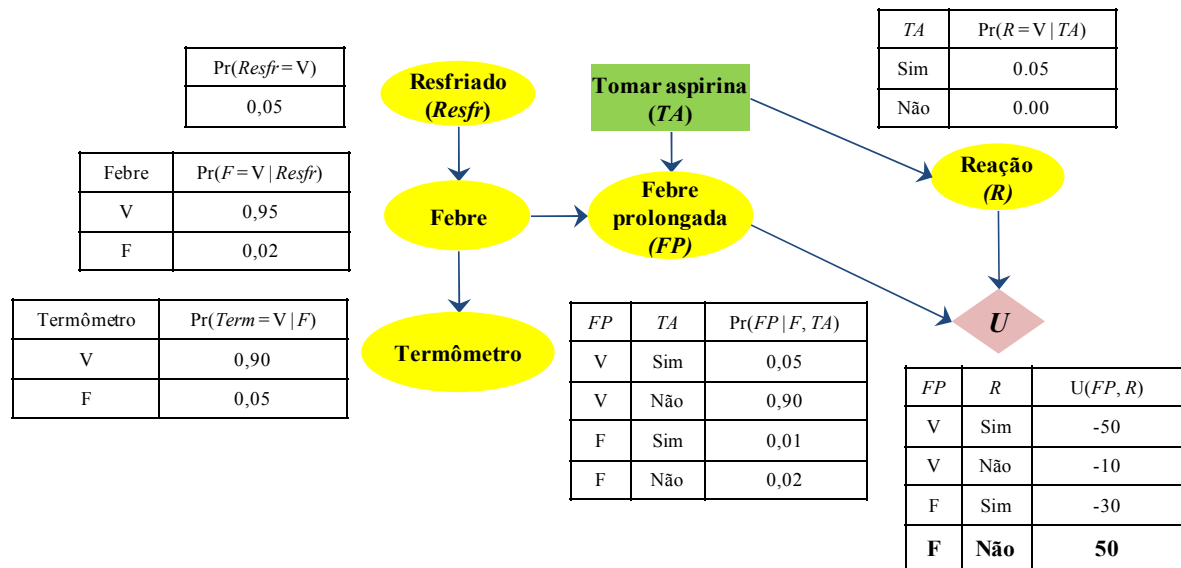


Figura 6. Exemplo genérico de diagrama de influência aplicado a uma decisão clínica. Adaptado de Korb e Nicholson (2004).

3.4 APRENDIZAGEM BAYESIANA

Na modelagem guiada por dados, aplicam-se métodos de aprendizagem, visando identificar a rede Bayesiana que fornece a maior probabilidade marginal *a posteriori* para a variável aleatória de interesse (e estado), condicionando o conjunto de dados (COOPER e HERSKOVITS, 1992). Os métodos de aprendizagem requerem um conjunto de dados retrospectivos formado por instâncias e atributos.

Basicamente, a atividade de modelagem de uma rede Bayesiana possui um componente qualitativo considerado durante o levantamento das variáveis aleatórias para o problema modelado, incluindo suas relações causais, e um componente quantitativo formado pela distribuição de probabilidades condicionais e não condicionais. Na maioria das vezes, a estrutura é definida pelo especialista e os parâmetros correspondentes são estimados através um método de aprendizagem Bayesiana utilizando um conjunto de dados de treinamento. A aprendizagem Bayesiana é considerada um problema de complexidade computacional NP - Difícil, abrangendo a aprendizagem da estrutura e parâmetros Bayesianos, também denominada descoberta da rede Bayesiana, ou aprendizagem apenas dos parâmetros. Os algoritmos de descoberta da rede Bayesiana são descritos por duas abordagens: algoritmos de descoberta que utilizam (1) testes de independência, e os que utilizam (2) métodos de busca e pontuação.

As redes Bayesianas são construídas baseadas em relações de independência, conforme descrito na Seção 3.2. Entre os algoritmos de descoberta baseados em testes de independência, está algoritmo denominado PC (SPIRITES *et al.*, 2000). O algoritmo PC testa a independência de cada par de variáveis aleatórias. A condição assumida inicialmente é um grafo não dirigido, com segmentos sendo substituídos por segmentos dirigidos conforme resultado do teste de independência. Para o teste de independência, em geral utiliza-se a função *chi-quadrado* $(X, Y | \mathbf{Z}, \mathbf{N}_{XYZ})$, onde \mathbf{N}_{XYZ} representa o número de vezes que a variável aleatória X assumiu o valor x , Y assumiu o valor y , e \mathbf{Z} a configuração \mathbf{z} na base de dados. Dash e Druzdzel (2003) descrevem uma nova função de teste de independência para o algoritmo PC, denominando teste de independência híbrido.

Um método de busca e pontuação pode ser descrito com três componentes (VISWESWARAN, 2007): (1) uma função de pontuação (score) que mede a qualidade da estrutura da rede em relação aos dados, (2) um método heurístico de busca para explorar o espaço das estruturas da rede, e (3) um estimador para parâmetros de aprendizagem da distribuição de probabilidades condicionais associadas a cada estrutura de rede.

Nos próximos parágrafos descrevemos as três medidas ou funções de pontuação mais utilizadas em algoritmos de busca para aprendizagem da rede Bayesianas. A medida de verossimilhança (FRIEDMAN, 1997) mostrada na Equação 22 é aplicada para estimar qual distribuição de parâmetros $\hat{\theta}_G$ maximiza a função verossimilhança (L) para uma estrutura de rede G , e dados D . O ponto de máxima verossimilhança é conhecido por MLE (vide Equação 10).

$$score_{ML}(G; D) = \log \left[L(\langle G; \hat{\theta}_G \rangle; D) \right] \quad (22)$$

O MDL (*Minimum Description Length*) (RISSANEN, 1978) mostrado na Equação 23 é utilizado para avaliar o tamanho da rede Bayesianas, avaliado em termos da quantidade de nodos e parâmetros, em relação ao número de atributos da base de treinamento e dados observados.

$$score_{DL}(G; D) = \frac{\dim[G]}{2} \cdot \log[N] - \log \left[L(\langle G; \hat{\theta}_G \rangle; D) \right] \quad (23)$$

onde $\dim[G]$ é o número de parâmetros independentes na rede Bayesianas e N é o número de atributos da base de dados D .

A medida conhecida por escore Bayesiano (B) (COOPER e HERSKOVITS, 1992), mostrado na Equação 24, é baseado na distribuição de probabilidade *a posteriori* de uma estrutura de rede Bayesiana G dado os dados D .

$$escor_{\mathcal{B}}(G; D) = \log[P(D|G)] + \log[P(G)] \quad (24)$$

onde o primeiro termo representa a verossimilhança marginal, calculado pela Equação 25, e o segundo é a distribuição de probabilidade *a priori*, dada uma estrutura. O ponto onde o escore Bayesiano é máximo é conhecido por MAP (*Maximum a posteriori Estimation*), citado para descrever a relação mostrada na Equação 19.

$$P(D|G) = \int_{\theta_G} P(D|\theta_G, G) \cdot P(\theta_G|G) d\theta_G \quad (25)$$

onde $P(D|\theta_G, G)$ é a verossimilhança dos dados, dado a rede Bayesiana, e $P(\theta_G|G)$ é a distribuição *a priori* sobre os possíveis valores para os parâmetros.

A

Tabela 4 apresenta quatro cenários possíveis para o problema da aprendizagem da rede Bayesiana (MURPHY, K, 2001). No primeiro cenário, o objetivo é buscar os parâmetros da rede Bayesiana (distribuição de probabilidades condicionais) que maximizam o escore de log-verossimilhança do conjunto de dados de treinamento. Assim, o conjunto de dados contém m casos independentes, $\Sigma = (x_1, \dots, x_m)$, onde $x_i = (x_{i1}, \dots, x_{in})^T$, e o conjunto de parâmetros $\Theta = (\theta_1, \dots, \theta_n)$, onde θ_i é um vetor de verossimilhança para a distribuição condicional da variável X_i , e o escore obtido pela função log-verossimilhança do conjunto de treinamento pela soma de cada nó, temos:

$$\log L(\Theta|\Sigma) = \sum \sum \log P(x_{ii} | \pi_i, \theta_i) \quad (26)$$

Tabela 4. Quatro cenários mostrando problemas de aprendizagem da rede Bayesiana.

Cenário	Estrutura da rede Bayesiana	Observações	Métodos de aprendizagem apropriados
1	Conhecida	Completa	Estimativa de máxima verossimilhança (MLE)
2	Conhecida	Parcial	EM (<i>Expectation-Maximization</i>), ou MCMC (<i>Markov Chain Monte Carlo</i>)
3	Desconhecida	Completa	Métodos de busca
4	Desconhecida	Parcial	EM com métodos de busca

A função de log-verossimilhança pode ser decomposta conforme a estrutura da rede. No primeiro cenário, o problema de aprendizagem consiste em encontrar o conjunto de parâmetros θ que maximiza a contribuição da função log-verossimilhança de cada nó de forma independente (COVER e THOMAS, 1991). Uma outra alternativa é atribuir uma função densidade probabilidade *a priori* para cada vetor de parâmetros e utilizar os dados de treinamento para maximizar a distribuição *a posteriori* dos parâmetros θ . Tal método é denominado “Tamanho de Amostra Equivalente” (ESS – *Equivalent Sample Size*) (COOPER e HERSKOVITS, 1992).

No segundo cenário (estrutura conhecida e observações parciais), o método de busca conhecido por algoritmo EM (*Expectation-Maximization*) é mais adequado para estimar a distribuição de parâmetros θ que maximiza o escore de log-verossimilhança ($score_{ML}$), ou MLE (Equação 22) (DEMPSTER *et al.*, 1977; RAMONI e SEBASTIANI, 2001). Ainda neste cenário, outra abordagem é conhecida por EM baseado em MCMC (*Markov chain Monte Carlo*). A ideia deste método é obter no passo E (*Expectation*) através de uma amostra da distribuição *a posteriori*, e calcular as estimativas amostrais das características desta distribuição, utilizando técnicas de simulação iterativa baseadas em técnicas de amostragem de Monte Carlo e cadeias de Markov (NEAPOLITAN, 2004). A ideia deste método é obter uma amostra da distribuição *a posteriori* e calcular estimativas amostrais de características desta distribuição, utilizando técnicas de simulação iterativa baseadas em cadeias de Markov. Este método de aprendizagem dos parâmetros pode apresentar um custo computacional menor em relação ao EM, principalmente nos casos de redes Bayesianas mais complexas.

No processo de modelagem adotado nesta tese, consideramos uma rede Bayesiana com estrutura definida manualmente, guiada por uma estrutura genérica multinível, e distribuição de probabilidades estimada a partir da aprendizagem Bayesiana. Consideramos também uma rede Bayesiana com variáveis aleatórias discretas e uma base de treinamento incompleta. O processo de modelagem da rede Bayesiana, incluindo a estrutura Bayesiana genérica, será detalhado no Capítulo 4. Em relação à aprendizagem, o cenário mais apropriado ao nosso problema de aprendizagem é o segundo cenário apresentado na

Tabela 4 (estrutura conhecida e parâmetros desconhecidos). Utilizaremos o método EM, pois a rede Bayesiana construída durante a modelagem da rede de decisão não é considerada complexa, não justificando assim a adoção do método MCMC. Uma rede Bayesiana é considerada complexa se contiver mais de 20 variáveis aleatórias (KORB e NICHOLSON, 2004).

Concluindo, no terceiro cenário, o objetivo é descobrir a estrutura de rede Bayesiana que melhor explica os dados de treinamento. O método mais simples é assumir que todas as variáveis aleatórias são condicionalmente independentes e relacionadas a um único nó-pai, em uma estrutura denominada *Näive Bayes*. Para o quarto cenário, um método é utilizar um algoritmo de busca, como EM, e avaliar a rede utilizando o escore Bayesiano (Equação 24) e MDL (Equação 23), método também conhecido por BIC (*Bayesian Information Criterion*) (GEIGER *et al.*, 1998). Nos próximos parágrafos, descreveremos o método de aprendizagem Bayesiana EM.

Conforme mencionado anteriormente, utilizamos uma rede Bayesiana de variáveis aleatórias discretas, multinomiais, justificando a escolha pela distribuição a priori das probabilidades a distribuição de Dirichlet. O método EM utiliza a medida de máxima verossimilhança (ou ML), conforme mostrado na Equação 22. O objetivo da aprendizagem com base na medida de verossimilhança é buscar por uma hipótese h_{ML} que maximize $P(D|h_i)$, onde D representa os vetor dados, ou seja, encontrar a distribuição de parâmetros que melhor represente o conjunto de dados D . Por se tratar de um modelo estatístico paramétrico, podemos ter em D atributos com valores ausentes ou instâncias com dados incompletos, como veremos no próximo capítulo, ocorrência bastante comum no conjunto casos clínicos utilizados como base de treinamento.

O algoritmo EM é um método iterativo de otimização descrito por Dempster *et. al.* (1977) utilizado para estimar um vetor de parâmetros θ para máxima verossimilhança, considerando observações parciais. O algoritmo EM é adequado também para distribuição discreta, ou seja, cada nível discreto é representado por um valor nominal da variável aleatória.

Cada iteração é dividida em dois passos: no passo E (*Expectation*) estimam-se as frequências esperadas dos valores nominais não preenchidos dos atributos, visando completar a amostra de dados. A estimativa das frequências esperadas é feita com base na distribuição de valores nominais preenchidos nos atributos. Sabemos que cada instância é composta por valores nominais associados aos atributos. Então, um valor nominal ausente é definido como aquele que, por alguma razão, não aparece preenchido na posição do atributo daquela instância.

Assim, seja a rede Bayesiana com n nós, o vetor de parâmetros da rede θ pode ser decomposto em $\theta = (\theta_1; \dots; \theta_n)$, onde θ_i representa o conjunto de parâmetros da distribuição de probabilidades condicionais $P(X_i | X_{\pi(i)})$ associada a cada nó X_i . Cada parâmetro θ_i é

decomposto em $\theta_i = (\theta_{i1}, \dots, \theta_{iq_i})$ onde q_i é o número de níveis discretos ou valores nominais possíveis de $X_{\pi(i)}$, e θ_{ij} representa a distribuição de probabilidade condicional $P(X_i | X_{\pi(i)} = j)$.

Seja Q a função log-verossimilhança, o valor esperado é obtido pela Equação 27.

$$Q(\theta | \theta^{(t)}) = E_{D_M | D_X, \theta^{(t)}} [\log P(D_X | \theta)] \quad (27)$$

onde D_X refere-se ao subconjunto dos valores nominais preenchidos, D_M refere-se ao subconjunto de valores nominais ausentes, $\theta^{(t)}$ é vetor de parâmetros estimado na iteração t .

No passo M (*Maximization*), aplica-se o método de busca por gradiente ascendente, visando um conjunto de parâmetros θ que maximize a função de verossimilhança Q , conforme mostrado na Equação 28.

$$\theta^{(t+1)} = \max_{\theta} Q(\theta | \theta^{(t)}) \quad (28)$$

O processo segue para próxima iteração, até que um critério de parada seja atendido. Um critério de parada pode ser a diferença do valor da função de verossimilhança $Q^{(t+1)}$ em relação ao valor da iteração anterior $Q^{(t)}$, ou um número máximo de iterações seja atingido.

Este capítulo apresentou a fundamentação teórica do modelo de decisão proposto nesta tese. No próximo capítulo, iremos detalhar o processo de modelagem da rede Bayesiana e modelo de decisão aplicado ao diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve.

CAPÍTULO 4 - MODELAGEM DA REDE DE DECISÃO

Um dos objetivos deste trabalho é representar os critérios de diagnóstico de cada doença de interesse na modelagem da rede de decisão proposta. Redes Bayesianas são bastante adequadas a realizar tal representação. Como comentado anteriormente, em uma rede Bayesiana, pela análise da distribuição de probabilidades das variáveis aleatórias após a entrada das evidências, é possível verificar as evidências mais relevantes para o resultado indicado. Esta informação pode ser utilizada por especialistas para validar a robustez do modelo, ou justificar a escolha no processo de tomada de decisão. Além das informações mais relevantes, o modelo Bayesiano pode apresentar quais informações ainda não observadas seriam importantes na identificação do diagnóstico. Estas características são diferenciais do modelo Bayesiano em relação a outros mecanismos de inferência, justificando a escolha de redes Bayesianas para a proposta deste trabalho.

Este capítulo apresenta o modelo de decisão proposto nesta tese, baseado em redes Bayesianas. A Seção 4.1 descreve as etapas de modelagem da rede de decisão. A Seção 4.2 apresenta os critérios de diagnóstico das doenças em foco, Demência, Doença de Alzheimer e Transtorno Cognitivo Leve. O processo de diagnóstico é apresentado na Seção 4.3. A Seção 4.4 descreve os subconjuntos de casos clínicos utilizados para aprendizagem dos parâmetros da rede Bayesiana. Na Seção 4.5, descrevemos as atividades de preparação da base de treinamento. Na Seção 4.6, apresentamos os resultados da modelagem Bayesiana. A discussão dos resultados é realizada na Seção 4.7.

4.1 ETAPAS DA MODELAGEM

Korb e Nicholson (2004) propõem um processo genérico de construção de uma rede Bayesiana subdividido em três etapas sequenciais: (1) construção da estrutura da rede Bayesiana com identificação das variáveis aleatórias e relação causal entre as mesmas, (2) estimativa dos parâmetros, ou distribuição de probabilidades (condicionais e *a priori*), (3) validação do modelo. Eles mostram também a combinação de técnicas de levantamento e mapeamento do conhecimento do especialista, conhecidas por técnicas de elicitación, na modelagem de redes Bayesianas, identificando os problemas mais comuns de cada etapa de construção.

Lauria e Duchessi (2007) mostram o processo de modelagem da rede Bayesiana em seis etapas sequenciais após a definição do problema, abrangendo também a sua avaliação e

implantação: (1) redução da dimensão do problema, (2) redimensionamento ou discretização das variáveis aleatórias numéricas, (3) identificação das restrições semânticas, buscando identificar algum conhecimento anterior que pode influenciar a configuração da rede, (4) seleção de modelos candidatos, preferencialmente utilizando alguma métrica de desempenho, ou função de pontuação, algoritmo de busca ou subconjunto de dados, (5) comparação dos resultados com modelos de referência e (6) implantação do modelo selecionado.

Propomos um processo de modelagem da rede Bayesiana, apresentado na Figura 7, utilizando uma combinação de técnicas de elicitación (levantamento do conhecimento do especialista) e modelagem orientada por dados. Utilizamos para representação gráfica do processo a notação BPMN (*Business Process Modeling Notation*) (WHITE e MIERS, 2008).

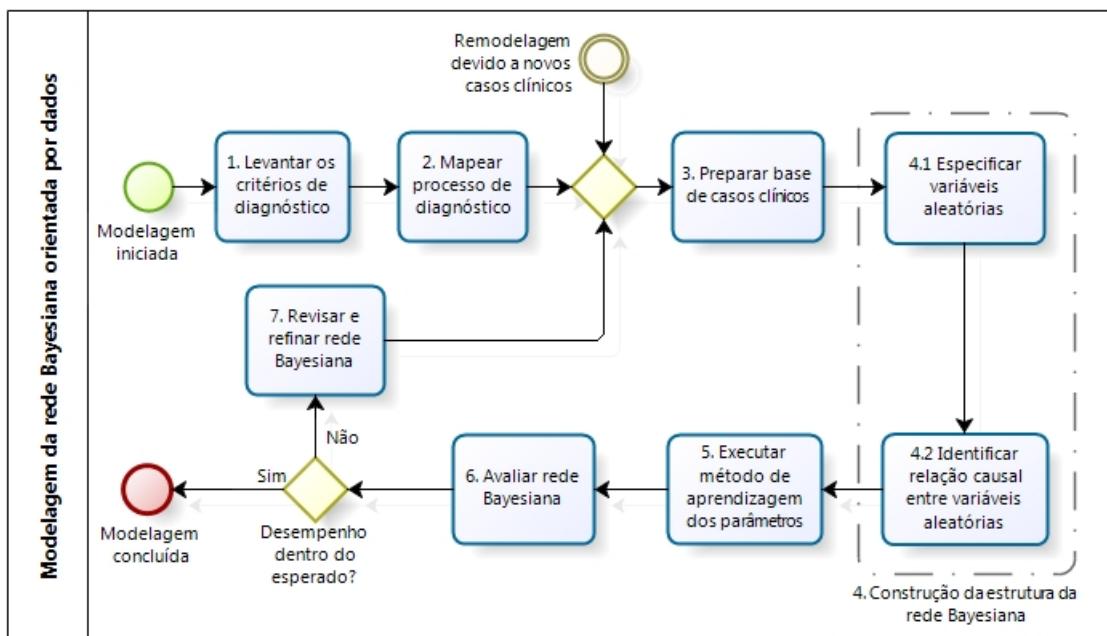


Figura 7. Processo de modelagem da rede Bayesiana orientado por dados.

Este processo é iniciado com a descrição dos critérios de diagnóstico da doença de interesse e o mapeamento do processo de diagnóstico, relacionando as principais características consideradas para tomada de decisão, e em qual o momento ou sequência estas são avaliadas ao longo do diagnóstico (Figura 7, Atividades 1 e Atividade 2).

A atividade seguinte é preparar a base de casos clínicos de modo a permitir o seu processamento por algoritmos computacionais (Figura 7, Atividade 3). A base de casos clínicos deverá conter os dados clínicos de pacientes com diagnóstico positivo e negativo da doença de interesse, preferencialmente, em quantidades próximas. Caso contrário, será necessária a execução de algum método de balanceamento da base. Nesta etapa, é possível

aplicar filtros na base de casos clínicos, visando melhorar o desempenho e robustez da rede Bayesiana. O método de balanceamento e outros filtros serão descritos nas próximas seções.

Depois de preparada a base de casos clínicos, a próxima etapa será construir a estrutura da rede Bayesiana manualmente, especificando as variáveis aleatórias e seus relacionamentos causais (Atividade 4). Para esta atividade, utilizamos uma estrutura de rede Bayesiana genérica em três níveis como guia e a ajuda de especialistas médicos considerando a doença de interesse. A rede Bayesiana proposta deve ser validada pelo especialista em sua versão final. O conjunto de parâmetros da rede Bayesiana, composto principalmente pela distribuição das probabilidades, foi estimado através de um método de aprendizagem supervisionada, utilizando os casos clínicos como base de treinamento (Atividade 5). Aplicamos medidas de desempenho e robustez para avaliar os resultados da classificação obtidos com a rede Bayesiana (Atividade 6). Consideramos uma atividade de revisão ou refinamento, caso o desempenho da rede Bayesiana esteja aquém do esperado (Atividade 7).

Depois de concluída a modelagem, e admitindo a implantação da rede Bayesiana em uma rotina clínica, deve-se considerar a necessidade de executar a rotina de modelagem novamente devido ao aparecimento de novos casos clínicos, dentro de um ciclo de melhoria contínua. Esta situação está representada no processo de modelagem por evento intermediário, indicando a remodelagem devido a novos casos clínicos.

Nas próximas seções iremos descrever cada etapa do processo de modelagem, considerando as doenças de interesse Demência, DA e TCL.

4.2 CRITÉRIOS DE DIAGNÓSTICO

Os critérios de diagnóstico de Demência mais utilizados são os publicados pela Associação Psiquiátrica Americana, na sua 4ª edição revisada, conhecidos pela sigla DSM-IV-TR (*Diagnostic and Statistical Manual of mental disorders – Fourth Edition – Text Revision*) (2000).

Os critérios de diagnóstico de DA mais utilizados são os divulgados pela NINCDS-ADRDA (*National Institute of Neurological and Communicative Disorders and Stroke - Alzheimer's Disease and Related Disorders Association*), publicado por McKhann *et al.* (1984). O NINCDS-ADRDA associa sintomas relacionados a déficits cognitivos e comportamentais ao diagnóstico de provável DA. A confirmação definitiva de DA é obtida normalmente apenas por exame histopatológico (MIRRA *et al.*, 1991; MORRIS, J. C. *et al.*, 1989) *post mortem*. Com o desenvolvimento da tecnologia e melhor entendimento etiológico

da DA, havia um consenso entre os pesquisadores sobre a necessidade de revisão destes critérios. Dubois *et al.* (2007) publicaram um revisão do critério de diagnóstico de DA, definindo as características e estágios de gravidade da DA, e incluindo marcadores biológicos e análise de imagens médicas. Em 2011, o Instituto Nacional de Envelhecimento dos Estados Unidos (NIA – *National Institute on Aging*), organismo vinculado ao Departamento de Serviços de Saúde dos Estados Unidos (*U.S. Department of Health and Human Services*), e a Associação de Alzheimer (*Alzheimer's Association*), organizaram em três publicações os critérios de diagnóstico de Demência, demência devido a DA (JACK *et al.*, 2011; MCKHANN, G. M. *et al.*, 2011) e TCL (ALBERT *et al.*, 2011), enumerando suas principais características, sinais, sintomas e técnicas mais recentes de triagem e diagnóstico destas doenças. As principais características aparecem listadas de forma resumida na Tabela 5.

Tabela 5. Principais características da Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, conforme critérios publicados pelo Instituto Nacional de Envelhecimento (Estados Unidos).

Características	D ⁽¹⁾	DA	TCL
Declínio gradual das funções cognitivas.	X	X	
Delírios.	X	X	
Déficits cognitivos relatados pelo paciente ou informante de confiança.	X	X	
Dificuldades de adquirir ou recordar novas informações.	X	X	
Dificuldades em lidar com tarefas complexas.	X	X	
Dificuldade em utilizar raciocínio visual-espacial.	X	X	
Dificuldade de utilizar elementos de linguagem.	X	X	
Mudanças de personalidade e comportamento.	X	X	
Modificações em testes de avaliação psiconeurológicas. Ex.: MMSE, CDR, Fluência Verbal, Trilhas, <i>etc.</i>	X		
Alterações nos biomarcadores do líquido cerebral: tau total, formas fosforiladas de tau, peptídeos A β 42 (formas de proteínas).		X	
Diminuição da fluordesoxiglicose medido em Tomografia por Emissão de Pósitrons (PET – <i>Positrons Emission Tomography</i>).		X	
Atrofia no volume hipocampal medido em aquisições em Ressonância Magnética do crânio (MRI – <i>Magnetic Resonance Image</i>).		X	
Imagem de perfusão em Tomografia Computadorizada por Emissão de Fóton Único (SPECT – <i>Single-Photon Emission Computed Tomography</i>).		X	
Presença de citocinas inflamatórias.			X
Estresse oxidativo (isoprostanos).			X
Histórico de DA na família.			X

(1) Doença a que a característica é associada: D=Demência, DA=Doença de Alzheimer, TCL=Transtorno Cognitivo Leve.

4.3 PROCESSO DE DIAGNÓSTICO

O processo de diagnóstico das doenças de interesse Demência, DA e TCL está apresentado em alto nível na Figura 8. Nesta notação gráfica, os pontos de decisão são representados por losangos, as atividades por retângulos, os eventos de início e término do processo por círculos. O processo começa com o atendimento aos pacientes nas unidades ou postos de saúde, sobre o regime ambulatorial ou internação. No primeiro atendimento, o agente de saúde aplica a anamnese. A anamnese consiste de uma entrevista com o paciente (autoanamnese) ou acompanhante (heteroanamnese), podendo ou não seguir um exame físico, com objetivo de estabelecer um ponto inicial no diagnóstico de uma doença. Com o resultado da anamnese, realiza-se a triagem, ou seja, o paciente é, ou não, conduzido a um grupo de atendimento mais especializado, conforme hipótese inicial de diagnóstico. No caso da suspeita de demência, aplicam-se avaliações específicas para confirmar ou refutar esta hipótese. Em caso de diagnóstico de demência negativo, verifica-se a hipótese de TCL. Em caso de confirmação da suspeita de demência, aplicam-se avaliações ainda mais específicas, buscando identificar o subtipo da demência, dentre eles, DA.

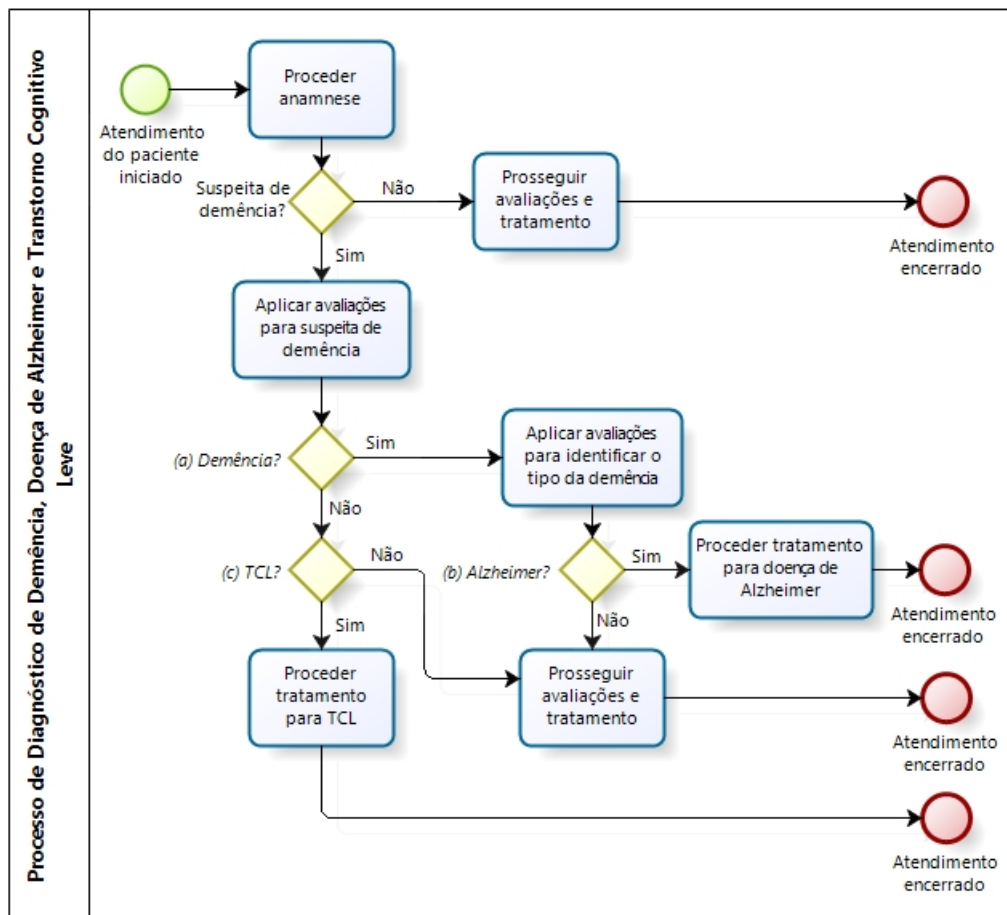


Figura 8. Processo de diagnóstico de Demência, Alzheimer (DA), Transtorno Cognitivo Leve (TCL).

O CDSS proposto nesta tese terá como função apoiar os pontos de decisão representados na Figura 8 pelas letras (a), (b) e (c).

Para cada ponto de decisão, foi modelada uma rede Bayesiana. Para cada rede Bayesiana, associamos uma função utilidade e, pelo menos, uma caixa de decisão, conforme ilustrado na Figura 9. Como vimos anteriormente, o processo de modelagem é orientado a dados, por isso, propomos uma rede Bayesiana para cada subconjunto de treinamento. Os subconjuntos de treinamento e as redes Bayesianas serão descritas nas próximas seções.

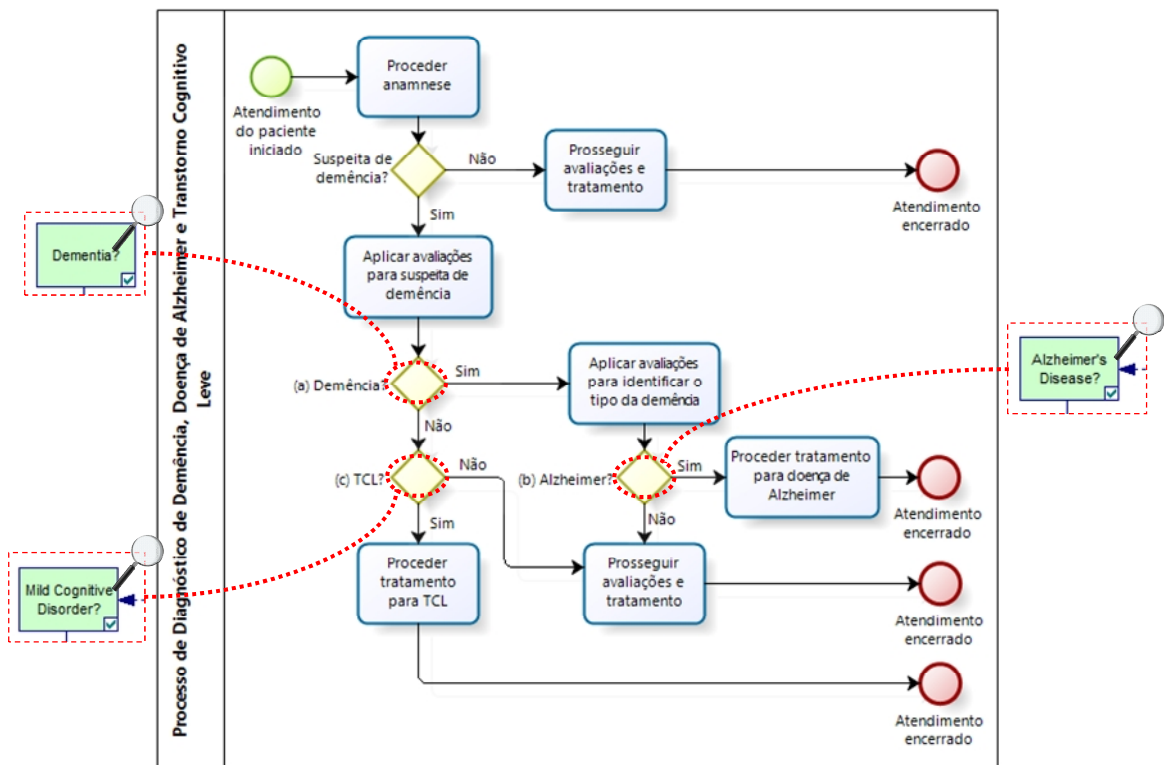


Figura 9. Representação da relação das caixas de decisão do diagrama de influência com os pontos de decisão representados no processo de diagnóstico.

4.4 DESCRIÇÃO DOS SUBCONJUNTOS DE TREINAMENTO

Conforme mencionado anteriormente, utilizamos duas bases de casos clínicos: a base do CERAD e a base do CDA. O CERAD foi fundado em 1986 com apoio do NIA (*National Institute on Aging*) visando estabelecer padronizações para avaliar e diagnosticar pacientes com DA. Participaram do projeto 24 instituições, envolvendo centros de pesquisa e universidades dos Estados Unidos. A base de casos clínicos, disponibilizada mediante solicitação, inclui dados de pacientes reais, contendo os resultados de instrumentos de avaliação, sintomas, sinais, dados de familiares, e exames histopatológicos, com a

confirmação ou não do diagnóstico de provável DA (FILLENBAUM *et al.*, 2008). A maioria publicações envolvendo a base CERAD são da área médica, e tem por objetivo estudar a sensibilidade de exames ou bateria de exames neuropsicológicos na identificação de pacientes com DA, identificar pontos de corte, ou mesmo propor alterações ou padronizações de testes utilizados na triagem, diagnóstico de Demência, DA e TCL (BECK *et al.*, 2012; PAAJANEN *et al.*, 2010; ROSSETTI *et al.*, 2010; SOTANIEMI *et al.*, 2012).

Alguns trabalhos envolvendo rede Bayesiana e a base CERAD foram conduzidos pelo Departamento de Informática Aplicada da Universidade de Fortaleza (ARAÚJO DE CASTRO *et al.*, 2008; PINHEIRO *et al.*, 2008). Castro *et al.* (2008) estudaram a análise do problema de decisão envolvendo o diagnóstico de DA com base nos dados disponíveis no CERAD. O objetivo foi identificar quais questões presentes nos exames neuropsicológicos do CERAD produzem maior impacto (maior atratividade) em relação a cada estágio de DA, determinado pelo CDR (*Clinical Dementia Rating*). O CDR é um exame neuropsicológico utilizado para avaliar o estágio de demência e possui uma alta sensibilidade para o diagnóstico de DA (MORRIS, J.C., 1997). O resultado deste teste consiste em cinco classificações representadas pelos escalares 0 a 3: 0 (nenhuma alteração), 0,5 (questionável), 1 (demência leve), 2 (demência moderada) e 3 (demência grave) (MONTAÑO e RAMOS, 2005). Uma modificação proposta pelo CERAD acrescentou mais duas classificações a este exame: 4 (demência profunda) e 5 (demência terminal). A estrutura da rede Bayesiana foi definida de maneira a representar o modelo de decisão proposto, e as probabilidades calculadas avaliando as respostas às questões dos exames. As questões foram representadas por variáveis aleatórias discretas, e as respostas associadas a valores nominais. Através da entrada das evidências na rede Bayesiana, é possível prever o valor do CDR e identificar as questões mais atrativas. Por outro lado, no trabalho realizado nesta tese, utilizamos todos os atributos e exames disponíveis na base CERAD, contudo, não consideramos as respostas às questões presentes nestes exames, apenas sua pontuação final. Ainda, a rede Bayesiana proposta nesta tese tem a função de fornecer a probabilidade de diagnóstico para apoiar um ponto de decisão, e não prever o resultado do CDR. Uma semelhança é que também fornecemos os exames mais relevantes (ou atrativos), ainda não observados, que ajudariam a identificação do diagnóstico. Mas, em vez da aplicação de um método multicritério, utilizamos informações da entropia para fazer tal indicação.

O CDA (Centro da Doença de Alzheimer) realiza o atendimento em pacientes com qualquer tipo de problema no domínio psicológico, psiquiátrico e neuropsiquiátrico, incluindo demências, depressão, ansiedade e psicoses. O CDA é vinculado ao Instituto de Psiquiatria, da

Universidade Federal do Rio de Janeiro. Os casos clínicos cedidos pelo CDA para o desenvolvimento desta tese representam dados de pacientes reais e indivíduos de controle normal registrados por médicos, alunos e colaboradores. Os atributos incluem pontuações de exames neuropsicológicos, dados clínicos, e informações demográficas do paciente (ex., idade, gênero, formação escolar). Para utilização da base de dados do CDA, foi necessária a submissão do projeto de pesquisa número 284/2010 a Comissão Nacional de Ética em Pesquisa (CONEP), e obter sua aprovação pelos Comitês de Ética em Pesquisa da UFF e da UFRJ.

Ambas as bases de casos clínicos (CERAD e CDA) foram separadas em subconjuntos aplicados a aprendizagem dos parâmetros da respectiva rede Bayesiana. Utilizamos o atributo contendo a informação de diagnóstico com valor binomial (diagnóstico positivo ou negativo), e definimos um critério de seleção para cada subconjunto, de forma coerente com o processo de diagnóstico apresentado na Figura 8. A Figura 10 mostra o critério para seleção dos subconjuntos aplicados no aprendizado dos parâmetros de cada rede Bayesiana. O conjunto total de casos de pacientes é utilizado para estimar os parâmetros da rede Bayesiana associada ao ponto de decisão *(a)*. O subconjunto de casos de pacientes com diagnóstico positivo para demência, ou outro subtipo da demência, é utilizado como base de treinamento da rede Bayesiana associada ao ponto de decisão *(b)*: se o caso se tratar de DA, classifica-se "DA positivo", e caso contrário, "DA negativo". O subconjunto de casos de pacientes com diagnóstico negativo para demência é utilizado para estimar os parâmetros da rede Bayesiana associada ao ponto de decisão *(c)*: se o caso se tratar de TCL, classifica-se "TCL positivo", e caso contrário, "TCL negativo".

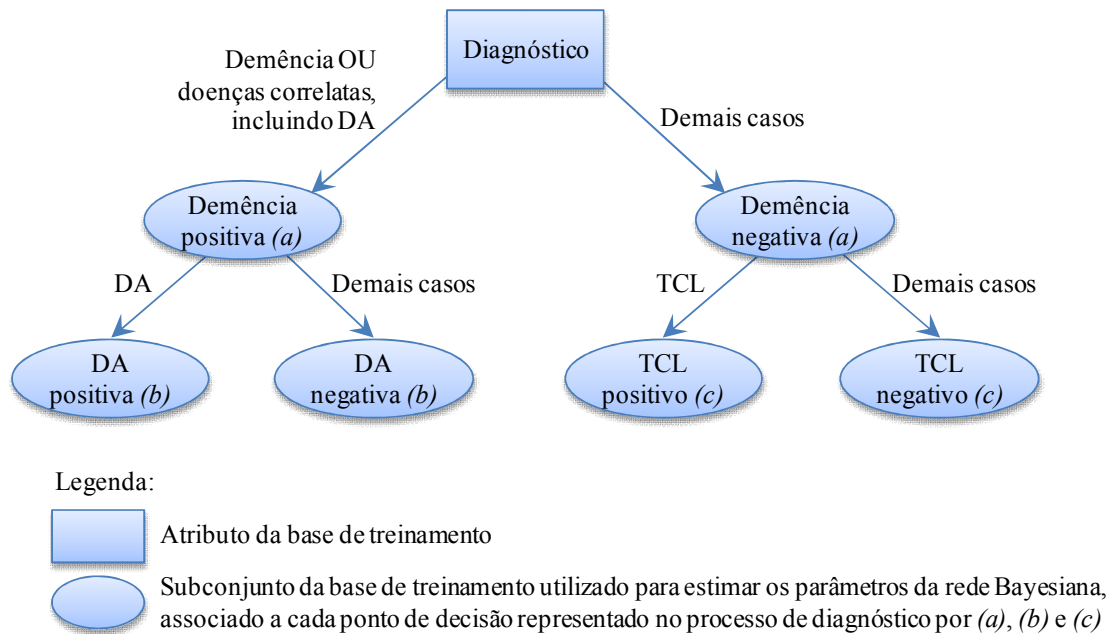


Figura 10. Critério de seleção dos casos de pacientes da base de treinamento pelos grupos de interesse.

4.5 PREPARAÇÃO DAS BASES DE CASOS CLÍNICOS

As próximas subseções descrevem as atividades de preparação das bases de treinamento (CERAD e CDA). Estas atividades foram executadas de forma independente. Visando facilitar a leitura e comparações entre as duas bases, iremos descrevê-las em uma mesma subseção.

4.5.1 INTEGRAÇÃO E REFORMATAÇÃO DAS BASES

A base CERAD estava formatada em SAS (*Statistical Analysis Software*), da SAS[®] (<http://www.sas.com/>). Uma rica documentação dos atributos e os modelos de formulários aplicados no levantamento dos dados acompanhava a base CERAD. Cada paciente é identificado por uma chave numérica presente em cada instância da base. É comum o monitoramento dos pacientes ao longo do curso da doença e período de tratamento. Assim, os dados do CERAD estão organizados por visita do paciente ao hospital ou posto de saúde (geralmente, uma vez por ano), onde alguns exames são reaplicados, e os novos resultados armazenados na base. Por uma limitação do escopo desta tese, não utilizamos estas informações temporais, tampouco modelamos características temporais na rede Bayesiana. Esta atividade foi proposta como trabalho futuro. Assim, consideramos apenas os dados coletados no primeiro ano de tratamento do paciente.

Os dados estavam distribuídos por várias tabelas codificadas no SAS. Visando facilitar a análise e processamento dos dados, integramos os dados em uma única tabela e mesclamos os atributos com semântica idêntica, mas identificados por nomes distintos. A Figura 11 mostra a distribuição de casos de pacientes com demência e controle normal ou diagnóstico negativo de demência do CERAD. Dos casos de demência, a maioria dos pacientes é classificado com o diagnóstico positivo ou negativo de DA, e alguns casos não possuem classificação. A Tabela 6 mostra a quantidade de instâncias de cada subconjunto utilizado como base de treinamento, e o percentual de dados incompletos ou ausentes do subconjunto em relação a sua quantidade total de dados. Um dado incompleto ou ausente é aquele que não possui valor disponível na base de dados.

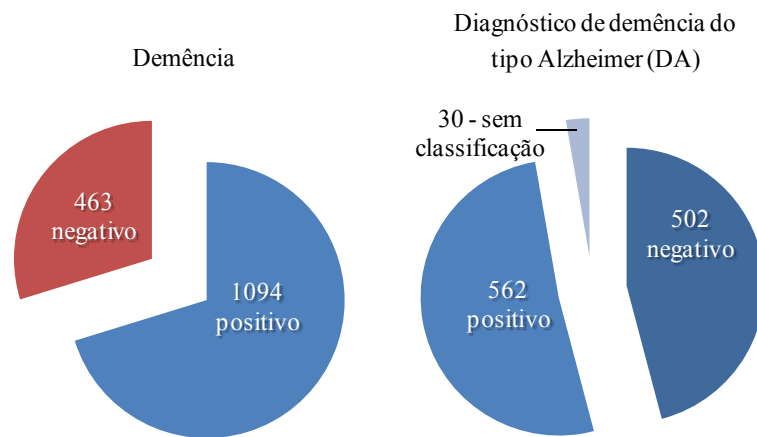


Figura 11. Quantidade de instâncias por grupo e doença da base CERAD.

Tabela 6. Quantidade de casos de pacientes por grupo e doença da base CERAD.

Doença:	Demência		DA		
	Negativo	Positivo	Negativo	Positivo	Sem classificação
Classificação ^(a) :					
Número de casos:	463	1094	502	562	30
Percentual de dados incompletos ^(b) :	48%	44%	47%	42%	-

(a) Critérios de classificação apresentados na Figura 10.

(b) Percentual de dados incompletos ou ausentes em relação ao conjunto total do subconjunto de treinamento.

Os dados clínicos dos pacientes do CDA foram encaminhados em planilhas eletrônicas, cedidas separadamente por diferentes pesquisadores. O primeiro passo foi integrar estas planilhas em uma única base de dados e mesclar atributos com mesmo significado, e instâncias referenciando o mesmo paciente. O paciente era identificado pelo seu nome completo, e, posteriormente, codificado, ocultando qualquer forma de identificação,

atendendo aos critérios de segurança e privacidade do projeto de pesquisa. Os dados foram então processados, retirando erros, ruídos, e substituindo textos livres por caracteres codificados. A Figura 12 mostra a quantidade de instâncias por subconjunto, seguindo o critério mostrado na Figura 10. Inicialmente, a base tinha 461 instâncias e 45 atributos. Retiramos as instâncias e atributos pertencentes a outros projetos de pesquisa como, por exemplo, estudos da influência da depressão e condicionamento físico na saúde de pacientes com suspeita de demência, restando 247 instâncias e 25 atributos. A Tabela 7 mostra a quantidade de instâncias de cada subconjunto utilizado como base de treinamento, e o percentual de dados incompletos ou ausentes em relação ao conjunto total de dados daquele subconjunto.

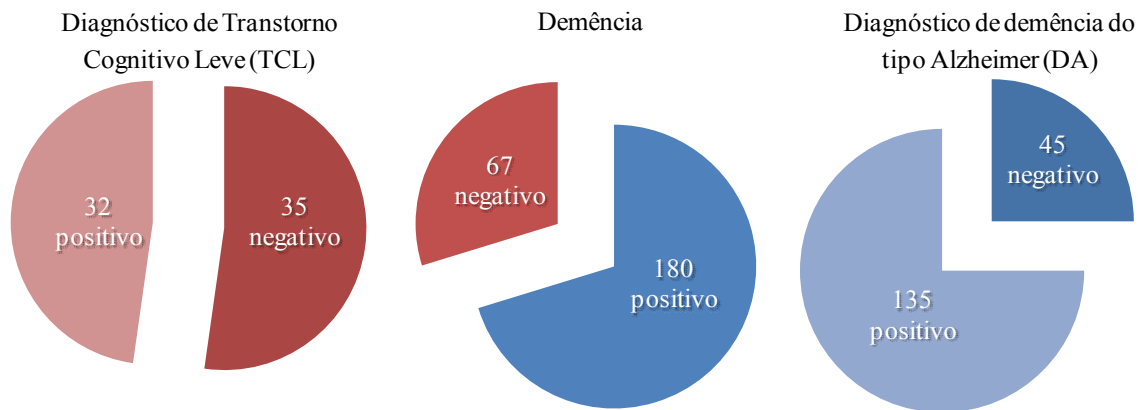


Figura 12. Quantidade de instâncias por grupo e doença da base CDA.

Tabela 7. Quantidade de casos de pacientes por grupo e doença da base CDA.

Doença:	Demência		DA		TCL	
	Negativo	Positivo	Negativo	Positivo	Negativo	Positivo
Classificação:	Negativo	Positivo	Negativo	Positivo	Negativo	Positivo
Subconjunto ^(a) :	Demência negativa	Demência positiva	Alzheimer negativo	Alzheimer positivo	TCL negativo	TCL positivo
Número de casos:	67	180	45	135	35	32
Percentual de dados incompletos ^(b) :	29%	24%	29%	22%	37%	21%

(a) Critérios de classificação apresentados na Figura 10.

(b) Percentual de dados incompletos ou ausentes em relação ao conjunto total do subconjunto de treinamento.

4.5.2 BALANCEAMENTO DOS SUBCONJUNTOS DE TREINAMENTO

Analisando as Tabelas 6 e 7, é possível notar que, o subconjunto de demência, da base CERAD, e os subconjuntos de Demência e DA, da base CDA, aparecem desbalanceados. Uma base de treinamento está desbalanceada quando a ocorrência de instâncias de uma classe

é significativamente superior que a ocorrência de instâncias da outra. Neste caso, a ocorrência em maior número é a de pacientes com diagnóstico positivo para demência (em relação aos negativos), e pacientes com diagnóstico positivo para DA (em relação aos negativos).

Diferentes técnicas para o balanceamento de base de treinamento foram propostas por diversos autores. As técnicas mais básicas consistem em subamostrar a classe majoritária (retirar um conjunto de instâncias desta classe), ou superamostrar a classe minoritária (replicar um conjunto de instâncias desta) (KUBAT e MATWIN, 1997; LING e LI, 1998). Utilizamos uma técnica de balanceamento por superamostragem conhecida por SMOTE (*Synthetic Minority Over-sampling Technique*) (CHAWLA *et al.*, 2002). Em vez replicar simplesmente instâncias da classe minoritária, esta técnica cria instâncias artificiais, formadas pela combinação linear de atributos numéricos, de forma não alterar as características globais da base de treinamento. A Figura 13 ilustra as quantidades de instâncias das bases em que o método de balanceamento foi aplicado.

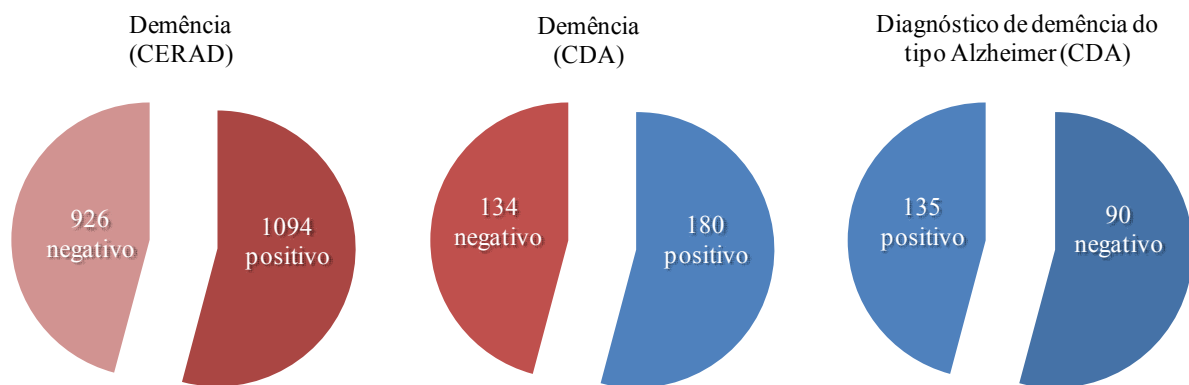


Figura 13. Quantidade de instâncias após aplicação do método de balanceamento das bases de treinamento.

4.5.3 SELEÇÃO DOS ATRIBUTOS

Para reduzir a dimensão do problema e melhorar o desempenho do classificador, aplicamos uma técnica baseada em entropia, para remover atributos de cada base com ganho de informação abaixo de um valor de limiar (MACKAY, 1992).

O ganho de informação é utilizado para medir a redução da incerteza de um evento Y , dado a observação de uma variável aleatória X (ANDERSON *et al.*, 1986). Neste caso, utilizamos a entropia H como medida de incerteza sobre uma variável aleatória, conforme mostra a Equação 29 sobre a variável aleatória X . A Figura 14 mostra o gráfico da função

entropia para $|H(X)|$, assumindo uma variável aleatória X com dois estados, por exemplo, Positivo e Negativo.

$$H(X) = -\sum_{x \in X} P(x) \cdot \log[P(x)] \quad (29)$$

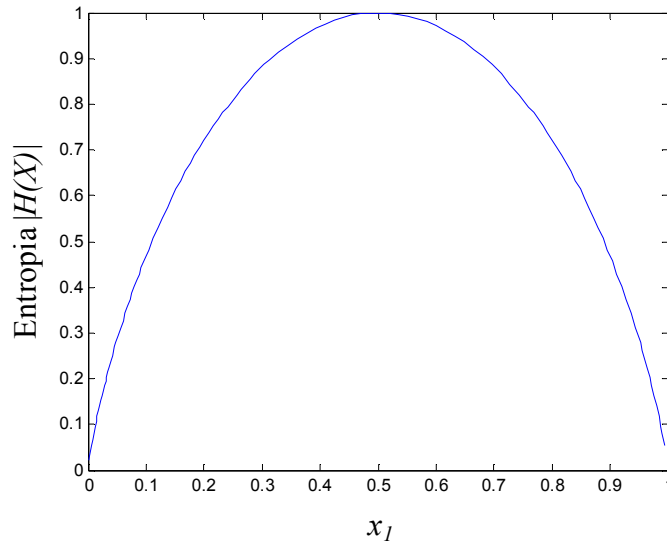


Figura 14. Função entropia $|H(X)|$ com a variável aleatória X contendo dois estados.

O ganho de informação é obtido pela redução da entropia de Y com a observação do estado da variável aleatória X , mostrado na Equação 30. Esta relação também é conhecida divergência de Kullback-Leibler (KULLBACK e LEIBLER, 1951).

$$GI(Y, X) = H(Y) - H(Y | X) \quad (30)$$

onde X representa os k atributos encontrados em cada instância da base de treinamento, com $(X, y) = (x_1, x_2, x_3, \dots, x_k, y)$, Y a classe correspondente a instância, e GI o ganho de informação. A classe é o atributo da base de treinamento que contém o diagnóstico do paciente.

Para avaliação do ganho de informação, utilizamos a implementação disponível no aplicativo WEKA, desenvolvido pela Universidade de Waikato³. Os atributos foram primeiramente reorganizados por ordem decrescente dos respectivos valores de ganho de informação, em seguida, estabelecido um valor de corte: os atributos abaixo do valor de corte foram removidos do modelo. Em geral, o valor de corte é definido por dois critérios: (1) a quantidade máxima de atributos desejada no modelo, e (2) a definição de um valor de limiar. No WEKA, valores abaixo da quinta casa decimal são desprezados. Assim, especificamos um

³ Endereço web: <http://www.cs.waikato.ac.nz/ml/weka/>

valor de limiar de 0,0001, ou seja, atributos com ganho de informação inferior a este valor de limiar foram considerados irrelevantes para o resultado em questão.

Em relação aos dados incompletos ou ausentes, o problema geral é estimar a distribuição de dados incompletos ou ausentes D mais apropriada, admitindo uma base de treinamento completa D_C .

Seja $X_i = x_{ik}$ e $\Pi_i = \pi_{ij}$, com i representando o índice do nodo, k o índice do estado discreto de x_i , e j o índice do conjunto de combinações dos valores ou estados discretos dos nodos-pais de x_{ik} , representado por π_{ij} . Seja a base de treinamento de n casos $D_C = (c_1, \dots, c_n)$. Seja também $n(x_{ik} | \pi_{ij})$ a frequência de casos em que cada variável aleatória X_i aparece na base de treinamento com estado x_{ik} , junto com a combinação π_{ij} dos nodos-pais de Π_i . Dado que D_C , então a probabilidade condicional $p(x_{ik} | \pi_{ij}) = \frac{n(x_{ik} | \pi_{ij})}{n(\pi_{ij})}$.

Admitindo agora uma base incompleta (com valores ausentes), a probabilidade condicional pode ser reescrita conforme mostra equação abaixo:

$$\hat{p}(x_{ik} | \pi_{ij}) = \frac{\alpha_{ijk} + n(x_{ik} | \pi_{ij})}{\alpha_{ij} + n(\pi_{ij})} \quad (31)$$

sendo α_{ijk} a frequência atribuída ao valor x_{ik} , admitindo uma base de treinamento completa.

Para atribuir um valor a α_{ijk} , os métodos de aprendizagem podem assumir algumas premissas em relação à distribuição dos dados incompletos ou ausentes pelos atributos da base de treinamento. Rubin (1976) enumerou três tipos de distribuição:

- D₁: A distribuição dos dados ausentes é completamente aleatória;
- D₂: A distribuição dos dados ausentes é função dos dados presentes na base de treinamento;
- D₃: A distribuição dos dados ausentes é função dos dados presentes e dados ausentes (ex., nenhuma ocorrência de um valor ou estado discreto de uma determinada variável aleatória).

O método de aprendizagem EM utilizado para a aprendizagem dos parâmetros da rede Bayesiana, descrito na Seção 3.4, assume a distribuição D₂ para os dados ausentes. Ramoni e Sebastiani (2001) avaliaram o impacto de dados ausentes no aprendizado de redes Bayesianas, com base em uma medida de confiança $1-w(x_{ik} | \pi_{ij})$, variando de 0 a 1, com valores próximos de zero indicando falta de confiança, e valores próximos de 1, alta confiança. Eles mediram uma confiança abaixo de 0,45 com 75% de dados ausentes, considerando a distribuição D₂.

As técnicas de tratamento aos dados ausentes do conjunto de treinamento podem ser agrupadas em duas categorias: (1) técnicas de preenchimento artificial dos dados ausentes e (2) utilização de modelos mais reduzidos. A ideia das técnicas de preenchimento é estimar os dados ausentes de uma instância a partir dos dados presentes (HASTIE *et al.*, 2001). Em contraste, o objetivo das técnicas de redução é manter no modelo de classificação as características mais importantes ou conhecidas, retirando os atributos não relevantes ou com dados ausentes (FRIEDMAN, 1997). Saar-Tsechansky *et al.* (2007) estudaram algumas técnicas de tratamento aos dados ausentes utilizadas em modelos baseados em árvore de decisão. As técnicas baseadas em modelos reduzidos apresentaram um desempenho superior em relação às baseadas no preenchimento de dados.

Assim, visando melhorar a robustez e o desempenho da rede Bayesiana, adotamos uma abordagem reduzida, excluindo do modelo os atributos com percentual de dados incompletos ou ausentes acima de um valor de limiar. Dadas as características das bases (CERAD e CDA), e o número elevado de atributos com dados ausentes e incompletos, especificamos este valor de limiar em 70%. Este valor de limiar apresenta uma medida de confiança dentro do esperado (RAMONI e SEBASTIANI, 2001).

4.5.4 DISCRETIZAÇÃO DOS ATRIBUTOS NUMÉRICOS

Os atributos da base de treinamento podem ser nominais ou numéricos (FAYYAD e IRANI, 1993). Os atributos nominais representam categorias ou um conjunto limitado de níveis codificados por valores numéricos prefixados. Os atributos numéricos representam números inteiros ou reais, medidos em um determinado domínio numérico. Alguns classificadores, como aqueles baseados em árvore de decisão, redes Bayesianas, ou regras, possuem dificuldades em lidar com atributos numéricos (ELOMAA e ROUSU, 2004; QUINLAN, 1996). Uma solução para este problema é subdividir os atributos numéricos em vários intervalos menores e tratar cada intervalo como uma categoria. O processo de transformar atributos numéricos em valores categóricos ou nominais é denominado discretização (KOTSIANTIS e KANELLOPOULOS, 2006). O resultado do processo de discretização é a estimativa do conjunto de pontos de corte, representando os limites inferior e superior dos intervalos numéricos adjacentes, abrangendo todo o domínio de valores do atributo. Em geral, há um compromisso entre o número de intervalos e o desempenho do aprendizado.

Publicações mostraram um levantamento de diferentes métodos de discretização, classificando-os nas seguintes categorias (KOTSIANTIS e KANELLOPOULOS, 2006; LIU, H. *et al.*, 2002): (1) não supervisionado ou supervisionado, (2) direto ou incremental, (3) global ou local, (4) estático ou dinâmico, (5) de cima para baixo, ou baixo para cima. O método de discretização não supervisionado não utiliza informações derivadas da classe da instância para estimar os pontos de corte. Alguns autores atribuem um desempenho melhor aos métodos de discretização supervisionados, ou seja, aqueles que utilizam informações da classe da instância (DOUGHERTY *et al.*, 1995). Os métodos de discretização diretos subdividem inicialmente o domínio numérico do atributo em k intervalos simultaneamente e de igual tamanho. Já os métodos incrementais iniciam com poucos intervalos numéricos, acrescentando pontos de corte recursivamente (ELOMAA e ROUSU, 2004). Os métodos globais consideram o processo de discretização como uma etapa pré-processamento do aprendizado ou indução do classificador, enquanto que nos métodos locais a estimativa dos intervalos ocorre durante o aprendizado ou indução (LIU, H. *et al.*, 2002). A diferença entre os métodos estáticos e dinâmicos consiste em utilizar outras características durante o processo de discretização (STECK e JAAKKOLA, 2004). Ao contrário dos métodos dinâmicos, nos métodos estáticos, os pontos de corte são estimados de forma independente das características de outros atributos da base de treinamento. Os métodos denominados de cima para baixo consideram a definição de um intervalo contendo todo o domínio numérico do atributo, que segue subdividindo-o de forma recursiva, guiado por algum algoritmo de busca heurístico (KERBER, 1992). Por outro lado, nos métodos de baixo para cima, é determinado inicialmente o número máximo de intervalos, agregando-os de forma recursiva, guiado por alguma heurística.

Um ponto importante no método é a escolha da medida de qualidade dos resultados da discretização. O MDL (*Minimum Description Length*) oferece uma solução geral ao problema da seleção de modelos (GRÜNWALD *et al.*, 2005). Quando vários modelos competem pela explicação de um conjunto de dados, o princípio do MDL é que devemos selecionar aquele modelo que melhor comprime os dados. A capacidade de compressão de um conjunto de dados está diretamente relacionada à descoberta de regularidades nesse conjunto de dados. Quando maior a compressão, maior o aprendizado sobre os dados.

O método de discretização selecionado foi a nova medida MDL descrita por Kononenko (1995). Conforme as características descritas anteriormente, este método é supervisionado, incremental, local, estático, e considera inicialmente o menor número de intervalos, denominado, portanto, de cima para baixo. A Equação 31 mostra a medida MDL.

$$MDL = \frac{1}{n} \left(\log \binom{n}{n_1, \dots, n_C} - \sum_j \log \binom{n_j}{n_{1j}, \dots, n_{Cj}} + \log \binom{n+C-1}{C-1} - \sum_j \log \binom{n_j+C-1}{C-1} \right) \quad (32)$$

sendo n a quantidade total de instâncias, n_j a quantidade de instâncias com j -ésimo valor, dado o atributo, n_i a quantidade de instâncias da i -ésima classe, e C a quantidade de classes.

Uma função busca é executada com objetivo minimizar MDL para diferentes pontos de corte candidatos no domínio numérico do atributo. Utilizamos a implementação deste método de discretização disponível no aplicativo WEKA.

4.5.5 DESCRIÇÃO DOS ATRIBUTOS SELECIONADOS

Em relação à base do CERAD, a Tabela 8 e a Tabela 9 apresentam a relação de atributos selecionados como variável aleatória da rede Bayesiana modelada para o diagnóstico de demência e DA, respectivamente. A coluna “Nível” representa a posição da variável aleatória na estrutura causal da rede Bayesiana, e será detalhado na Seção 4.6. A tabela apresenta também o percentual de dados incompletos ou ausentes de cada atributo e o valor do ganho de informação $H(X)$ calculado na fase de preparação da rede.

Tabela 8. Relação de atributos selecionados da base do CERAD como variável aleatória da rede Bayesiana para diagnóstico de Demência.

Nível ^(a)	Ordem e variável ou atributo	Estados ^(b)	Descrição	Dados incompletos %	Ganho de informação H(X)
D	Demência	1: Negativo 2: Positivo	Indicação de demência	-	-
B	1. Educação	1: (0-9] 2: [10-16] 3: [17-19] 4: [20-inf)	Educação (em anos)	0	0,0396
F	2. CDR (<i>Clinical dementia rating scale</i>)	1: 0 2: 0,5 3: 1 4: 2 5: 3 6: 4 7: 5	<i>Clinical Dementia Rating</i> Adaptação proposta pelo CERAD, incluindo ao CDR originalmente proposto por Morris (1993), os níveis 4 (profunda) e 5 (terminal).	0	0,0178
B	3. Etnia	1:caucasiano 2:afrodescendente	Etnia	0	0,0167
F	4. MMSE (<i>Mini mental state score</i>)	1: (0-27] 2: [28-30]	Mini-exame do estado mental (FOLSTEIN <i>et al.</i> , 1975)	1	0,0115
F	5. Nomeação de Boston (<i>Boston</i>	1: (0-14]	Teste de nomeação de Boston (WILLIAMS <i>et al.</i> , 1989)	1	0,0093

	<i>naming</i>)	2: [15-inf)			
F	6. Blessed	1: (0-5] 2: [6-inf)	Pontuação ponderada de Blessed (<i>Total Weighted Score-Short Blessed</i>) (BALL <i>et al.</i> , 1999)	1	0,0067
F	7. Déficits cognitivos relatados pelo paciente ou informante	1: Negativo 2: Positivo	Déficits cognitivos relatados pelo paciente ou informante	54	0,0041
F	8. AVD (Atividades da Vida Diária)	1: 0 2: 0,5 3: 1 4: 2	Atividades da Vida Diária (<i>Reason-Perform Household tasks – total score</i>) (GALASKO <i>et al.</i> , 1997)	15	0,0040
F	9. Recordação de lista de palavras	1: 0 2: > 0	Recordação de lista de palavras (<i>Word list recall</i>) (WELSH <i>et al.</i> , 1991)	14	0,0034
F	10. Dificuldades no uso de elementos de linguagem	1: sem dificuldade 2: dificuldade leve 3: dificuldade moderada	Dificuldades no uso de elementos de linguagem	1	0,0027
B	11. Gênero	1: Masculino 2: Feminino	Gênero	0	0,0026
F	12. Declínio gradual das funções cognitivas	1: Negativo 2: Positivo	Declínio gradual das funções cognitivas	19	0,0021
F	13. Habilidades funcionais	1: Negativo 2: Positivo	Preservação da independência e habilidades funcionais (<i>Preservation of independence in functional abilities</i>) (WELSH <i>et al.</i> , 1991)	35	0,0012
F	14. Memória de lista de palavras	1: 0 2: > 0	Memória de lista de palavras (<i>Word list memory</i>) (WELSH <i>et al.</i> , 1991)	13	0,0005
F	15. Afasia progressiva primária	1: Negativo 2: Positivo	Afasia progressiva primária	50	0,0004

- (a) Indica o nível que variável aleatória foi posicionada na rede Bayesiana, onde *B* (*Background information*) refere-se ao primeiro nível (dados demográficos, fatores de pré-disposição ou exclusão), *D* (*Disease*) ao segundo nível representando as doenças de interesse, *F* (*Findings*) refere-se ao terceiro nível (sinais, sintomas, resultados de testes).
- (b) Relaciona os níveis discretos representando intervalos numéricos (para atributos numéricos) ou estados. Os intervalos foram obtidos da aplicação do método de discretização supervisionada.

Tabela 9. Relação de atributos selecionados da base do CERAD como variável aleatória da rede Bayesiana para diagnóstico da Doença de Alzheimer.

Nível ^(a)	Ordem e variável ou atributo	Estados ^(b)	Descrição	Dados incompletos %	Ganho de informação H(X)
D	Doença de Alzheimer	1: Negativo 2: Positivo	Indicação de Doença de Alzheimer	-	-
F	1. Blessed	1: (0-8] 2: [9-inf)	Pontuação ponderada de Blessed (<i>Total Weighted Score-Short Blessed</i>) (BALL <i>et al.</i> , 1999)	1	0,0435
F	2. Dificuldades no uso de elementos de linguagem	1: sem dificuldade 2: dificuldade leve 3: dificuldade moderada	Dificuldades no uso de elementos de linguagem	1	0,0154
F	3. Nomeação de Boston (<i>Boston</i>)	1: (0-11]	Teste de nomeação de Boston	1	0,0112

	<i>naming</i>)	2: [12-13] 3: 14 4: [15-inf)	(WILLIAMS <i>et al.</i> , 1989)		
F	4. MMSE (<i>Mini mental state score</i>)	1: (0-14] 2: [15-30]	Mini-exame do estado mental (FOLSTEIN <i>et al.</i> , 1975)	1	0,0096
F	5. CDR (<i>Clinical dementia rating scale</i>)	1: 0 2: 0,5 3: 1 4: 2 5: 3 6: 4 7: 5	<i>Clinical Dementia Rating</i> Adaptação proposta pelo CERAD, incluindo ao CDR originalmente proposto por Morris (1993), os níveis 4 (profunda) e 5 (terminal).	0	0,0079
F	6. Recordação de lista de palavras	1: 0 2: > 0	Recordação de lista de palavras (<i>Word list recall</i>) (WELSH <i>et al.</i> , 1991)	15	0,0075
F	7. Declínio gradual das funções cognitivas	1: Negativo 2: Positivo	Declínio gradual das funções cognitivas	19	0,0053
B	8. Etnia	1:caucasiano 2:afrodescendente	Etnia	0	0,0052
B	9: Acidente cerebrovascular	1: Negativo 2: Positivo	Acidente cerebrovascular	53	0,0032
F	10: Reconhecimento de lista de palavras	1: 0 2: > 0	Reconhecimento de lista de palavras (<i>Word list recognition</i>) (WELSH <i>et al.</i> , 1991)	15	0,0027
B	11. Gênero	1: Masculino 2: Feminino	Gênero	0	0,0025
F	12: Fluência verbal	1: Negativo 2: Positivo	Verificação de fluência verbal	12	0,0017
F	13: Mudanças na personalidade e comportamento	1: Negativo 2: Positivo	Mudanças na personalidade e comportamento	9	0,0008
F	14. AVD (Atividades da Vida Diária)	1: 0 2: 0,5 3: 1 4: 2	Atividades da Vida Diária (<i>Reason-Perform Household tasks – total score</i>) (GALASKO <i>et al.</i> , 1997)	15	0,0008
F	15. Memória de lista de palavras	1:0 2: > 0	Memória de lista de palavras (<i>Word list memory</i>) (WELSH <i>et al.</i> , 1991)	13	0,0008
F	16. Déficits cognitivos relatados pelo paciente ou informante	1: Negativo 2: Positivo	Déficits cognitivos relatados pelo paciente ou informante	54	0,0008
B	17: Educação	1: (0-9] 2: [10-16] 3: [17-19] 4: [20-inf)	Educação (em anos)	0	0,0004

(a) Indica o nível que variável aleatória foi posicionada na rede Bayesiana, onde *B* (*Background information*) refere-se ao primeiro nível (dados demográficos, fatores de pré-disposição ou exclusão), *D* (*Disease*) ao segundo nível representando as doenças de interesse, *F* (*Findings*) refere-se ao terceiro nível (sinais, sintomas, resultados de testes).

(b) Relaciona os níveis discretos representando intervalos numéricos (para atributos numéricos) ou estados. Os intervalos foram obtidos da aplicação do método de discretização supervisionada.

Em relação à base CDA, a Tabela 10 apresenta a relação de atributos selecionados como variável aleatória da rede Bayesiana modelada para o diagnóstico de Demência, DA, e TCL, respectivamente. Os atributos foram selecionados com base no ganho de informação $H(X)$, detalhado na Seção 4.5.3, e o percentual de dados incompletos ou ausentes. Medimos o ganho de informação para cada subconjunto de treinamento. Selecionamos os atributos com medida de ganho de informação maior que zero em pelo menos um subconjunto de treinamento. Os valores estão apresentados na Tabela 10.

Tabela 10. Relação de atributos selecionados da base do CDA como variável aleatória da rede Bayesiana para diagnóstico da Demência, Doença de Alzheimer e Transtorno Cognitivo Leve.

N ^(a)	Ordem e variável ou atributo	Descrição	DI ^(b) %	Ganho de Informação H(X)			A ^(c)
				D	DA	TCL	
D	Demência (<i>Dementia</i>)	Indicação de Demência	-	-	-	-	C
D	Doença de Alzheimer (<i>Alzheimer's Disease</i>)	Indicação de Doença de Alzheimer	-	-	-	-	C
D	Transtorno cognitivo leve (<i>Mild cognitive disorder</i>)	Indicação de Transtorno Cognitivo Leve	-	-	-	-	C
F	1. MMSE (<i>Mini-mental state examination score</i>)	Mini-exame do estado mental (FOLSTEIN <i>et al.</i> , 1975)	5	0,2791	-	0,4250	A
F	2. CDR (<i>Clinical dementia rating scale</i>)	Escala do teste neuropsicológico <i>Clinical Dementia Rating</i> (MORRIS, J. C., 1993)	11	0,2441	0,0041	0,1693	A
F	3. Pfeffer (<i>Pfeffer questionnaire score</i>)	Questionário Pfeffer ou a escala Bayer de Atividades da Vida Diária (<i>Bayer Scale for Activities of Daily Living</i>) (PFEFFER <i>et al.</i> , 1982)	12	0,2074	-	0,1952	A
F	4. VFT (<i>Verbal fluency test score</i>)	Teste de fluência verbal (PASSOS <i>et al.</i> , 2011)	8	0,1665	-	0,2448	A
F	5. CDT (<i>Clock drawing test scale</i>)	Teste do desenho do relógio	12	0,0881	0,0327	0,1360	A
F	6. TMT (<i>Trial making test</i>)	Teste de trilha: instrumento para avaliar atenção visuomotora	40	0,0829	0,0184	0,1368	A
B	7. Idade (<i>Age</i>)	Faixa etária	4	0,0647	-	-	A
F	8. Lawton (<i>Lawton scale</i>)	Escala de Lawton, usada como medida de qualidade de vida para demência	58	0,0384	-	0,0450	A
F	9. IQCode (<i>IQCode score</i>)	Questionário do Informante sobre Declínio Cognitivo em Idosos (<i>The Informant Questionnaire on Cognitive Decline in the Elderly</i>) (JORM, 2004)	56	0,0342	-	0,0414	A
F	10. Stroop (<i>Stroop color word test</i>)	Teste Stroop desempenho de cor-palavra	60	0,0324	-	0,0726	A
F	11. Berg (<i>Berg balance scale</i>)	Escala de Berg para equilíbrio funcional	60	0,0209	-	0,0645	A
B	12. Gênero (<i>Gender</i>)	Masculino ou feminino	9	0,0001	0,0038	0,0403	A

F	13. Depressão (<i>Depression</i>)	Presença de sintomas relacionados à depressão	16	0,0001	0,0154	0,0079	A
B	14. Educação (<i>Education</i>)	Nível educacional (em anos)	2	-	0,0423	0,1781	A
-	15. Figura Rey (<i>Rey Complex Figure</i>)	Instrumento de avaliação neuropsicológica (JAMUS e MÄDER, 2005)	78	0,0181	-	-	F
-	16. CAMCOG (<i>Cambridge Cognitive Examination</i>)	Exame cognitivo de Cambridge	79	-	-	-	F
-	17. Símbolos e Dígitos (<i>Digit Symbol</i>)	Teste de símbolos e dígitos (<i>Digit symbol test</i>) (HART, R. P. <i>et al.</i> , 1987)	81	-	-	-	F
-	18. NPI (<i>Neuropsychiatric inventory</i>)	Inventário neuropsiquiátrico (CUMMINGS <i>et al.</i> , 1994)	56	-	-	-	R
-	19. Cornell (<i>Cornell depression scale</i>)	Escala de depressão de Cornell (CARTHERY-GOULART <i>et al.</i> , 2007)	62	-	-	-	R
-	20. TUG	Teste de mobilidade funcional básica para idosos (<i>Timed Up and Go</i>) (PODSIADLO e RICHARDSON, 1991)	64	-	-	-	R
-	21. POMA	Avaliação de desempenho de mobilidade orientada	85	-	-	-	RF
-	22. STS	Avalia coordenação motora, equilíbrio e mobilidades básicas (<i>Sit-to-Stand test</i>) (JANSSEN <i>et al.</i> , 2002)	97	-	-	-	RF
-	23. <i>Digit Span</i>	Avalia a memorização e sequência de dígitos (<i>Digit span test</i>) (CONKLIN <i>et al.</i> , 2000)	62	-	-	-	R
-	24. RAVL	Teste de aprendizagem auditivo-verbal de Rey (<i>Rey Auditory-Verbal Learning test</i>) (BARZOTTI <i>et al.</i> , 2004)	93	-	-	-	RF
-	25. Volume de estruturas do encéfalo	Volumes de estruturas do encéfalo medido em imagens MRI	83	-	-	-	RF

(a) Indica o nível que variável aleatória foi posicionada na rede Bayesiana, onde *B* (*Background information*) refere-se ao primeiro nível (dados demográficos, fatores de predisposição ou exclusão), *D* (*Disease*) ao segundo nível representando as doenças de interesse, *F* (*Findings*) refere-se ao terceiro nível (sinais, sintomas, resultados de testes).

(b) DI = Dados incompletos. Contém o percentual de dados incompletos no subconjunto de demência.

(c) A = Análise, sendo: C = Classe; A = atributo selecionado; F = atributos não selecionado devido ao elevado percentual de dados faltantes; R = atributo não selecionado devido à baixa relevância obtida no filtro Ganho de Informação. RF = atributo não selecionado devido à baixa relevância e elevado percentual de dados faltantes.

O método de discretização foi executado para cada subconjunto de treinamento. Utilizamos os mesmos níveis discretos para as variáveis aleatórias pertencentes à rede Bayesiana para diagnóstico de Demência e DA: os pontos de corte estimados pelo método de discretização do subconjunto de treinamento para rede Bayesiana de DA estavam contidos no subconjunto de treinamento para rede Bayesiana de Demência. Isto não ocorreu no caso da rede Bayesiana utilizada para diagnóstico de TCL. A Tabela 11 mostra os níveis discretos para as três redes Bayesianas.

Tabela 11. Estados ou níveis discretos definidos para cada variável aleatória das redes Bayesianas da base do CDA.

Ordem e variável ou atributo	Estados	
	Demência, DA	TCL
Demência	1: Negativo 2: Positivo	1: Negativo 2: Positivo
Doença de Alzheimer	1: Negativo 2: Positivo	1: Negativo 2: Positivo
Transtorno cognitivo leve	1: Negativo 2: Positivo	1: Negativo 2: Positivo
1. MMSE	1: [0-17] 2: [18-26] 3: [27-30]	1: [0-28] 2: [29-30]
2. CDR	1: 0 2: 0,5 3: 1 4: 2 5: 3	1: 0 2: 0,5 3: 1 4: 2 5: 3
3. Pfeffer	1: 0 2: [1-2] 3: [3-inf)	1: [0-1] 2: [2-inf) 3: [3-inf)
4. VFT	1: [0-4] 2: [5-11] 3: [12-inf)	1: [0-15] 2: [16-inf)
5. CDT	1: 0 2: 1 3: 2 4: 3 5: 4 6: 5	1: 0 2: 1 3: 2 4: 3 5: 4 6: 5
6. TMT	1: [0-16] 2: [17-59] 3: [60-inf)	1: [0-36] 2: [37-inf)
7. Idade	1: [0-72] 2: [73-inf)	1: [0-69] 2: [70-inf)
8. Lawton	1: [0-9] 2: [10-inf)	1: [0-14] 2: [15-inf)
9. IQCode	1: [0-3,55] 2: [3,56-inf)	1: [0-3,32] 2: [3,33-inf)
10. Stroop	1: [0-15] 2: [16-inf)	1: [0-17] 2: [18-inf)
11. Berg	1: [0-51] 2: [52-inf)	1: [0-54] 2: 55 3: [56-inf)
12. Gênero	1: Feminino 2: Masculino	1: Feminino 2: Masculino
13. Depressão	1: Ausência 2: Presença	1: Ausência 2: Presença
14. Educação	1: [0-13] 2: [14-inf)	1: [0-15] 2: [16-inf)

4.6 MODELAGEM DA REDE BAYESIANA

As próximas subseções descrevem a definição da estrutura Bayesiana, a estimativa da distribuição dos parâmetros, e a verificação de independência entre as variáveis aleatórias da rede Bayesiana.

4.6.1 DEFINIÇÃO DA ESTRUTURA DA REDE BAYESIANA

Cada atributo selecionado foi associado a uma variável aleatória (nodo) da rede Bayesiana. No total, foram construídos cinco modelos Bayesianos, um para cada diagnóstico e base de treinamento. São eles: (1) Demência e (2) DA utilizando a base CERAD, (3) Demência, (4) DA e (5) TCL utilizando a base CDA. A modelagem da rede Bayesiana ocorre em duas etapas: mapeamento das variáveis aleatórias e identificação do relacionamento causal entre as mesmas. Conforme mencionado anteriormente, utilizamos cada atributo selecionado para cada base e doença como variável aleatória da rede Bayesiana correspondente.

Para apoiar a identificação do relacionamento causal entre as variáveis aleatórias, utilizamos uma estrutura de rede Bayesiana genérica composta de três níveis, proposta por Shwe *et al.* (1991) e Pradhan *et al.* (1994), mostrada na Figura 15. No primeiro nível, posicionamos as variáveis aleatórias que representam os fatores de pré-disposição ou exclusão da doença de interesse, incluindo dados demográficos do paciente (ex., idade, gênero, grupo étnico), onde B_m (*Background information*) expressa a m -ésima variável aleatória. As variáveis aleatórias deste nível não são condicionadas a nenhum fator causal. As variáveis aleatórias que representam as doenças de interesses aparecem no segundo nível, onde Q_n (*Query node*) expressa a n -ésima variável aleatória, com dois estados discretos apenas: *Positivo* e *Negativo*. Estes estados representam a diagnóstico médico da doença de interesse, e sua distribuição de probabilidades é condicionalmente dependente das variáveis aleatórias localizadas no primeiro nível da rede. A consulta da distribuição das probabilidades *a posteriori* (após a entrada das evidências) deverá indicar a probabilidade do paciente manifestar, ou não, a doença de interesse, confirmando ou refutando uma hipótese de diagnóstico. No terceiro nível, estão as variáveis aleatórias que representam os fatores de manifestação relacionados à doença, onde F_o (*findings*) expressa a o -ésima variável aleatória. Os fatores de manifestação incluem sinais, sintomas e resultados de testes de avaliações neuropsicológicas, e sua distribuição de probabilidades é condicionalmente dependente da variável aleatória que representa a doença de interesse. A Figura 15 mostra também a estrutura genérica representada em diagrama de influência. O diagrama de influência

representa uma extensão da rede Bayesiana, agregando ao modelo, os nodos U_n (*Utility*), expressando a n -ésima função utilidade, relacionada ao nó Q_n , e os nodos D_n (*Decision node*), expressando o n -ésimo ponto de decisão, relacionados ao nó U_n .

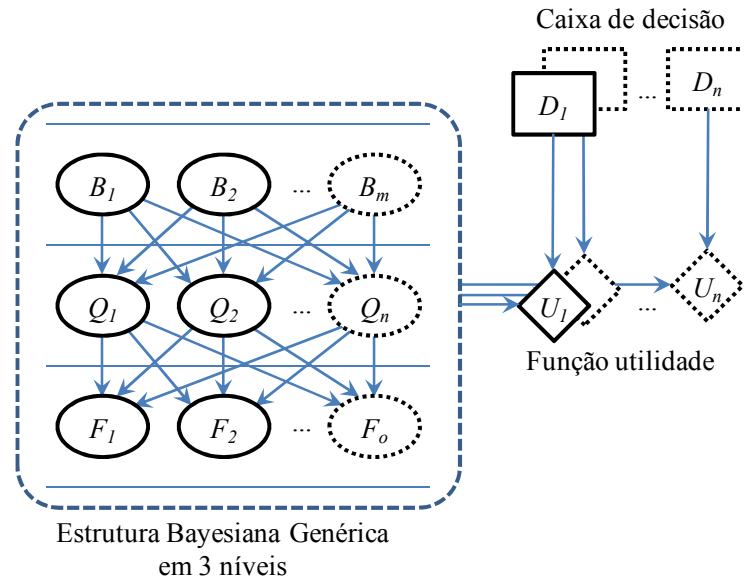


Figura 15. Representação gráfica dos níveis da rede Bayesiana, mostrando a dependência causal entre as variáveis e demais elementos de uma rede de decisão: função utilidade e caixa de decisão.

O nível que cada variável aleatória foi posicionada na rede Bayesiana dentro desta estrutura genérica é mostrado nas Tabelas 8, 9 e 10, na coluna *Nível*. A definição de qual nível da rede Bayesiana aquela variável aleatória pertence foi baseada na análise da causalidade do atributo com base nos critérios de diagnóstico da doença de interesse, e com a validação de um especialista do domínio do conhecimento. Os especialistas são médicos psiquiatras, com especialidade em doenças de transtornos mentais, que compõem a equipe do CDA na UFRJ.

4.6.2 ESTIMATIVA DOS PARÂMETROS

Cada variável aleatória é atribuída uma distribuição de probabilidades condicionais ou não condicionais. A distribuição das probabilidades condicionais pondera a relação causal entre duas (ou mais) variáveis aleatórias. Conforme Wang (2007), considerando que a construção da estrutura da rede Bayesiana já é um desafio na elicitación, a ponderação das variáveis aleatórias exige a especificação de inúmeras probabilidades condicionais. Uma maneira de obter os valores para probabilidades condicionais é através da entrevista ao especialista médico no momento da elicitación. Contudo, o julgamento humano sobre a

probabilidade muitas vezes é propenso a erros sistemáticos motivados por uma variedade de fatores (TVERSKY e KAHNEMAN, 1974). Aplicamos, portanto, um algoritmo de aprendizagem computacional para estimar os parâmetros (distribuição de probabilidades) da rede Bayesiana. Utilizamos as bases de casos clínicos (CERAD e CDA) como base de treinamento, avaliadas separadamente, obtendo, no total, cinco redes Bayesianas, conforme descrito anteriormente.

Como algoritmo aprendizagem Bayesiana, utilizamos o EM (*Expectation Maximization*), pelas seguintes razões: (1) adequação às distribuições discretas, (2) bom desempenho para observações parciais, ou seja, base de treinamento com dados ausentes (estados nunca observados no subconjunto de treinamento), e incompletos (variáveis cujos estados às vezes aparecem na base de treinamento, e às vezes não). A distribuição *a priori* dos parâmetros é a distribuição de Dirichlet por ser mais adequada para dados discretos e/ou multinomiais, conforme exposto na Seção 3.1. Sendo α_{ijk} o hiperparâmetro da distribuição de Dirichlet para variável aleatória i assumindo o j -ésimo estado e nodos pais, o k -ésimo estado. Assumimos uma distribuição inicial *a priori* uniforme. Então, para a variável aleatória X_i , utilizamos o hiperparâmetro $\alpha_{ijk} = \frac{\alpha_0}{q_i \cdot r_i}$, onde α_0 representa a ponderação da distribuição de Dirichlet, q_i o número de estados das variáveis aleatórias pais de X_i , e r_i o número de estados de X_i (BUNTINE, 1991). Atribuímos $\alpha_0 = 1$, admitindo uma distribuição inicial uniforme. Assim, α_{ijk} tem como valor inicial o tamanho equivalente da amostra dos parâmetros de verossimilhança.

Utilizamos a implementação do algoritmo EM disponível na ferramenta BNT (*Bayesian Network Toolbox*), desenvolvida no Mathworks MATLAB® (<http://code.google.com/p/bnt/>) (MURPHY, K, 2001). As Figuras 16 e 17 apresentam as curvas de aprendizado, utilizando subconjuntos de treinamento distintos. Estas curvas foram obtidas através do valor esperado da função log-verossimilhança (Equação 27) em cada iteração do algoritmo EM. Utilizamos validação cruzada com 4 subdivisões (*folds*). Cada curva de aprendizado representa um *fold*. Nota-se que, a partir da segunda iteração, a diferença do valor esperado em relação à iteração anterior é pequena, onde é possível concluir que o valor de máxima log-verossimilhança é aproximado em poucas iterações, comprovando assim a eficácia do algoritmo EM. O método de validação será descrito no Capítulo 5.

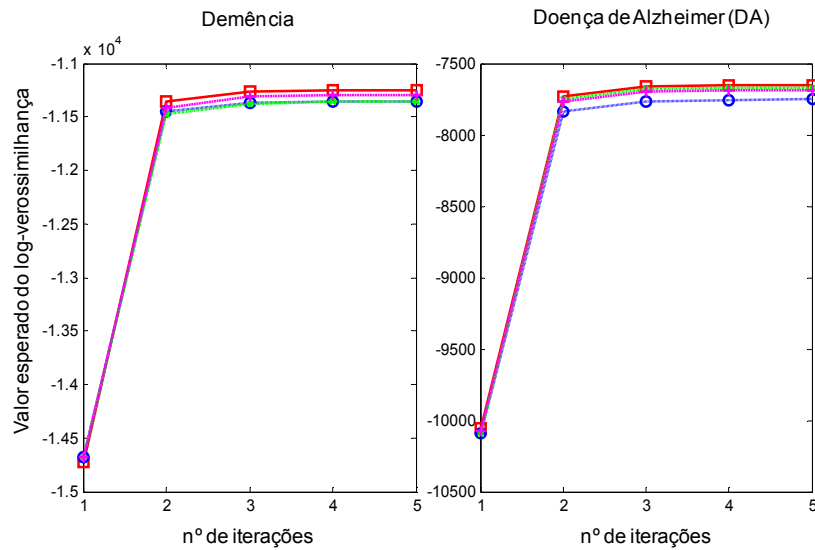


Figura 16. Curva de aprendizado por *fold*, utilizando subconjunto de dados para Demência e DA, base CERAD.

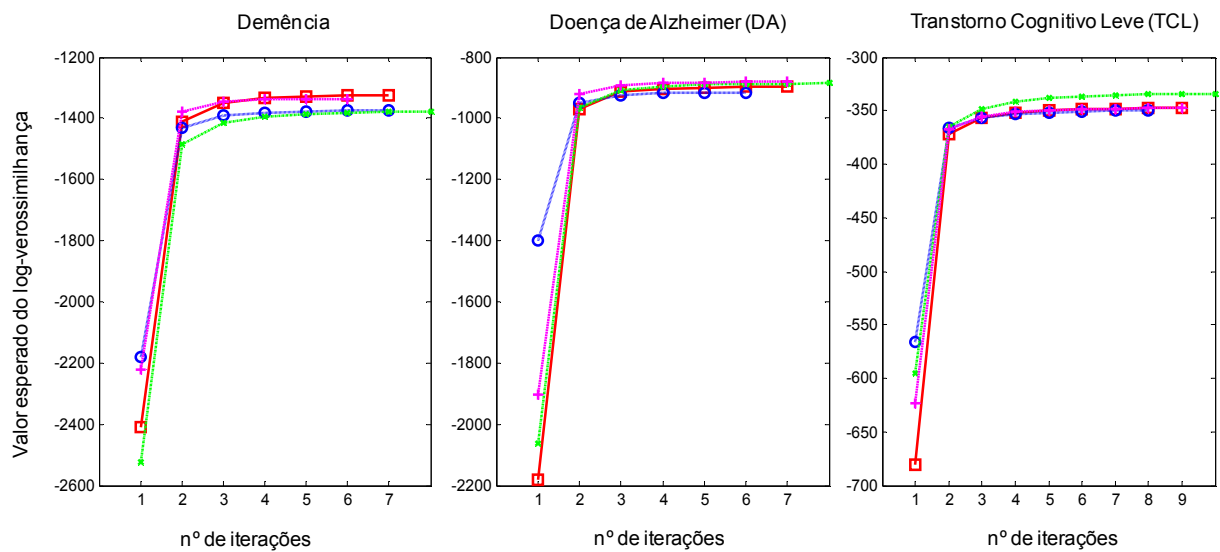


Figura 17. Curva de aprendizado por *fold*, utilizando subconjunto de dados para Demência, DA e TCL, base CDA.

As Figuras 18 e 19 mostram o diagrama de influência, incluindo a rede Bayesiana, modeladas para o diagnóstico de Demência e DA, respectivamente, utilizando a base CERAD. A Figura 20 mostra o diagrama de influência para o diagnóstico de Demência, utilizando a base CDA. Os diagramas de influência para o diagnóstico de DA e TCL possuem a mesma estrutura do diagrama de influência para Demência, mas com distribuição de probabilidades distintas. No caso de TCL, os níveis discretos das variáveis aleatórias são diferentes (ver Tabela 11). Utilizamos a ferramenta de autoria e inferência GeNIe / SMILE,

desenvolvida pelo laboratório de Sistemas de Decisão da Universidade de Pittsburg (<http://genie.sis.pitt.edu>). A variável aleatória central, representada pelo nodo Diagnóstico (*Diagnosis*), expressa a probabilidade de diagnóstico da doença de interesse, dado um conjunto de evidências. A variável aleatória central é conectada ao nó Utilidade (*Utility*). O nó Utilidade representa a função utilidade, utilizada como forma de parametrizar os impactos nos custos e riscos a saúde do paciente, em caso de erro no diagnóstico, conforme detalhado na Seção 3.3. No caso, utilizamos parâmetros simétricos para função utilidade (mesmo valor de utilidade para acertos e erros de classificação), mostrados na Tabela 12. Estes valores maximizam as medidas de desempenho de classificação da rede Bayesiana, e podem ser posteriormente ajustados, utilizando-os como um fator de ponderação, dentro de um processo de revisão e/ou refinamento, se a opção do médico é melhorar a taxa de verdadeiros-positivos (TPR – *True Positive Rate*), em detrimento a taxa de verdadeiros-negativos (TNR – *True Negative Rate*), ou vice-versa.

Os valores representados pelas barras horizontais dos estados das variáveis aleatórias expressam as probabilidades marginais da respectiva variável aleatória obtida pelo mecanismo de inferência, considerando uma rede Bayesiana sem evidências. Estes valores são calculados através das Tabelas de Distribuição de Probabilidades (CPTs – *Conditional Probability Tables*) associadas a cada variável aleatória e estimadas pelo método de aprendizagem a partir da base de treinamento. A utilização de subconjuntos de treinamento específicos para cada diagnóstico gerou CPTs distintas para cada rede Bayesiana, por consequência, probabilidades marginais também distintas. Isso justifica a necessidade de modelar uma rede Bayesiana para cada diagnóstico e/ou base de treinamento.

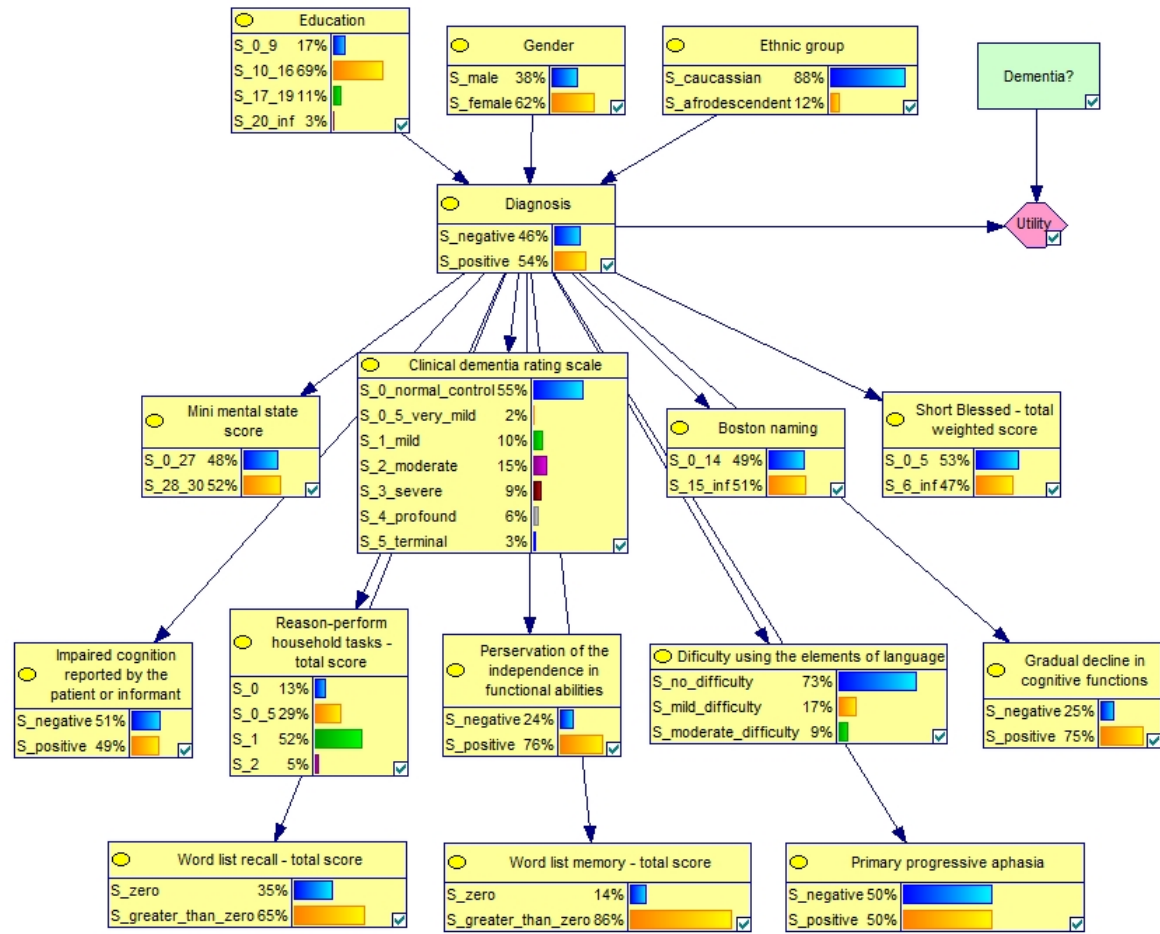


Figura 18. Diagrama de influência modelado para o diagnóstico de Demência, utilizando a base CERAD.

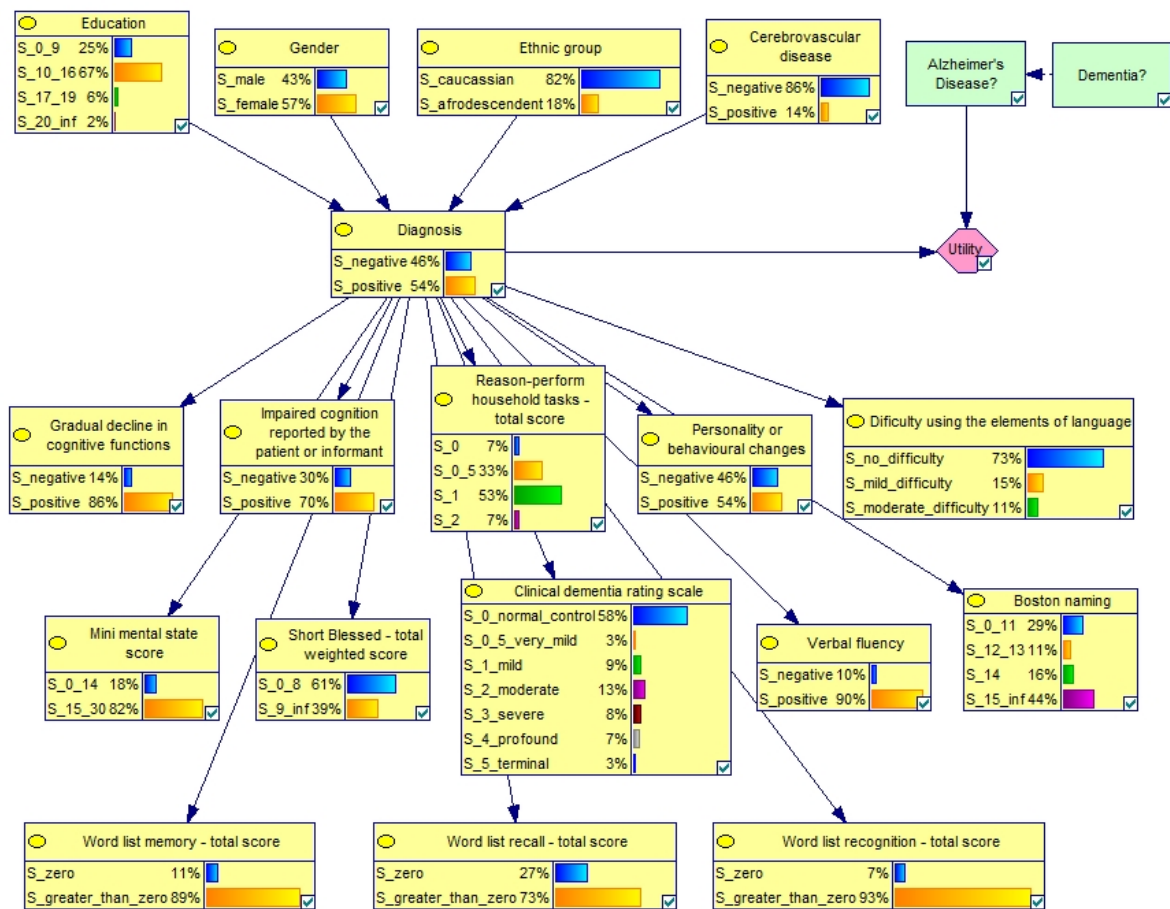


Figura 19. Diagrama de influência modelado para o diagnóstico da Doença de Alzheimer, utilizando a base CERAD.

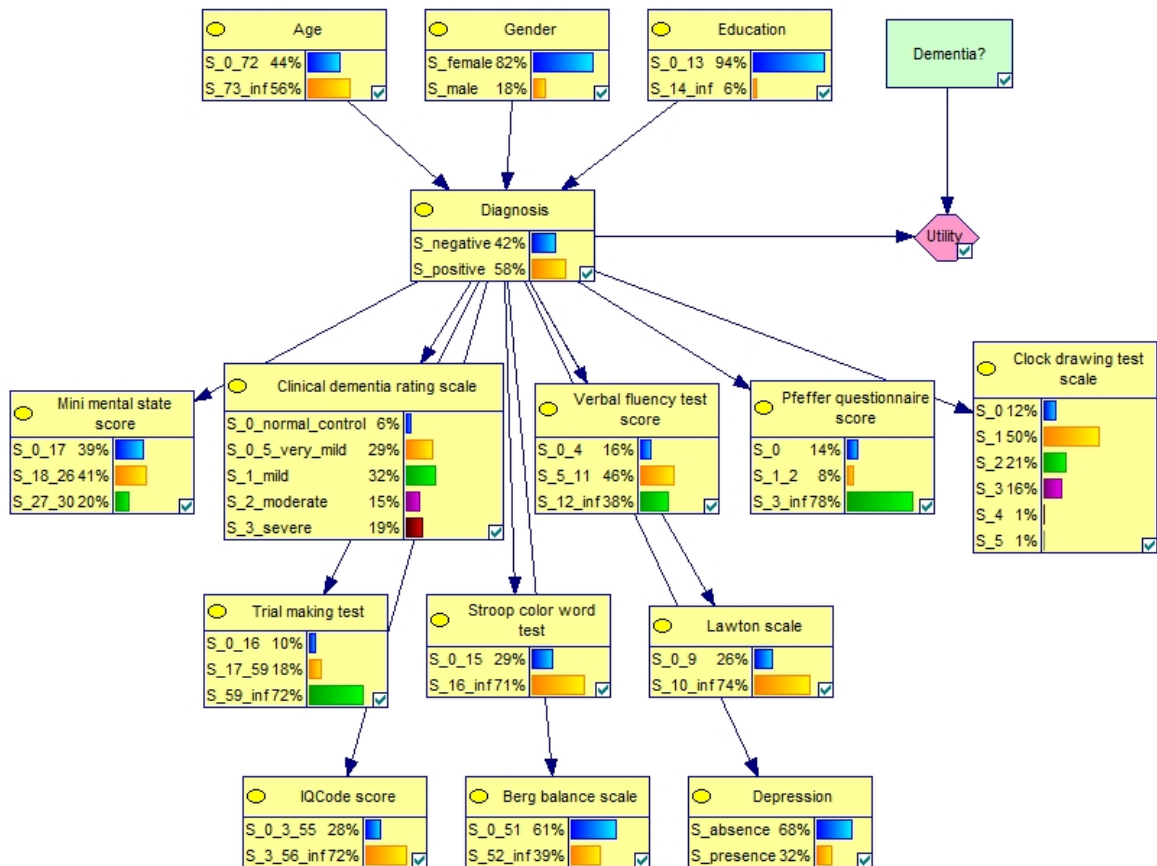


Figura 20. Diagrama de influência modelado para o diagnóstico da Demência, utilizando a base CDA.

Tabela 12. Parâmetros genéricos utilizados na função utilidade.

Diagnóstico:	<i>Positivo</i>		<i>Negativo</i>	
	<i>Positivo</i>	<i>Negativo</i>	<i>Positivo</i>	<i>Negativo</i>
Estados da variável aleatória central ^(a) :				
Valor:	1	-1	-1	1

(a) A variável aleatória central, representada pelo nodo (*node*) Diagnóstico (*Diagnosis*), representando a probabilidade de manifestação, ou não, da doença pelo paciente, dado um conjunto de evidências.

A rede Bayesiana, o nodo Utilidade e as caixas de decisão completam o diagrama de influência. Na Figura 19, é possível notar que a caixa de decisão de diagnóstico de DA está relacionada à caixa de decisão de diagnóstico de Demência. Isto significa que o diagnóstico positivo de Demência deve anteceder a avaliação de diagnóstico de DA, conforme mostrado no processo de diagnóstico (ver Figura 8).

Devido ao número reduzido de instâncias de casos classificados como TCL na do CERAD, não consideramos a modelagem e aprendizagem da rede Bayesiana para o

diagnóstico desta doença nesta base. Este diagnóstico foi considerado apenas na base do CDA.

4.6.3 TESTE DE INDEPENDÊNCIA DAS VARIÁVEIS ALEATÓRIAS

A estrutura Bayesiana assume como premissa, relações de independência condicional na ordem causal das variáveis aleatórias. Seja X_i uma variável aleatória (nodo) da rede Bayesiana. X_i da rede deve ser condicionalmente independente de seus predecessores (ou pais), de modo que a distribuição conjunta para as variáveis aleatórias $P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | Pais(X_i))$, tal qual descrito na Seção 3.2 (Equação 18).

Para verificar esta premissa, levantamos o coeficiente de correlação entre variáveis aleatórias ou nodos do mesmo nível da estrutura Bayesiana, aplicando a Equação 33. Se os atributos não estiverem correlacionados (coeficiente de correlação próximo de zero), é possível concluir que os dados são distribuídos de forma independente (GLYMOUR *et al.*, 1996).

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)} \cdot \sqrt{\text{var}(Y)}} = \frac{\sum_{i=1}^n (x_i - E[X]) \cdot (y_i - E[Y])}{\sqrt{\sum_{i=1}^n (x_i - E[X])^2 \cdot \sum_{i=1}^n (y_i - E[Y])^2}} \quad (33)$$

onde $\text{cov}(X,Y)$ representa função de covariância da distribuição de X e Y , $\text{var}(X)$ e $\text{var}(Y)$ a variância de X , e Y , respectivamente. $E[X]$ e $E[Y]$ representam os valores esperados da distribuição de X e Y , respectivamente.

As variáveis aleatórias não condicionadas (1º nível, ou B_i conforme a Figura 15), e as variáveis aleatórias condicionadas à doença (3º nível, ou F_i conforme a Figura 15), foram avaliadas separadamente. As Tabelas 13 e 14 apresentam os valores de correlação para o diagnóstico de Demência, utilizando as bases CERAD e CDA, respectivamente. Para as variáveis aleatórias condicionadas à doença, utilizamos os subconjuntos de Demência negativo e positivo.

A Figura 21 mostra o gráfico de Boxplot das distribuições dos coeficientes de correlação. No caso da base CERAD, o valor máximo obtido para o coeficiente de correlação (0,66) fora considerado *outlier* pelo critério de construção do Boxplot (BENJAMINI, 1988). Um valor é representado como *outlier* quando o mesmo for superior a 1,5 vezes o 3º quartil, ou inferior a 1,5 vezes o 1º quartil, sendo considerado atípico e fora da construção gráfica do Boxplot. É possível notar que a maioria dos valores de correlação para a base CERAD se

concentrou próximo de zero, o que leva a conclusão de que a maioria das variáveis aleatórias é não correlacionada, ou independente. Já no caso da base CDA, os valores de correlação também se concentraram próximo de zero, contudo, a dispersão dos valores foi significativamente maior em relação à base CERAD. O maior valor de correlação dois a dois apresentado pela base CDA foi 0,83.

Tabela 13. Correlação dois a dois das variáveis aleatórias da rede Bayesiana para diagnóstico de Demência, utilizando a base CERAD.

	3 ^(a)	11	4	5	6	7	8	9	10	12	13	14	15
1º nível (B_i)^(b):													
1	-0,21	-0,12											
3		0,10											
3º nível (F_i Demência = negativo):													
2			0,15	0,06	-0,04	-0,34	0,07	-0,03	0,01	-0,22	-0,07	0,00	-0,20
4				0,60	-0,37	-0,16	-0,15	0,20	-0,31	-0,02	-0,06	0,05	-0,13
5					-0,22	-0,16	-0,13	0,22	-0,19	-0,04	-0,04	0,10	-0,02
6						0,07	0,18	-0,24	0,52	-0,22	0,23	-0,03	0,21
7							0,05	0,01	0,15	0,66	0,41	-0,03	0,13
8								0,06	-0,02	-0,08	-0,07	0,10	-0,03
9									-0,19	0,06	-0,01	0,50	-0,06
10										-0,13	0,25	-0,08	0,10
12											0,03	0,01	-0,06
13												-0,04	0,02
14													-0,08
3º nível (F_i Demência = positivo):													
2			0,06	-0,05	-0,02	-0,09	0,00	-0,03	-0,05	-0,06	-0,08	-0,10	0,10
4				0,55	-0,21	-0,19	-0,05	0,09	-0,19	-0,13	-0,02	0,01	0,00
5					-0,16	-0,15	-0,03	0,07	-0,19	-0,12	-0,02	0,03	-0,02
6						0,24	0,02	-0,10	0,60	0,04	0,03	-0,03	-0,01
7							0,06	0,01	0,16	0,62	0,50	-0,05	-0,01
8								0,04	0,05	0,04	-0,01	-0,01	0,06
9									-0,06	0,06	-0,02	0,51	0,05
10										0,00	0,04	-0,02	0,02
12											-0,03	0,00	0,03
13												0,00	-0,02
14													-0,02

(a) Os números representam os atributos relacionados na Tabela 8.

(b) Representa o nível da rede Bayesiana, mostrado na Figura 15. No 1º nível, as variáveis aleatórias não estão condicionadas. No 3º nível, as variáveis aleatórias aparecem condicionadas ao nível do diagnóstico, e a análise de correlação é, portanto, separada.

Tabela 14. Correlação dois a dois das variáveis aleatórias da rede Bayesiana para diagnóstico de Demência, utilizando a base CDA.

	12 ^(a)	14	2	3	4	5	6	8	9	10	11	13
1º nível (B_i)^(b):												
7	0,13	-0,18										
12		0,038										
3º nível (F_i Demência = negativo):												
1			-0,65	-0,42	0,42	0,61	-0,49	-0,60	-0,40	-0,59	0,64	0,13
2				0,67	-0,37	-0,16	0,04	0,35	0,62	-0,10	-0,42	-0,14
3					-0,57	-0,44	0,22	0,83	0,54	0,73	-0,54	-0,08
4						0,54	-0,36	-0,61	0,06	-0,54	0,08	-0,02
5							-0,72	-0,35	-0,15	-0,65	0,41	-0,08
6								0,45	0,07	0,67	-0,62	-0,21
8									0,00	0,36	-0,65	-0,84
9										0,00	0,00	-0,37
10											-0,33	0,29
11												0,63
3º nível (F_i Demência = positivo):												
1			-0,57	-0,16	0,50	0,52	0,37	-0,18	-0,02	-0,27	0,25	0,11
2				0,24	-0,54	-0,53	-0,46	0,18	-0,06	0,15	-0,41	-0,05
3					-0,09	-0,23	-0,08	0,39	0,15	0,29	-0,29	0,03
4						0,44	0,31	-0,12	-0,06	-0,39	0,33	-0,12
5							0,28	-0,16	0,14	-0,03	0,39	0,03
6								-0,22	-0,09	0,11	0,28	-0,02
8									0,00	-0,06	-0,35	-0,05
9										0,00	0,00	-0,01
10											0,02	0,22
11												0,17

(a) Os números representam os atributos relacionados na Tabela 10

(b) Representa o nível da rede Bayesiana, mostrado na Figura 15. No 1º nível, as variáveis aleatórias não estão condicionadas. No 3º nível, as variáveis aleatórias aparecem condicionadas ao nível do diagnóstico, e a análise de correlação é, portanto, separada.

Portanto, para se concluir o atendimento ao critério de independência das variáveis aleatórias no contexto da rede Bayesiana, é necessário um estudo mais exaustivo, incluindo outros testes de independência (CHIN-JEN e FINE, 2006). É possível também considerar a utilização de métodos de descoberta automática de estrutura da rede Bayesiana para melhorar os resultados dos testes de independência, em vez da estrutura genérica em três níveis descrita na Subseção 4.6.1 (CAMPOS, 2006; DASH e DRUZDZEL, 2003). Uma desvantagem da aplicação destes métodos é o aumento da complexidade da estrutura da rede, dificultando a interpretação e validação dos critérios de diagnóstico pelo especialista do domínio do conhecimento. Além disso, em comparação com a rede Bayesiana de estrutura pré-definida,

não conseguimos observar uma melhoria significativa nos resultados das medidas de classificação. O método de descoberta automática da estrutura de rede e os resultados de desempenho serão mostrados na Seção 5.2.

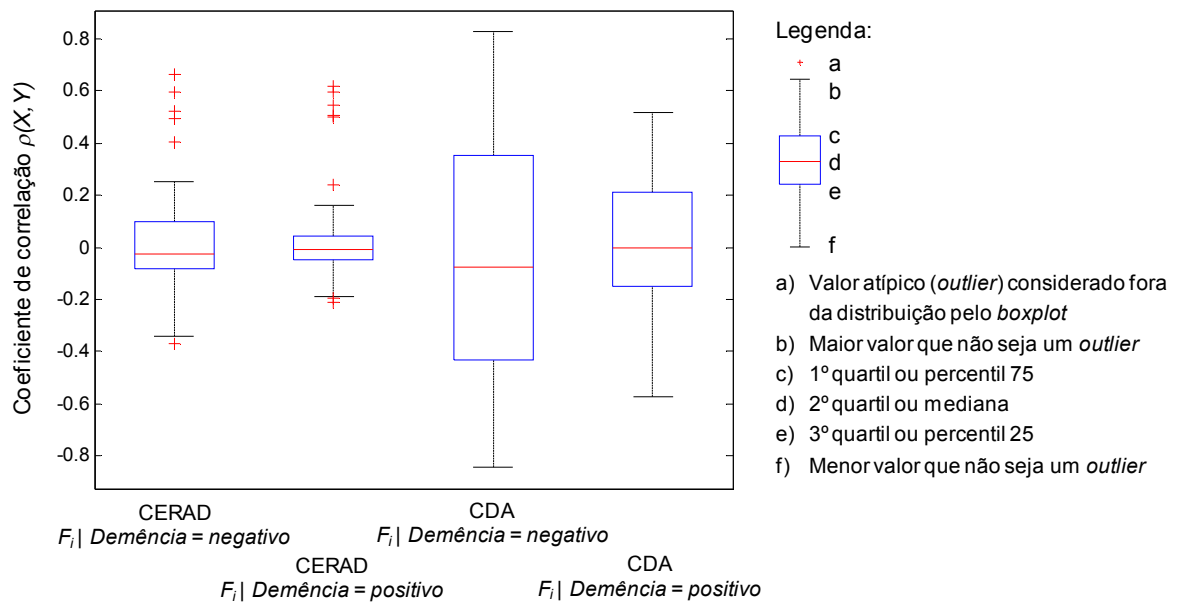


Figura 21. Gráfico Boxplot mostrando distribuição dos coeficientes de correlação obtidos da base CERAD e base CDA. As variáveis aleatórias da distribuição estão condicionadas ao nodo da rede Bayesiana que representa a doença sobre análise.

4.7 DISCUSSÃO

O CDSS proposto nesta tese utiliza como modelo de decisão um diagrama de influência, composto por uma rede Bayesiana, função utilidade e caixas de decisão, associadas a pontos de decisão sequenciais representados em um processo de diagnóstico. Considerando os diferentes tipos de raciocínio apresentados na Figura 4 (Diagnóstico, Causal, Intercausal e Misto), utilizamos o raciocínio misto nesta proposta. No Capítulo 2, apresentamos alguns CDSSs, classificando-os em três categorias: (1) CDSSs baseados em regras, (2) CDSSs baseados em diretrizes clínicas, e (3) CDSSs baseados em redes. Em comparação com o CDSS proposto nesta tese, que é baseado em redes, as principais diferenças em relação aos trabalhos relacionados encontram-se no processo de modelagem da rede Bayesiana.

As técnicas de modelagem Bayesiana aplicáveis ao domínio clínico são: (1) modelagem orientada pelo conhecimento, (2) modelagem orientada por dados e (3) modelagem híbrida, reunindo características das modelagens (1) e (2) (MANI *et al.*, 1997).

Exemplos de modelagem orientada pelo conhecimento podem ser encontrados nos sistemas QMR/Internist-I (MILLER, R.A. *et al.*, 1982), Diaval (DIEZ, F. J. *et al.*, 1997), BayPAD (LUCIANI *et al.*, 2007), PDS (PYPER *et al.*, 2008) e TiMeDDx (DENEKAMP e PELEG, 2010). Um exemplo de modelagem orientada por dados pode ser encontrado no sistema IMASC (CZIBULA *et al.*, 2008). Liu *et al.* (2006) mostraram uma modelagem de rede Bayesiana com estrutura prédefinida e parâmetros estimados por aprendizagem supervisionada, com aplicações na análise genética. Utilizamos uma modelagem orientada por dados, mas com a estrutura de rede Bayesiana definida manualmente, com base na estrutura genérica estendida de rede Bayesiana descrita em Pradhan *et al.* (1994). Shwe *et al.* (1991) propuseram uma adaptação do sistema QMR/Internist-I para um modelo baseado em rede Bayesiana, e adotaram uma estrutura genérica similar, com a diferença que a distribuição de probabilidades foi exclusivamente definida por especialistas do domínio do conhecimento, aplicando um modelo de estrutura local conhecido por *Noisy-Or* (DIEZ, F.J., 1993). A utilização de estruturas Bayesianas genéricas ou canônicas tem por objetivo principal simplificar o processo de aquisição e validação do conhecimento junto ao especialista do domínio do conhecimento (FRIEDMAN *et al.*, 1997).

Cada atributo selecionado da base de treinamento foi utilizado para descrever cada variável aleatória da rede Bayesiana. Avaliando todo o seu conjunto, as variáveis aleatórias são necessárias e suficientes para o diagnóstico das doenças de interesse em todas as suas formas de manifestação, utilizando três fontes de informação para a validação: (1) os critérios clínicos de diagnóstico, (2) o significado dos atributos da base de casos clínicos, e (3) o conhecimento de especialistas médicos do CDA. A distribuição de probabilidades associadas a cada variável aleatória foi estimada por método de aprendizagem supervisionada.

A Tabela 15 mostra a associação em referência cruzada dos critérios de diagnóstico apresentados na Tabela 5 e a relação de atributos da base de casos clínicos do CDA apresentados na Tabela 10. As características presentes no critério para o diagnóstico de DA, tais como, alterações nos biomarcadores do líquido cerebral, diminuição da fluordesoxiglicose, e alterações em SPECT, e as características do critério para o diagnóstico de TCL, presença de citocinas inflamatórias, estresse oxidativo e histórico de DA na família, não estão representadas na base de treinamento do CDA. Seria interessante a utilização de uma base de casos clínicos com atributos que atendam na totalidade as características presentes nos critérios de diagnóstico destas doenças (DA e TCL). Este item será endereçado como trabalho futuro.

Tabela 15. Associação cruzada de características presentes nos critérios de diagnóstico de Demência, DA e TCL, com atributos da base CDA.

	MMSE ⁽²⁾	CDR	VFT	Pfeffer	CDT	TMT	Stroop	Lawton	IQCode	Berg	Depressão
Declínio gradual das funções cognitivas ⁽¹⁾	X	X	X		X	X	X				
Delírios											
Prejuízos cognitivos relatados pelo paciente				X				X	X		
Dificuldades de adquirir novas informações	X			X					X		
Dificuldades em lidar com tarefas complexas					X	X	X	X	X		
Dificuldade em utilizar raciocínio visual-espacial	X				X	X					
Dificuldade em utilizar elementos de linguagem	X		X								
Mudanças de personalidade e comportamento											
Modificações em testes de avaliação neuropsicológicas	X	X	X	X	X	X	X	X	X	X	
Alterações nos biomarcadores de líquido cerebral											
Diminuição da fluordesoxiglicose											
Atrofia no volume hipocampal											
Alterações em SPECT											
Presença de citocinas inflamatórias											
Estresse oxidativo											
Histórico de DA na família											

(1) Representa a característica presente nos critérios de diagnóstico de Demência, Doença de Alzheimer e Transtorno Cognitivo Leve, conforme Tabela 5.

(2) Atributos relacionados na base de casos clínicos do CDA .

Comparando a base de treinamento fornecida pelo CERAD, com a do CDA, a Tabela 16 mostra a associação em referência cruzada dos atributos presentes nestas duas bases (CERAD e CDA). Não foram identificadas associações com os atributos Etnia, *Blessed* e AVD da base CERAD em relação à base CDA. O atributo "Recordação de lista de palavras" seria associado com RAVLT da base CDA. Contudo, este exame neuropsicológico foi

excluído da base devido ao alto percentual de dados incompletos. Uma solução é conseguir uma outra base de casos clínicos com maior percentual de dados preenchidos.

A modelagem inicial da rede Bayesiana, incluindo a definição da estrutura, relação de variáveis aleatórias e avaliação do desempenho, foi apresentada em Seixas *et al.* (2011) e Seixas *et al.* (2012). Em Seixas *et al.* (2011) fundamentamos o modelo Bayesiano para diagnóstico clínico utilizando a estrutura genérica de rede Bayesiana (detalhado na Subseção 4.6.1) e, em Seixas *et al.* (2012), mostramos a modelagem da rede Bayesiana utilizando a base CERAD. Comparando os resultados iniciais publicados nos artigos, é possível notar um desempenho inferior ao apresentado nesta tese. As justificativas são: (1) consideramos inicialmente como subconjunto de treinamento da rede Bayesiana para diagnóstico de DA, apenas os casos de DA definitivo, ou seja, pacientes com comprovação histopatológica da doença por exame *post-mortem*. Tal critério reduziu significativamente o número de instâncias da base de treinamento. Nesta tese, estamos utilizando casos de provável DA, o que inclui pacientes diagnosticados com DA de acordo com os critérios publicados pelo NINCDS-ADRDA (vide Seção 4.2). (2) Nos artigos, não é executada a análise de ganho de informação detalhada na Seção 4.5.3. Esta análise teve como objetivo excluir atributos sem relevância para a classe considerada.

Tabela 16. Associação em referência cruzada dos atributos da base CERAD com os atributos da base CDA.

	MMSE ⁽²⁾	CDR	VFT	Pfeffer	CDT	TMT	Stroop	Lawton	IQCode	Berg	Depressão
CDR ⁽¹⁾		X									
Etnia											
MMSE	X										
Boston			X								
Blessed											
Déficits cognitivos	X	X	X		X	X			X		
AVD											
Recordação de lista de palavras											
Dificuldades no uso da linguagem			X								
Declínio gradual das funções cognitivas	X	X	X		X	X			X		
Habilidades funcionais					X						
Memória de lista de palavras	X			X					X		
Afasia progressiva primária	X		X								

(1) Atributo da base de casos clínicos do CERAD.

(2) Atributo da base de casos clínicos do CDA.

Por se tratar de uma modelagem orientada por dados, notamos uma forte dependência da qualidade da base de casos clínicos e seus atributos. É possível haver dois ou mais diagramas de influência para o diagnóstico da mesma doença, mas com um conjunto de variáveis aleatórias diferentes. O conjunto de variáveis aleatórias é formado pelas características observadas que deverão apoiar uma decisão. Estas características devem ser respaldadas pelos critérios de diagnóstico daquela doença, amplamente aceitos pela comunidade médica. O médico (usuário do sistema) deve escolher o diagrama de influência mais adequado ao seu contexto clínico. Comparando os atributos da base fornecida pelo CERAD com os atributos da base do CDA, é possível notar a coincidência de poucos atributos. Os atributos coincidentes são: Educação, Idade, CDR e MMSE. Ainda, para os atributos Educação e MMSE, a coincidência foi parcial, pois seus estados discretos representam diferentes intervalos numéricos.

Quanto à modelagem do CDSS, para incluir alguma característica ou critério a ser tratado no diagnóstico que não foi considerado na modelagem inicial da rede Bayesiana, há duas estratégias:

- (1) Incluir a variável aleatória “diretamente” na rede Bayesiana. Esta atividade pode ser feita utilizando alguma ferramenta de autoria da rede Bayesiana, como o aplicativo GeNIe/SMILE, ou reeditando o arquivo que contém os dados da rede Bayesiana. Se o arquivo for baseado em texto, sua modificação pode ser feita através de qualquer editor de texto. Além da variável aleatória, será necessário especificar a distribuição de probabilidades condicionais, ou tabela de probabilidades condicionais, o que nem sempre é uma tarefa trivial. Por exemplo, se a variável aleatória representar algum sintoma ou sinal, provavelmente ela será posicionada no terceiro nível da estrutura genérica. Neste caso, considerando variáveis aleatórias binomiais (dois estados discretos mutuamente exclusivos), teremos que especificar quatro probabilidades condicionais. Se aumentar o número de estados discretos, o número de probabilidades condicionais aumentará de forma exponencial (KORB e NICHOLSON, 2004).
- (2) Incluir a variável aleatória na rede Bayesiana, reunir um número representativo de casos clínicos em que a nova característica esteja presente, e reaplicar o método de aprendizagem supervisionada, no caso, EM. Este é o caminho mais recomendado, representado no processo da Figura 8 na etapa de revisão e refinamento da rede Bayesiana.

Representamos também o processo de diagnóstico das doenças de interesse através de um fluxo de trabalho (*workflow*), formatado em BPMN, e relacionamos univocamente as

caixas de decisão do diagrama de influência aos pontos de decisão do processo de diagnóstico. Em contraste com os trabalhos relacionados no Capítulo 2, a decisão modelada no CDSS proposto nesta tese é sequencial e não concorrente (ao contrário do sistema QMR/Internist-I) (MILLER, R.A. *et al.*, 1982). Então, a caixa de decisão modelada no diagrama de influência apresenta apenas dois estados, diagnóstico positivo e diagnóstico negativo, seguindo caminhos distintos devidamente representados no *workflow*. A formalização de fluxos de trabalho para o diagnóstico de diferentes doenças, considerando a modelagem BPMN, e associação dos pontos de decisão do fluxo às caixas de decisão do diagrama de influência, serão tratadas como trabalhos futuros.

No próximo capítulo, iremos detalhar os métodos de avaliação de desempenho e robustez dos modelos Bayesianos propostos.

CAPÍTULO 5 - AVALIAÇÃO DOS RESULTADOS

Este capítulo avalia o desempenho e a robustez da rede Bayesiana. Foram adotadas três estratégias para realizar a avaliação. A Seção 5.1 mostra os resultados de desempenho da rede Bayesiana, utilizando medidas de classificação. Os resultados da rede Bayesiana são apresentados para as bases de treinamento CERAD e CDA separadamente, e comparados com outros classificadores conhecidos. A Seção 5.2 descreve o meta classificador heurístico AdaBoostM1 e compara seus resultados em conjunto com os classificadores mostrados na seção anterior. A Seção 5.3 compara os resultados de desempenho, utilizando descoberta da estrutura de rede por métodos computacionais automáticos a partir dos dados. A Seção 5.4 avalia a robustez do modelo de decisão, utilizando análise de sensibilidade.

5.1 AVALIAÇÃO DO DESEMPENHO POR MEDIDAS DE CLASSIFICAÇÃO

A avaliação do desempenho da rede Bayesiana foi realizada através de quatro medidas, duas medidas baseadas na discriminação e duas medidas baseadas em erro. A avaliação é feita na distribuição *a posteriori* da variável aleatória associada ao diagnóstico da doença de interesse, ou variável aleatória central. Conforme mencionado anteriormente, esta variável aleatória possui dois estados, *Positivo* e *Negativo*, representando a sugestão de diagnóstico, e a probabilidade marginal *a posteriori*, associada ao estado da variável aleatória, representando a incerteza, dado o conjunto de evidências.

As medidas de desempenho baseadas na discriminação têm por objetivo avaliar como a rede Bayesiana discrimina entre dois possíveis estados, dadas as evidências (distribuição *a posteriori*). Utilizamos a área da curva ROC (*Receiver Operating Characteristics*), conhecida pela sigla AUC (*Area Under Curve*), e a medida F_1 (média harmônica da precisão e *recall*) como medida de desempenho. A regra de classificação Bayesiana é mostrada na Equação 34,

$$Z = \textit{positivo} \text{ se } P(z_{\textit{pos}} | E) = \max_{i=\textit{neg},\textit{pos}} P(z_i | E) \quad (34)$$

onde Z é a variável aleatória consultada, E representa o conjunto de evidências, e $P(z_i | E)$ é a probabilidade *a posteriori* estimada para cada estado i .

As medidas de desempenho baseadas no erro comparam o valor previsto com a probabilidade *a posteriori* estimada para a variável aleatória central. Utilizamos o erro quadrático médio (MSE – *Mean Squared Error*) e a entropia cruzada (MXE – *Mean Cross-*

Entropy) como medida de desempenho. A Tabela 17 mostra um resumo das medidas de desempenho utilizadas para avaliar as redes Bayesianas propostas.

A curva ROC mostra o gráfico da taxa de verdadeiros positivos (TPR) pela taxa de verdadeiros negativos (TNR) para todos os limiares possíveis de probabilidade da variável aleatória em análise assumir o estado *Positivo* (HANLEY, 1982). Assim, seja o número de instâncias positivas verdadeiras (TP – *True Positive*), instâncias negativas verdadeiras (TN – *True Negative*), instâncias positivas falsas (FP – *False Positive*) e instâncias negativas falsas (FN – *False Negative*), a Tabela 18 mostra uma representação matricial destes indicadores denominada matriz confusão, e as Equações 35 e 36 mostram a obtenção dos indicadores de sensibilidade e especificidade, respectivamente. A Figura 22 ilustra a área sob a curva ROC (AUC).

Tabela 17. Resumo das medidas de desempenho da rede Bayesiana.

Medida de desempenho	Abreviação	Intervalo	Melhor score
Área da curva do ROC	AUC	[0, 1]	1
Média harmônica da precisão e <i>recall</i>	F ₁	[0, 1]	1
Erro médio quadrado	MSE	[0, 1]	0
Entropia cruzada	MXE	[0, ∞)	0

Tabela 18. Matriz confusão.

		Real	
		<i>Positivo</i>	<i>Negativo</i>
Classe prevista	<i>Positivo</i>	<i>TP</i>	<i>FP</i>
	<i>Negativo</i>	<i>FN</i>	<i>TN</i>

$$TPR = \frac{TP}{TP + FN} \quad (35)$$

$$TNR = \frac{TN}{TN + FP} \quad (36)$$

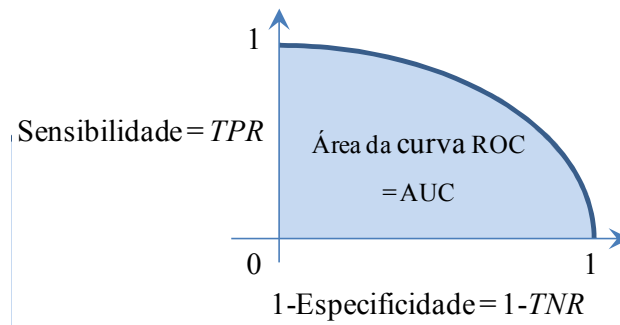


Figura 22. Curva ROC e a área sob a curva ROC (AUC)

A medida F_1 considera precisão e *recall*. Precisão é definida como medida de acurácia para uma determinada classe (Equação 37). *Recall* é definido como medida da capacidade do modelo inferir instâncias de uma determinada classe, com base em um conjunto de dados. Para variáveis aleatórias binomiais (dois estados: *Positivo* e *Negativo*), *recall* apresenta os mesmos resultados do TPR (Equação 35). Por fim, a medida F_1 pode ser utilizada como medida única, como mostrado na Equação 38 (OLSON e DELEN, 2008).

$$\text{precisão} = \frac{TP}{TP + FP} \quad (37)$$

$$F_1 = 2 \cdot \frac{\text{precisão} \cdot \text{recall}}{\text{precisão} + \text{recall}} \quad (38)$$

Seja Y a variável aleatória central, y_i a classificação de cada instância i . Atribuindo $y_i = 0$ para estado *Negativo*, e $y_i = 1$ para estado *Positivo*, e seja $P(x = 1)$ a probabilidade *a posteriori* de y_i assumir o valor 1. Então, o erro quadrático médio para n instâncias (BRIER, 1950) é dado pela Equação 38.

$$MSE = \frac{1}{n} \sum_{i=1}^n [y_i - P(x=1)]^2 \quad (39)$$

A entropia cruzada média para n instâncias é obtida pela Equação 39 (WITTEN e FRANK, 2005).

$$MXE = \frac{1}{n} \sum_{i=1}^n -y_i \cdot \log[P(x=1)] - (1-y_i) \cdot \log[1-P(x=1)] \quad (40)$$

Conforme exposto, para a validação, o conjunto de dados é subdividido em dois subconjuntos: um de treinamento e outro de teste. Utilizamos o método conhecido por validação cruzada (KOHAVI, 1995b). No método de validação cruzada, a amostra é estratificada em k subconjuntos (*folds*) de igual tamanho. O treinamento é repetido k vezes, cada um, utilizando um subconjunto para teste e $k-1$ subconjuntos para o treinamento

propriamente dito. O objetivo de utilizar dois conjuntos de dados (um de treinamento e outro de teste) mutuamente exclusivos é evitar o superajuste (*overfitting*) decorrente do processo de aprendizagem, bem como a inconsistência da validação do modelo (BELLAZZI e ZUPAN, 2008). Cada subconjunto deve conter o mesmo percentual de classes em relação ao conjunto total de instâncias. Utilizamos validação cruzada com 4 subconjuntos (*folds*), devido ao tamanho total de instâncias do conjunto de treinamento (no caso da base CDA, 247 instâncias, conforme mostra a Tabela 7) (KOHAVI, 1995b).

A Tabela 19 apresenta os resultados das medidas de desempenho para cada *fold*, utilizando a rede Bayesiana modelada pela base CERAD. Para Demência, a área de curva ROC média (AUC) foi de 0,82. Para DA, AUC foi de 0,67. Para medida F_1 , os valores obtidos foram 0,78 e 0,65, para diagnóstico de Demência e DA, respectivamente. As medidas de TPR e TNR, no caso de Demência, foram de 0,81 e 0,69, respectivamente. No caso de DA, os valores de TPR e TNR foram de 0,66 e 0,57, respectivamente. Estes resultados, incluindo os resultados obtidos para AUC, representam um desempenho regular da rede Bayesiana utilizada para o diagnóstico de Demência e DA, devido a sua aproximação de 0,50 (considerado o pior valor para TPR e TNR). Uma forma de melhorar o desempenho é considerar atributos mais sensíveis (ou relevantes) ao diagnóstico de DA, ou considerar uma base de treinamento mais representativa. Em relação ao MSE e MXE, os resultados foram 0,11 e 0,24 para Demência, e 0,25 e 0,30 para DA, respectivamente.

Tabela 19. Medidas de desempenho e valores da matriz confusão obtidos em cada *fold* considerando a rede Bayesiana obtida pela base CERAD.

Doença	Fold ⁽²⁾	AUC ^(3†)	F ₁ ^(†)	TPR ^(†)	TNR ^(†)	TP	FP	TN	FN	MSE ^(↓)	MXE ^(↓)
Demência	1	0,82	0,79	0,84	0,67	179	58	119	35	0,17	0,23
	2	0,84	0,81	0,82	0,76	171	43	134	37	0,17	0,22
	3	0,84	0,78	0,81	0,70	171	56	130	40	0,16	0,21
	4	0,79	0,74	0,77	0,62	160	67	109	48	0,20	0,27
	Média (DP) ⁽¹⁾	0,82 (0,02)	0,78 (0,03)	0,81 (0,03)	0,69 (0,06)	-	-	-	-	-	0,18 (0,02)
Doença de Alzheimer	1	0,66	0,66	0,69	0,54	97	57	68	44	0,24	0,30
	2	0,67	0,64	0,62	0,65	87	44	81	54	0,23	0,29
	3	0,67	0,67	0,71	0,52	100	60	66	40	0,23	0,29
	4	0,68	0,63	0,64	0,56	89	55	71	51	0,23	0,29
	Média (DP)	0,67 (0,01)	0,65 (0,02)	0,66 (0,05)	0,57 (0,05)	-	-	-	-	-	0,23 (0,00)

(1) Média (DP) = média aritmética (desvio padrão).

(2) *Fold* = estrato do método de validação cruzado (*Cross Validation*).

(3) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F₁ = medida F₁, melhor escore é 1. TPR = Sensibilidade, ou taxa de verdadeiros positivos, melhor escore é 1. TNR = Especificidade, ou taxa de verdadeiros falsos, melhor escore é 1. TP = quantidade de verdadeiros positivos. FP = quantidade de falsos positivos. TN = quantidade de verdadeiros negativos. FN = quantidade de falsos negativos. MSE = Erro médio quadrado (*Mean Squared Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

A Tabela 21 mostra as medidas de desempenho da rede Bayesiana obtida da base CERAD, comparado com o desempenho de outros classificadores conhecidos. O objetivo da comparação é verificar se os resultados apresentados pela rede Bayesiana são compatíveis com outros classificadores. Para comparação com a rede Bayesiana, foram selecionados os seguintes classificadores, todos implementados na ferramenta WEKA (*Waikato Environment for Knowledge Analysis*), desenvolvimento pela Universidade de Waikato: (1) *Näive Bayes* (JOHN, G. H. e LANGLEY, 1995), (2) modelo de regressão logístico com estimador de crista (*Logistic*) (LE CESSIE e VAN HOUWELINGEN, 1992), (3) redes neurais Perceptron multicamadas retropropagação (*Multilayer Perceptron*) (WITTEN e FRANK, 2005), (4) tabela de decisão (*Decision Table*) (KOHAVI, 1995a), (5) árvore de decisão de um nível (*decision stump*) (IBA e LANGLEY, 1992), (6) árvore de decisão C4.5 (*J48*) (QUINLAN, 1993). Utilizamos validação cruzada com 4 *folds*, e os mesmos subconjuntos de treinamento e teste utilizados na avaliação da rede Bayesiana. A Tabela 20 lista os principais parâmetros utilizados nesses classificadores. A Tabela 21 mostra os resultados, resumidos pela média aritmética e desvio padrão obtidos pelas medidas de desempenho em relação aos *folds*.

Tabela 20. Principais parâmetros dos classificadores utilizados na comparação.

Classificador	Descrição dos Parâmetros
1. Nãive Bayes	Não utilizar estimador de núcleo (<i>kernel</i>) para os atributos numéricos, em vez de distribuição normal.
	Não utilizar discretização supervisionada para converter atributos numéricos para valores nominais.
2. Modelo de regressão logístico (<i>Logistic</i>)	Observação: utiliza um modelo de regressão logística com estimador de crista. Neste classificador, os valores incompletos ou ausentes são substituídos por valores nominais, de modo não alterar as características da distribuição, e os atributos nominais são transformados em atributos numéricos. Número máximo de iterações = 1
	Valor ridge na operação de log-verossimilhança = 10^{-8}
3. Redes neurais multicamadas Perceptron (<i>Multilayer Perceptron</i>)	Não adicionar automaticamente camadas escondidas da rede.
	Não reduzir a taxa de aprendizagem proporcionalmente ao número de rodadas.
	Quantidade de camadas escondidas na rede neural = (quantidade de atributos + quantidade de classes) / 2
	Taxa de aprendizagem inicial = 0,3
	<i>Momentum</i> aplicado aos pesos sinápticos durante a atualização = 0,2
	Converter atributos nominais para atributos numéricos. Ajuda a melhorar o desempenho do modelo.
	Normalizar os atributos nominais para valores entre -1 e 1. Ajuda a melhorar o desempenho do modelo.
	Caso a classe for numérica, normalizar classe. Ajuda a melhorar o desempenho do modelo.
	Se o desempenho do modelo divergir do esperado ao longo das rodadas (processo iterativo), então reinicializar a rede.
	Numero máximo de rodadas = 500
4. Tabela de decisão (<i>Decision table</i>)	Utilizar método de validação <i>leave one out</i> .
	Utilizar acurácia como medida de avaliação do desempenho da combinação de atributos para tabela de decisão.
	Método de busca utilizado para encontrar as melhores combinações de atributos para tabela de decisão é o <i>BestFirst</i> . O método <i>BestFirst</i> utiliza uma heurística gulosa (<i>greedy hillclimbing</i>), e retrocesso (<i>backtracking</i>).
	Direção do método busca para frente.
	Número máximo utilizado para retrocesso no método de busca = 5
5. Arvore de decisão de um nível (<i>Decision stump</i>)	Não há parâmetros de configuração neste classificador.
6. Árvore de decisão J48	Não utilizar subdivisões binárias nos atributos nominais na construção das árvores.
	Fator de confiança utilizada no algoritmo de poda = 0,25
	Número mínimo de instâncias por folha = 2
	Quantidade de <i>folds</i> utilizados para reduzir o erro no algoritmo de poda = 3
	Considerar a criação de subárvores no algoritmo de poda.

Tabela 21. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CERAD.

Doença	Modelo	AUC ^(2↑)	F ₁ ^(↑)	TPR ^(↑)	TNR ^(↑)	MSE ^(↓)	MXE ^(↓)
Demência	Rede Bayesiana ⁽¹⁾	0,82 (0,02)	0,78 (0,03)	0,81 (0,03)	0,69 (0,06)	0,18 (0,02)	0,24 (0,02)
	1. <i>Náive Bayes</i>	0,82 (0,02)	0,78 (0,02)	0,81 (0,02)	0,69 (0,04)	0,24 (0,02)	0,18 (0,02)
	2. <i>Logistic</i>	0,82 (0,03)	0,78 (0,03)	0,82 (0,03)	0,66 (0,04)	0,22 (0,02)	0,17 (0,01)
	3. <i>Multilayer perceptron</i>	0,84 (0,02)	0,78 (0,02)	0,76 (0,04)	0,78 (0,06)	0,65 (0,07)	0,20 (0,01)
	4. <i>Decision table</i>	0,82 (0,01)	0,79 (0,02)	0,88 (0,05)	0,61 (0,05)	0,21 (0,01)	0,16 (0,01)
	5. <i>Decision stump</i>	0,73 (0,03)	0,79 (0,02)	0,89 (0,03)	0,58 (0,03)	0,19 (0,01)	0,24 (0,01)
	6. <i>J48</i>	0,81 (0,02)	0,78 (0,01)	0,84 (0,01)	0,64 (0,05)	0,23 (0,01)	0,17 (0,01)
Doença de Alzheimer	Rede Bayesiana	0,67 (0,01)	0,65 (0,02)	0,66 (0,05)	0,57 (0,05)	0,23 (0,00)	0,29 (0,00)
	1. <i>Náive Bayes</i>	0,67 (0,02)	0,64 (0,01)	0,66 (0,04)	0,57 (0,05)	0,29 (0,01)	0,23 (0,01)
	2. <i>Logistic</i>	0,66 (0,02)	0,65 (0,03)	0,66 (0,03)	0,57 (0,07)	0,28 (0,01)	0,23 (0,01)
	3. <i>Multilayer perceptron</i>	0,71 (0,03)	0,66 (0,03)	0,66 (0,04)	0,63 (0,03)	0,83 (0,10)	0,31 (0,03)
	4. <i>Decision Table</i>	0,70 (0,03)	0,66 (0,01)	0,68 (0,03)	0,59 (0,09)	0,27 (0,01)	0,22 (0,01)
	5. <i>Decision stump</i>	0,62 (0,04)	0,57 (0,04)	0,49 (0,04)	0,74 (0,05)	0,24 (0,01)	0,29 (0,01)
	6. <i>J48</i>	0,67 (0,02)	0,67 (0,02)	0,71 (0,03)	0,56 (0,05)	0,27 (0,01)	0,24 (0,01)

(1) Média (DP) = média aritmética (desvio padrão). Obtido considerando 4 *folds* da validação cruzada.

(2) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F₁ = medida F₁, melhor escore é 1. TPR = Sensibilidade ou taxa de verdadeiros positivos, melhor escore é 1. TNR = Especificidade ou taxa de verdadeiros falsos, melhor escore é 1. MSE = Erro médio quadrado (*Mean Square Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

Analisando os modelos de classificação obtidos para Demência e DA simultaneamente, os melhores AUC foram obtidos pelo classificador baseado em redes neurais multicamadas Perceptron (0,84 e 0,71). Do mesmo modo, os melhores valores para medida F₁ foram obtidos pelo classificador baseado em tabela de decisão (0,79 e 0,66). Os melhores valores de MSE foram obtidos pela rede Bayesiana (0,18 e 0,23). E no caso do MXE, pelo classificador baseado em tabela de decisão (0,16 e 0,22, respectivamente). Em geral, foram obtidos valores de TNR levemente menores que o TPR, mostrando uma dificuldade dos classificadores distinguirem casos negativos das doenças em questão (Demência e DA). Pode-se notar que nenhum dos classificadores obtém o melhor desempenho em todas as métricas analisadas e que o desempenho da rede Bayesiana é comparável aos demais classificadores.

A Tabela 22 apresenta os resultados das medidas de desempenho utilizando a rede Bayesiana modelada pela base CDA. Obtivemos uma área da curva ROC média (AUC) de 0,96 para Demência. No caso de diagnóstico de DA, AUC foi de 0,86. Os melhores resultados

de AUC foram obtidos para o diagnóstico de TCL, com 0,97, o que mostra o potencial da rede Bayesiana para auxiliar o diagnóstico de um estágio pré-clínico de DA. Já no caso da medida F_1 , os melhores resultados foram para diagnóstico de Demência (0,94). Os melhores resultados de TPR foram para o diagnóstico de Demência e TCL (0,94). Já para o TNR, o melhor resultado foi para o diagnóstico de Demência. Em relação ao MSE e MXE, os melhores resultados foram obtidos para o diagnóstico de Demência (0,07 e 0,16, respectivamente).

Tabela 22. Medidas de desempenho e valores da matriz confusão obtidos em cada *fold* considerando a rede Bayesiana obtida pela base CDA.

Doença	Fold ⁽²⁾	AUC ^(3↑)	F_1 ^(↑)	TPR ^(↑)	TNR ^(↑)	TP	FP	TN	FN	MSE ^(↓)	MXE ^(↓)
Demência	1	0,99	0,99	0,97	1,00	36	0	26	1	0,02	0,05
	2	0,93	0,88	0,89	0,82	33	5	23	4	0,13	0,24
	3	0,96	0,95	0,94	0,96	31	1	25	2	0,05	0,05
	4	0,96	0,92	0,97	0,79	35	5	19	1	0,08	0,30
	Média (DP) ⁽¹⁾	0,96 (0,03)	0,94 (0,05)	0,94 (0,04)	0,89 (0,10)	-	-	-	-	0,07 (0,04)	0,16 (0,13)
Doença de Alzheimer	1	0,83	0,82	0,82	0,74	23	5	14	5	0,16	0,18
	2	0,87	0,83	0,90	0,58	26	8	11	3	0,15	0,19
	3	0,82	0,83	0,81	0,76	22	4	13	5	0,16	0,38
	4	0,91	0,81	0,83	0,72	19	5	13	4	0,13	0,17
	Média (DP)	0,86 (0,04)	0,82 (0,01)	0,84 (0,04)	0,70 (0,08)	-	-	-	-	0,15 (0,02)	0,23 (0,10)
Transtorno cognitivo leve	1	0,90	0,84	0,89	0,75	8	2	6	1	0,18	0,81
	2	1,00	1,00	1,00	1,00	9	0	8	0	0,01	0,02
	3	0,97	0,93	0,88	1,00	7	0	8	1	0,05	0,08
	4	1,00	0,90	1,00	0,75	9	2	6	0	0,12	0,41
	Média (DP)	0,97 (0,05)	0,92 (0,07)	0,94 (0,07)	0,88 (0,14)	-	-	-	-	0,09 (0,14)	0,33 (0,36)

(1) Média (DP) = média aritmética (desvio padrão).

(2) Fold = estrato do método de validação cruzado (*Cross Validation*).

(3) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F_1 = medida F_1 , melhor escore é 1. TPR = Sensibilidade, ou taxa de verdadeiros positivos, melhor escore é 1. TNR = Especificidade, ou taxa de verdadeiros falsos, melhor escore é 1. TP = quantidade de verdadeiros positivos. FP = quantidade de falsos positivos. TN = quantidade de verdadeiros negativos. FN = quantidade de falsos negativos. MSE = Erro médio quadrado (*Mean Squared Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

A Tabela 23 mostra as medidas de desempenho da rede Bayesiana obtida da base CDA, comparadas com o desempenho de outros classificadores conhecidos, os mesmos utilizados na comparação anterior, com parâmetros listados na Tabela 20. Também utilizamos

validação cruzada com 4 *folds*, e os mesmos subconjuntos de treinamento e teste utilizados na avaliação da rede Bayesiana. No caso da Demência, o classificador baseado em tabela de decisão mostrou os melhores resultados para AUC (0,98), medida- F_1 (0,95), e TPR (0,97). Para TNR, o melhor valor foi obtido pelo NãiveBayes (0,93). Para MSE e MXE, o melhor resultado foi obtido pelo classificador baseado em tabela de decisão (0,05).

Tabela 23. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CDA.

Doença	Modelo	AUC ^(2↑)	F_1 ^(↑)	TPR ^(↑)	TNR ^(↑)	MSE ^(↓)	MXE ^(↓)
Demência	Rede Bayesiana ⁽¹⁾	0,96 (0,03)	0,94 (0,05)	0,94 (0,04)	0,89 (0,10)	0,18 (0,04)	0,16 (0,13)
	1. NãiveBayes	0,97 (0,04)	0,93 (0,05)	0,91 (0,04)	0,93 (0,09)	0,13 (0,11)	0,07 (0,06)
	2. Logistic	0,97 (0,04)	0,94 (0,03)	0,93 (0,02)	0,92 (0,11)	0,11 (0,06)	0,07 (0,03)
	3. Multilayer perceptron	0,96 (0,05)	0,92 (0,02)	0,90 (0,05)	0,92 (0,11)	0,20 (0,15)	0,08 (0,03)
	4. Decision table	0,98 (0,03)	0,95 (0,04)	0,97 (0,00)	0,90 (0,01)	0,09 (0,04)	0,05 (0,03)
	5. Decision stump	0,88 (0,08)	0,90 (0,04)	0,95 (0,01)	0,78 (0,15)	0,10 (0,05)	0,15 (0,06)
	6. J48	0,96 (0,04)	0,95 (0,06)	0,97 (0,05)	0,89 (0,11)	0,09 (0,05)	0,06 (0,04)
Doença de Alzheimer	Rede Bayesiana	0,86 (0,04)	0,82 (0,01)	0,84 (0,04)	0,70 (0,08)	0,15 (0,02)	0,23 (0,10)
	1. NãiveBayes	0,85 (0,03)	0,80 (0,02)	0,81 (0,04)	0,70 (0,08)	0,21 (0,02)	0,16 (0,01)
	2. Logistic	0,82 (0,03)	0,81 (0,02)	0,82 (0,01)	0,71 (0,06)	0,25 (0,08)	0,17 (0,02)
	3. Multilayer perceptron	0,90 (0,03)	0,86 (0,03)	0,84 (0,05)	0,82 (0,06)	0,29 (0,12)	0,14 (0,03)
	4. Decision table	0,87 (0,06)	0,77 (0,09)	0,69 (0,12)	0,87 (0,08)	0,20 (0,04)	0,15 (0,04)
	5. Decision stump	0,80 (0,01)	0,76 (0,01)	0,61 (0,01)	1,00 (0,00)	0,15 (0,00)	0,18 (0,00)
	6. J48	0,76 (0,06)	0,79 (0,04)	0,83 (0,07)	0,63 (0,07)	0,24 (0,04)	0,18 (0,03)
Transtorno cognitivo leve	Rede Bayesiana	0,97 (0,05)	0,92 (0,07)	0,94 (0,07)	0,88 (0,14)	0,09 (0,08)	0,33 (0,36)
	1. NãiveBayes	0,97 (0,03)	0,90 (0,07)	0,85 (0,06)	0,96 (0,07)	0,15 (0,13)	0,09 (0,05)
	2. Logistic	0,76 (0,10)	0,69 (0,08)	0,69 (0,05)	0,63 (0,22)	0,00 (0,00)	0,34 (0,11)
	3. Multilayer perceptron	0,90 (0,08)	0,79 (0,04)	0,77 (0,02)	0,79 (0,14)	0,37 (0,12)	0,19 (0,03)
	4. Decision table	0,82 (0,16)	0,78 (0,15)	0,74 (0,16)	0,83 (0,14)	0,21 (0,09)	0,16 (0,09)
	5. Decision stump	0,92 (0,04)	0,91 (0,06)	0,91 (0,11)	0,91 (0,06)	0,09 (0,05)	0,12 (0,08)
	6. J48	0,79 (0,09)	0,72 (0,15)	0,61 (0,24)	0,96 (0,07)	0,21 (0,05)	0,17 (0,04)

(1) Média (DP) = média aritmética (desvio padrão). Obtido considerando 4 *folds* da validação cruzada.

(2) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F_1 = medida F_1 , melhor escore é 1. TPR = Sensibilidade ou taxa de verdadeiros positivos, melhor escore é 1. TNR = Especificidade ou taxa de verdadeiros falsos, melhor escore é 1. MSE = Erro médio quadrado (*Mean Square Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

No caso de DA, os melhores resultados para AUC e medida- F_1 foram obtidos pelo classificador baseado em redes neurais Perceptron (0,90 e 0,86, respectivamente). No caso da TPR, os melhores resultados foram obtidos pela rede Bayesiana (0,84) e o classificador

baseada em redes neurais Perceptron (0,84). Já o TNR, o melhor resultado ficou para o classificador baseado em árvore de decisão de um nível denominado *Decision stump*. Entretanto, este resultado não é consistente, pois este mesmo classificador obteve o pior TPR (0,61). O segundo melhor resultado para TNR foi obtido pelo classificador baseado em tabela de decisão (0,87). O melhor resultado para MSE foi obtido pela rede Bayesiana e pelo *Decision stump* e MXE pelo classificador baseado em redes neurais Perceptron (0,14).

No caso do TCL, os melhores resultados para AUC, medida- F_1 e TPR foram obtidos pela rede Bayesiana (0,97, 0,92 e 0,94, respectivamente). Para TNR, os melhores resultados foram obtidos pelo classificador NãiveBayes e árvore de decisão J48 (0,96). Para MSE, o melhor resultado foi obtido pelo classificador de regressão logística. E para MXE, o melhor resultado foi obtido pelo classificador NãiveBayes. Este resultado mostra o potencial do modelo Bayesiano na identificação de TCL, que pode estar relacionada a um estágio pré-clínico da DA. A identificação precoce de DA pode antecipar o tratamento do paciente e melhorar sua qualidade de vida.

De uma forma geral, a comparação dos resultados de desempenho da rede Bayesiana com outros classificadores teve como objetivo principal comprovar que seus resultados são compatíveis aos de outros classificadores conhecidos. Apesar de nem sempre mostrar os melhores resultados de desempenho, há outras características que diferenciam o modelo Bayesiano de outros modelos de classificação. Uma característica desejável em um modelo de decisão é a capacidade de interpretação pelo especialista dos critérios que o envolvem (HART, A. e WYATT, 1990). A apresentação simplesmente do resultado de diagnóstico junto com uma incerteza obtida pelo mecanismo de inferência é, muitas vezes, insuficiente para o médico. A decisão clínica deve ser justificada com base em critérios ou argumentos, coerentes com o conhecimento tácito e bibliografia médica. Na rede Bayesiana, as variáveis que compõem o modelo de decisão e seus relacionamentos causais ficam explicitamente representadas em um diagrama gráfico, o que facilita a interpretação do modelo de decisão, tanto por especialistas do domínio do conhecimento, como por profissionais da área da computação e aplicativos computacionais de inferência. A capacidade de interpretação do modelo de decisão facilita também sua revisão contínua, modificando as variáveis existentes ou incorporando novos achados da área médica (BELLAZZI e ZUPAN, 2008; BIN *et al.*, 1999; KORB e NICHOLSON, 2004). A Seção 2.5 relaciona outras características diferenciais do modelo Bayesiano que justificaram sua escolha como modelo de decisão adotado nesta tese. Um dos objetivos deste trabalho foi representar o conhecimento médico e os critérios de diagnóstico da doença de interesse no modelo de decisão a ser proposto e não

simplesmente testar diversos classificadores em uma base de dados para verificar qual deles resulta em melhor desempenho. A comparação dos resultados de desempenho da rede Bayesiana com outros classificadores teve como objetivo mostrar que o modelo Bayesiano proposto apresenta resultados próximos e compatíveis com os demais.

5.2 AVALIAÇÃO UTILIZANDO DESCOBERTA DA ESTRUTURA DA REDE

Outra possibilidade aliada ao uso de redes Bayesianas é a definição automática da estrutura da rede descoberta a partir de uma base de dados, aplicando um método computacional. Para fins de comparação com os resultados obtidos pela proposta deste trabalho, de redes Bayesianas definidas conforme o conhecimento do domínio, também utilizamos um método automático para definição de redes Bayesianas. As Figuras 23, 24 e 25 mostram a estrutura da Rede Bayesiana descoberta de maneira automatizada. Utilizamos para esta análise a base de casos clínicos fornecida pelo CDA, e comparamos os resultados com o obtido pela aprendizagem dos parâmetros da rede Bayesiana apresentados na seção anterior. O método computacional de descoberta da estrutura e aprendizagem dos parâmetros é baseado em teste de independência, descrito por Dash e Druzdzel (2003), aplicando um método de busca gulosa denominado *Greedy Thick Thinning* (JONGSAWAT e PREMCHAIWADI, 2009). A Tabela 24 apresenta os resultados de desempenho das redes descobertas.

Em comparação com os resultados apresentados na Tabela 22, notamos que, de uma forma em geral, a rede Bayesiana descoberta apresentou resultados de desempenho bem próximos em relação à rede Bayesiana com estrutura previamente definida. Para medida AUC, a rede descoberta obteve 0,94 para Demência, indicando um resultado levemente inferior a rede com estrutura definida manualmente (0,96). Para diagnóstico de DA, os resultados foram superiores com a rede descoberta automaticamente para AUC, 0,89 contra 0,86. Em relação ao diagnóstico de TCL, os resultados foram bem próximos. Uma desvantagem do modelo com estrutura descoberta automaticamente é a complexidade das conexões causais, dificultando sua interpretação pelos especialistas médicos, refinamentos e revisões, por exemplo, motivadas por novos casos clínicos.

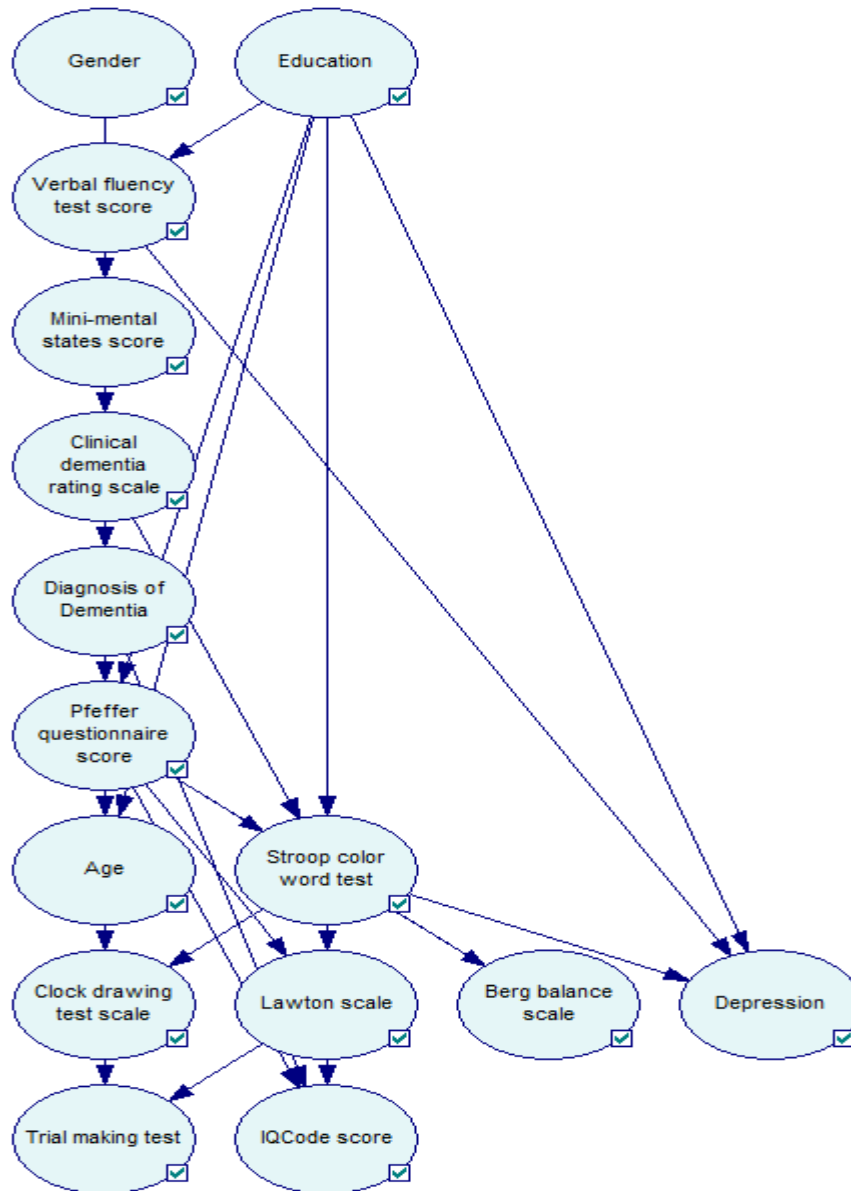


Figura 23. Rede Bayesiana descoberta por método computacional aplicada ao diagnóstico de Demência, utilizando base CDA.

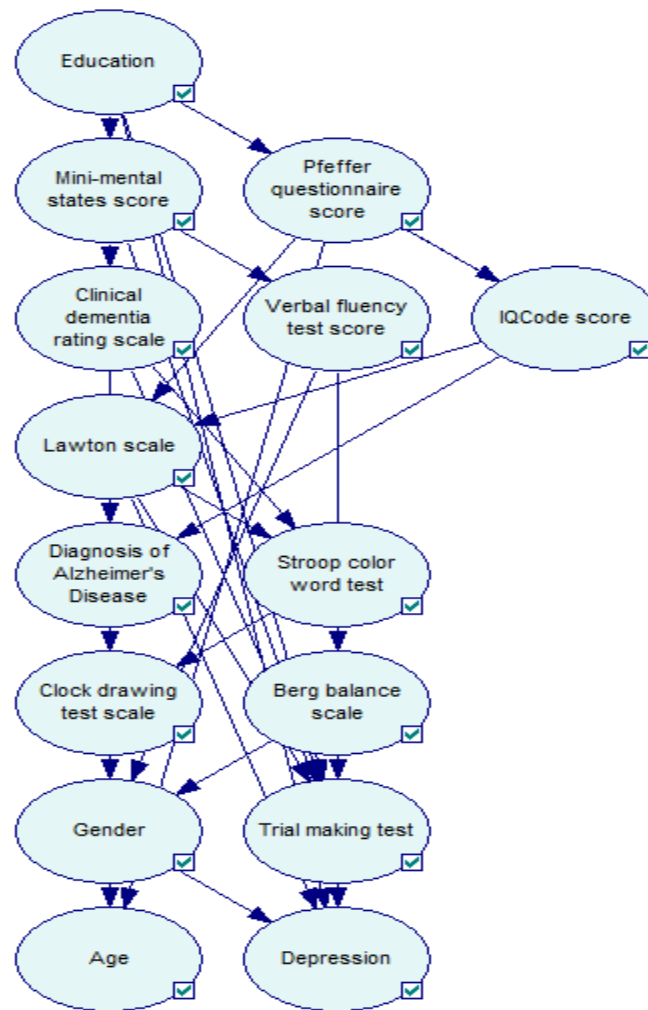


Figura 24. Rede Bayesiana descoberta por método computacional aplicada ao diagnóstico de DA, utilizando base CDA.

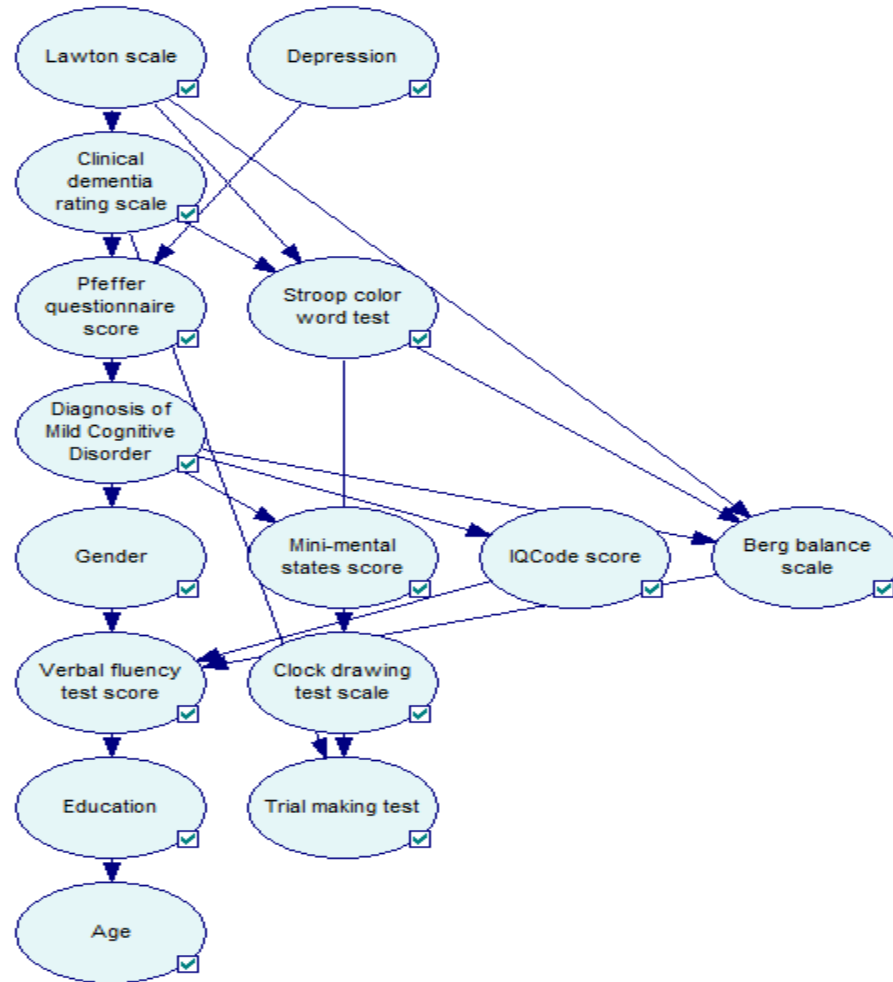


Figura 25. Rede Bayesiana descoberta por método computacional aplicado ao diagnóstico de TCL, utilizando base CDA.

Tabela 24. Desempenho apresentado pela rede Bayesiana descoberta por método computacional para o diagnóstico da doença de interesse, utilizando base CDA.

Doença	AUC ^(2↑)	F1 ^(↑)	TPR ^(↑)	TNR ^(↑)	MSE ^(↓)	MXE ^(↓)
Demência	0,94 (0,03) ⁽¹⁾	0,80 (0,05)	0,75 (0,11)	0,96 (0,04)	0,08 (0,02)	0,12 (0,03)
Doença de Alzheimer	0,89 (0,04)	0,87 (0,04)	0,91 (0,05)	0,46 (0,14)	0,12 (0,01)	0,16 (0,02)
Transtorno Cognitivo Leve	0,96 (0,03)	0,86 (0,07)	0,91 (0,12)	0,78 (0,19)	0,10 (0,04)	0,15 (0,07)

(1) Média (DP) = média aritmética (desvio padrão). Obtido considerando 4 *folds* da validação cruzada.

(2) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F₁ = medida F₁, melhor escore é 1. TPR = sensibilidade ou taxa de verdadeiros positivos, melhor escore é 1. TNR = especificidade ou taxa de verdadeiros falsos, melhor escore é 1. MSE = Erro médio quadrado (*Mean Squared Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

5.3 COMPARAÇÃO DO DESEMPENHO COM O ADABOOST

Um conjunto de algoritmos de aprendizagem computacional que utilizam alguma meta-heurística para melhorar o desempenho de outros classificadores é denominado *Boosting*. O *AdaBoosting*, nome derivado de *Adaptive Boosting* (em português, impulso ou estímulo adaptativo), formulado por Freund e Schapire (1995), consiste em um algoritmo *Boosting* generalista, com parâmetros obtidos por aprendizagem. O método de aprendizagem definido como forte é descrito como uma combinação linear de aprendizagens definidas como fracas. Assim, seja uma base de treinamento contendo m instâncias, ou $S = \langle (x_1, y_1), \dots, (x_m, y_m) \rangle$, onde $x_i \in X$ é representada por um vetor de valores de atributos da instância i , e $y_i \in Y$ o rótulo da classe associada a x_i . O método é iniciado através da seguinte sequência: (1) preenchimento do vetor de ponderação $D_1(i) = \frac{1}{m}$, (2) execução do método de aprendizagem fraca e (3) avaliação dos seus resultados. Para as instâncias classificadas incorretamente, é atribuído um peso igual a 1, e as instâncias classificadas corretamente um peso igual a β_t , que deve ser menor que 1. Estes valores são utilizados para atualizar o vetor de ponderação $D_t(i)$. A ideia é que, a cada iteração, para os candidatos classificados incorretamente, o peso seja incrementado, para que a aprendizagem priorize estas instâncias em detrimento das instâncias classificadas corretamente. Assim, a função da aprendizagem fraca é encontrar uma hipótese fraca $h_t : X \rightarrow Y$, e a função da aprendizagem forte é encontrar uma distribuição D_t mais apropriada para hipótese fraca. Há duas versões do algoritmo *AdaBoost*: o *AdaBoostM1* e o *AdaBoostM2*. A Figura 26 descreve as etapas do algoritmo *AdaBoostM1*. A cada iteração, a distribuição de pesos D_t é atualizada, até que seja alcançado algum critério de parada. A hipótese final (h_{fin}) representa a soma ponderada das hipóteses fracas, ou seja, para uma dada instância x , h_{fin} atribui uma classe y , de forma a maximizar a soma de pesos D_t das hipóteses fracas que predizem aquela classe.

O algoritmo *AdaBoostM2* é similar ao *AdaBoostM1*, com a diferença que procura estender os resultados obtidos do método de aprendizagem fraca, como, por exemplo, a utilização da plausibilidade ou grau de certeza que a classificação da instância é y . Outra diferença, em vez de avaliar a aprendizagem fraca através do erro de predição (ϵ_t), o *AdaBoostM2* calcula uma função pseudo-prejuízo em relação a distribuição sobre todos os pares de exemplos e classificações incorretas. Com isso, o algoritmo pode direcionar a aprendizagem fraca não somente para os exemplos difíceis de classificar, mas nas

classificações incorretas que são mais difíceis de discriminar. O objetivo do algoritmo é encontrar o conjunto de hipóteses fracas h_t com a menor função de pseudoprejuízo.

<p>Entradas:</p> <ol style="list-style-type: none"> 1. Sequência de m exemplos $S = \langle (x_1, y_1), \dots, (x_m, y_m) \rangle$, com classes $y_i \in Y = \{1, \dots, k\}$ 2. Método de aprendizagem fraca 3. Inteiro T especificando o número máximo de iterações <p>Inicialização:</p> $D_1(i) = \frac{1}{m}$ <p>Para cada iteração $t = 1, 2, \dots, T$:</p> <ol style="list-style-type: none"> 1. Executar método de aprendizagem fraca e avaliar os resultados das classificações com a distribuição D_t. 2. Admitir a hipótese $h_t : X \rightarrow Y$ 3. Calcular o erro da hipótese $h_t : \varepsilon_t = \sum_{i: h_t(x_i) \neq y_i} D_t(i)$. Se $\varepsilon_t > 1/2$, então $T = t-1$ e interromper o processo iterativo. 4. Calcular $\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$ 5. Atualizar a distribuição $D_t : D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} \beta_t & \text{se } h_t(x_i) = y_i \\ 1, & \text{caso contrário} \end{cases}$, onde Z_t é uma constante de normalização calculada de tal forma que D_{t+1} seja uma distribuição. <p>Saída:</p> <p>A hipótese final $h_{fin}(x) = \arg \max_{y \in Y} \sum_{t: h_t(x)=y} \log \frac{1}{\beta_t}$</p>
--

Figura 26. Descrição do algoritmo meta classificador AdaBoostM1.

Avaliamos nesta tese o algoritmo *AdaBoostM1*, comparando o desempenho combinado com os classificadores utilizados na comparação com a rede Bayesiana, com parâmetros listados na Tabela 20. Utilizamos para comparação as medidas de desempenho AUC (área da curva ROC) e MSE (erro médio quadrado). Utilizamos a implementação do *AdaBoostM1* disponível no WEKA. O objetivo da avaliação é comparar os resultados de desempenho do classificador isoladamente contra o mesmo classificador em conjunto com *AdaBoost*. Utilizamos na verificação o teste estatístico de hipótese nula, com nível de significância de 5%. A hipótese nula (h_0) afirma que não houve melhora no desempenho, e a rejeição da hipótese, ou hipótese alternativa (h_A), afirma uma melhora (ou piora) nos

resultados da medida de desempenho. A melhora (ou piora) do desempenho pode ser avaliada pela magnitude e sinal do valor da estatística t , obtida pela Equação 40.

$$t = \frac{d_m}{\frac{\sigma_m}{\sqrt{n}}} \quad (41)$$

sendo $d_m = d_{Am} - d_{Bm}$, ou seja, a média da medida de desempenho obtida nos classificadores, e σ_m o desvio padrão das medidas de desempenho, e n a quantidade de medidas realizadas na avaliação. Consideramos as medidas obtidas em cada *fold*. Estamos utilizando quatro *folds* na avaliação cruzada e, portanto, $n = 4$. Em um teste de hipótese, assumimos uma distribuição próxima da distribuição normal, considerando $\nu = n - 1$ graus de liberdade, portanto, $\nu = 3$. Para este valor, considerando o nível de confiança = 95%, bicaudal, obtemos a rejeição da hipótese nula para um valor $|t| > 3,182$, conforme função densidade da distribuição de *t-Student* (MURPHY, K. *et al.*, 2012). A rejeição da hipótese nula e um valor positivo da estatística t indica uma melhora do desempenho, e, por consequência, um valor negativo indica uma piora do desempenho, sempre em relação ao classificador de base.

A Tabela 25 mostra os valores da estatística t , comparando os resultados de AUC e MSE, considerando o classificador (isolado) e o mesmo classificador na função de classificador fraco do *AdaBoostM1*. Realçamos em negrito as comparações que representaram a rejeição da hipótese nula e, portanto, melhoria do resultado sobre análise. Consideramos também os subconjuntos de teste da validação cruzada do CERAD e CDA. É possível notar que os resultados do classificador *Decision stump* melhoram com a aplicação do *AdaBoost*.

Tabela 25. Comparação de resultados de AUC e MSE, considerando o classificador e o mesmo classificador em conjunto com AdaBoostM1.

	CERAD				CDA					
	Demência		DA		Demência		DA		TCL	
	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE
<i>Rede Bayesiana</i>	-2,97	1,07	-1,61	3,17	-0,91	0,74	-0,75	0,36	-2,64	0,63
<i>1.Naive Bayes</i>	-4,56	-0,78	-2,06	-1,18	-1,25	-1,59	-0,86	-1,70	-1,30	-0,98
<i>2.Logistic</i>	-8,75	-3,63	-3,58	-1,57	-1,79	-2,21	-0,79	-0,39	1,00	1,84
<i>3.Multilayer perceptron</i>	1,11	1,53	-2,15	-2,33	0,97	1,09	-0,95	-0,70	0	0
<i>4.Decision table</i>	2,50	-0,99	2,73	-1,10	-1,36	-2,25	1,36	1,13	0,40	-0,65
<i>5.Decision stump</i>	6,55	5,36	4,76	4,33	2,79	1,85	3,36	0,98	2,82	-8,42
<i>6.J48</i>	5,68	-2,04	1,47	-1,45	-0,55	-4,70	2,49	0,90	2,09	1,35

Avaliando os resultados da Tabela 25, é possível notar que a aplicação do AdaBoostM1 no classificador *Decision stump* (árvore de decisão de um nível), melhorou em geral seus resultados (a exceção da medida de MSE da rede para diagnóstico de TCL). Avaliando em especial os resultados do AdaBoostM1 juntamente com a rede Bayesiana, podemos concluir que não houve melhoras significativas nos resultados.

Foi realizado também um estudo de sensibilidade de alguns parâmetros utilizados no algoritmo *AdaBoostM1*, avaliando os resultados obtidos nos classificadores selecionados para comparação. Para este estudo, adotamos também o teste de hipótese nula com os mesmos parâmetros e critérios descritos anteriormente. Foi avaliada a sensibilidade na modificação dos parâmetros descritos nos itens abaixo:

- (1) Modificação da distribuição dos pesos aplicados diretamente no classificador fraco pela distribuição por reamostragem do subconjunto de treinamento (SEIFFERT *et al.*, 2008). A cada iteração, o algoritmo AdaBoostM1 atualiza a distribuição de pesos D_t e repassa diretamente ao algoritmo de aprendizagem fraca. Contudo, nem todos os algoritmos de aprendizagem possuem capacidade para lidar com informações do peso (FREUND e SCHAPIRE, 1996). Uma outra alternativa é utilizar a reamostragem. Na reamostragem, o subconjunto de treinamento é adaptado para refletir os pesos, replicando e substituindo as instâncias de forma proporcional ao respectivo peso. Ou seja, a probabilidade de encontrar uma instância na reamostragem no subconjunto de treinamento é proporcional ao peso atribuído a esta instância.
- (2) Modificação do número máximo de iterações do algoritmo AdaBoost, de 10 para 20 iterações.
- (3) Modificação da ponderação da distribuição de pesos D_t de 100% para 10%.

As Tabelas 26, 27 e 28 apresentam os valores da estatística t , realçando em negrito os testes que rejeitaram a hipótese nula e obtiveram a melhoria do resultado sobre análise. Avaliando os resultados, a melhoria pela distribuição dos pesos por reamostragem foi esparsa, não apresentando resultados consistentes ou significativos (Tabela 26). A modificação do número de iterações melhoraram os resultados de AUC e MSE da rede para diagnóstico de Demência do classificador *Logistic* (Tabela 27). Em geral, a modificação do valor ponderação trouxe melhorias também nos resultados do classificador *Decision stump* (Tabela 28).

Tabela 26. Comparação de resultados do AdaboostM1, avaliando a utilização direta da distribuição de pesos D_i pelo classificador fraco, e a distribuição de pesos por reamostragem do conjunto de treinamento.

	CERAD				CDA					
	Demência		DA		Demência		DA		TCL	
	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE
1.Naive Bayes	1,29	-0,67	-0,44	-0,42	1,51	-5,60	1,89	0,65	-0,31	-0,93
2.Logistic	0,18	2,29	2,01	-0,57	-0,50	-3,20	1,76	1,34	1,92	0,89
3.Multilayer perceptron	-0,05	-0,44	4,03	2,69	1,04	-1,74	1,39	-0,29	-3,94	0,06
4.Decision table	-4,51	-4,09	-3,16	-6,57	-0,29	1,98	-0,44	1,35	2,56	1,43
5.Decision stump	0,96	-1,37	0,34	0,56	0,99	1,18	1,98	2,19	1,16	-0,07
6.J48	-1,58	-0,99	-1,94	-2,91	-2,22	-0,66	-0,92	-0,69	-1,31	-1,24

Tabela 27. Comparação de resultados do AdaboostM1, avaliando a sensibilidade da modificação do número de iterações de 10 para 20 nos resultados do classificador.

	CERAD				CDA					
	Demência		DA		Demência		DA		TCL	
	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE
1.Naive Bayes	4,56	0,78	-0,25	0,60	-2,54	0,81	-5,19	0,46	0,23	-0,88
2.Logistic	6,75	3,63	-2,36	1,25	-2,23	1,03	-2,35	-0,04	-1,56	-1,84
3.Multilayer perceptron	-9,10	-1,56	-2,87	0,05	-3,66	-0,34	-2,76	-0,39	1,00	2,26
4.Decision table	-6,82	-2,32	-7,43	-0,70	-2,68	2,08	-6,19	-2,50	-0,47	2,86
5.Decision stump	-6,55	-5,36	-4,76	-4,33	-1,84	-1,27	-3,36	-4,02	-2,03	4,52
6.J48	-92,64	-0,09	-3,96	-1,84	-1,81	3,00	-6,92	-1,77	-2,09	-1,19

Tabela 28. Comparação de resultados do AdaboostM1, avaliando a sensibilidade da modificação do ponderação da distribuição de pesos D_i de 100% para 10% nos resultados do classificador.

	CERAD				CDA					
	Demência		DA		Demência		DA		TCL	
	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE	AUC	MSE
1.Naive Bayes	0	0	0	0	-1,00	-1,00	1,54	3,41	-1,39	-1,30
2.Logistic	0	0	0	0	0,04	1,92	0,27	0,30	1,00	1,03
3.Multilayer perceptron	-0,12	-2,10	0	0	0	0	0,58	-1,00	0	0
4.Decision table	1,88	0	2,39	0,40	-0,92	-2,84	0,48	5,11	2,31	0,47
5.Decision stump	4,89	-2,01	2,33	1,93	2,20	1,00	4,73	2,83	-0,87	-1,64
6.J48	-0,15	-9,42	-0,14	-3,25	-0,65	3,45	-2,96	-3,03	-1,00	-0,58

Considerando a melhoria geral nos resultados, optamos pela substituição do classificador *Decision stump*, pelo meta classificador *AdaBoostMI* em conjunto com *Decision stump*. A Tabela 29 mostra os resultados com os valores de *Decision stump* já substituídos pelo *AdaBoostMI*. Os resultados desta tabela mostram mais uma referência comparativa com a rede Bayesiana, o que não muda a análise realizada anteriormente. Contudo, é importante ressaltar que o objetivo desta tese não é defender o modelo Bayesiano como melhor classificador no contexto do CDSS, mas sim, defendê-lo como um modelo capaz de representar o conhecimento médico e oferecer um desempenho como classificador muito próximo a outros classificadores conhecidos.

Tabela 29. Comparação do desempenho da rede Bayesiana com outros modelos de predição, utilizando a base CDA.

Doença	Modelo	AUC ^(2↑)	F ₁ ^(↑)	TPR ^(↑)	TNR ^(↑)	MSE ^(↓)	MXE ^(↓)
Demência	Rede Bayesiana ⁽¹⁾	0,96 (0,03)	0,94 (0,05)	0,94 (0,04)	0,89 (0,10)	0,18 (0,04)	0,16 (0,13)
	1. <i>NäiveBayes</i>	0,97 (0,04)	0,93 (0,05)	0,91 (0,04)	0,93 (0,09)	0,13 (0,11)	0,07 (0,06)
	2. <i>Logistic</i>	0,97 (0,04)	0,94 (0,03)	0,93 (0,02)	0,92 (0,11)	0,11 (0,06)	0,07 (0,03)
	3. <i>Multilayer perceptron</i>	0,96 (0,05)	0,92 (0,02)	0,90 (0,05)	0,92 (0,11)	0,20 (0,15)	0,08 (0,03)
	4. <i>Decision table</i>	0,98 (0,03)	0,95 (0,04)	0,97 (0,00)	0,90 (0,01)	0,09 (0,04)	0,05 (0,03)
	5. <i>AdaBoostMI (Decision stump)</i>	0,96 (0,03)	0,93 (0,03)	0,95 (0,05)	0,86 (0,12)	0,07 (0,02)	0,12 (0,04)
	6. <i>J48</i>	0,96 (0,04)	0,95 (0,06)	0,97 (0,05)	0,89 (0,11)	0,09 (0,05)	0,06 (0,04)
Doença de Alzheimer	Rede Bayesiana	0,86 (0,04)	0,82 (0,01)	0,84 (0,04)	0,70 (0,08)	0,15 (0,02)	0,23 (0,10)
	1. <i>NäiveBayes</i>	0,85 (0,03)	0,80 (0,02)	0,81 (0,04)	0,70 (0,08)	0,21 (0,02)	0,16 (0,01)
	2. <i>Logistic</i>	0,82 (0,03)	0,81 (0,02)	0,82 (0,01)	0,71 (0,06)	0,25 (0,08)	0,17 (0,02)
	3. <i>Multilayer perceptron</i>	0,90 (0,03)	0,86 (0,03)	0,84 (0,05)	0,82 (0,06)	0,29 (0,12)	0,14 (0,03)
	4. <i>Decision table</i>	0,87 (0,06)	0,77 (0,09)	0,69 (0,12)	0,87 (0,08)	0,20 (0,04)	0,15 (0,04)
	5. <i>AdaBoostMI (Decision stump)</i>	0,86 (0,03)	0,79 (0,06)	0,72 (0,12)	0,86 (0,10)	0,14 (0,01)	0,18 (0,01)
	6. <i>J48</i>	0,76 (0,06)	0,79 (0,04)	0,83 (0,07)	0,63 (0,07)	0,24 (0,04)	0,18 (0,03)
Transtorno cognitivo leve	Rede Bayesiana	0,97 (0,05)	0,92 (0,07)	0,94 (0,07)	0,88 (0,14)	0,09 (0,08)	0,33 (0,36)
	1. <i>NäiveBayes</i>	0,97 (0,03)	0,90 (0,07)	0,85 (0,06)	0,96 (0,07)	0,15 (0,13)	0,09 (0,05)
	2. <i>Logistic</i>	0,76 (0,10)	0,69 (0,08)	0,69 (0,05)	0,63 (0,22)	0,00 (0,00)	0,34 (0,11)
	3. <i>Multilayer perceptron</i>	0,90 (0,08)	0,79 (0,04)	0,77 (0,02)	0,79 (0,14)	0,37 (0,12)	0,19 (0,03)
	4. <i>Decision table</i>	0,82 (0,16)	0,78 (0,15)	0,74 (0,16)	0,83 (0,14)	0,21 (0,09)	0,16 (0,09)
	5. <i>AdaBoostMI (Decision stump)</i>	0,96 (0,02)	0,85 (0,04)	0,83 (0,12)	0,88 (0,10)	0,11 (0,05)	0,19 (0,15)
	6. <i>J48</i>	0,79 (0,09)	0,72 (0,15)	0,61 (0,24)	0,96 (0,07)	0,21 (0,05)	0,17 (0,04)

(1) Média (DP) = média aritmética (desvio padrão). Obtido considerando 4 *folds* da validação cruzada.

(2) AUC = área da curva ROC (*Receiver Operating Characteristics*), melhor escore é 1. F₁ = medida F₁, melhor escore é 1. TPR = Sensibilidade ou taxa de verdadeiros positivos, melhor escore é 1. TNR = Especificidade ou taxa de verdadeiros falsos, melhor escore é 1. MSE = Erro médio quadrado (*Mean Square Error*), melhor escore é 0 (zero). MXE = Entropia cruzada (*Mean Cross Entropy*), melhor escore é 0 (zero).

5.4 AVALIAÇÃO DA ROBUSTEZ DA REDE BAYESIANA

A análise de robustez da rede Bayesiana, também denominada análise de sensibilidade, tem por objetivo avaliar como a rede Bayesiana reage às evidências e/ou alterações nos seus parâmetros (BERGER, 1985). Há dois tipos de análise de sensibilidade: sensibilidade às evidências e sensibilidade a mudanças de parâmetros (KORB e NICHOLSON, 2004). Utilizamos a análise de sensibilidade à evidência para verificar qual evidência é mais relevante na distribuição de probabilidade *a posteriori* da variável aleatória central (aquela que representa a probabilidade de diagnóstico). A evidência mais relevante será aquela que gerar maior redução da incerteza nesta variável aleatória. Além da avaliação de robustez, uma aplicação imediata da análise de sensibilidade é indicar para o médico qual informação ainda não observada seria mais relevante para ajudar o diagnóstico.

A entropia é a medida utilizada para avaliar a incerteza em um modelo probabilístico (conforme Equação 29). Seja X a variável aleatória central que representa a probabilidade de diagnóstico, $H(X)$ a entropia medida em X . Naturalmente, a variável aleatória X é binomial, com dois estados, *Negativo* e *Positivo*, representando a probabilidade de diagnóstico negativo e positivo da doença, respectivamente. A Tabela 30 mostra as variáveis aleatórias, listadas em ordem crescente pelo valor da entropia obtido em X , em dois grupos: diagnóstico negativo se $P(X=Negativo | E) > 0,5$, e diagnóstico positivo se $P(X=Positivo | E) > 0,5$, onde E representa a evidência sob análise.

Para o diagnóstico negativo de Demência, a evidência do CDR com resultado igual a 0 (sem alteração) aparece como mais relevante, seguida das avaliações neuropsicológicas Pfeffer, MMSE e Lawton. Uma descrição do CDR é dada na Seção 4.4. O MMSE (*Mini-Mental State Examination*), minixame do estado mental, é a escala de avaliação cognitiva mais amplamente utilizada em todo o mundo, sendo considerado por alguns pesquisadores um instrumento de rastreio (ALMEIDA, 1998; FOLSTEIN *et al.*, 1975). Através do MMSE, o paciente recebe uma pontuação de 0 a 30 (controle normal). Levantamentos mostram diferentes níveis de corte aplicados ao MMSE (BERTOLUCCI *et al.*, 2001) como critério de classificação do paciente. Na modelagem Bayesiana, obtivemos um MMSE com três e dois intervalos numéricos, conforme mostra a Tabela 11.

Já para o diagnóstico positivo de Demência, O CDT aparece no conjunto de evidências mais relevantes. Trata-se de um instrumento de avaliação cognitiva de fácil e de rápida aplicação, especialmente em idosos (HAM DAN e BUENO, 2005). O teste de fluência verbal (VFT) aparece alterado em várias doenças mentais, tais como as demências degenerativas do

tipo Alzheimer ou frontotemporal, em enfermidades psiquiátricas como a esquizofrenia e depressão (RODRIGUES *et al.*, 2008). Obtivemos a evidência VFT com pontuação abaixo de 4, o que indica uma alteração nestas dimensões cognitivas. Em seguida, TMT, teste para avaliação a atenção, consiste em conectar com lápis, no menor tempo possível e em ordem crescente, 25 números (forma A) e números alternados com letras (forma B), avaliando a atenção seletiva, processamento perceptual e flexibilidade mental (SPREEN e STRAUSS, 1998). O CDR aparece novamente como item de relevância para diagnóstico positivo de demência, com pontuação 3, indicando demência grave.

Tabela 30. Análise de sensibilidade mostrando as evidências mais relevantes para o diagnóstico (negativo e positivo) de Demência, DA e TCL, utilizando a base CDA.

Doença (X)	Diagnóstico	Evidência		Entropia H(X)
		Variável aleatória	Estado: valor	
Demência	Negativo	CDR	1: 0	0,00
		Pfeffer	1: 0	0,00
		MMSE	3: 27-30	0,27
		Lawton	3: 0-9	0,55
	Positivo	CDT	1: 0	0,00
		VFT	1: 0-4	0,00
		TMT	1: 0-16	0,00
		CDR	5: 3 (grave)	0,18
Doença de Alzheimer	Negativo	CDR	1: 0	0,41
		Lawton	1: 0-9	0,43
		Stroop	1: 0-15	0,80
	Positivo	Pfeffer	1: 0	0,23
		TMT	1: 0-16	0,53
		IQCode	1: 0-3,55	0,65
Transtorno cognitivo leve	Negativo	Berg	1: 0-54	0,04
		Lawton	2: 15-inf	0,06
		CDR	4: 2 (moderado)	0,16
		IQCode	2: 3,33-inf	0,18
	Positivo	TMT	1: 0-36	0,00
		CDT	1: 0	0,00
		Berg	3: 56-inf	0,02
		Stroop	1: 0-17	0,09

No caso da DA, o CDR aparece no conjunto de evidências mais relevantes para o grupo do diagnóstico negativo, com pontuação zero, indicando controle normal. Uma

característica da DA em estágios avançados é o impacto nas atividades de rotina diária do paciente. A escala Lawton consiste em uma avaliação psiconeurológica utilizada para avaliar a capacidade de realização das atividades de rotina diária, também conhecida pelo acrônimo escala IADL (*Instrumental Activity Daily Living*) (LAWTON, 1997). Obtivemos a evidência do teste de Lawton com pontuação na faixa de zero a 9 para diagnóstico negativo de DA. O teste neuropsicológico Stroop é utilizado para medir funções executivas e atenção. Também obtivemos a evidência do Stroop com pontuação na faixa de zero a 15 (STROOP, 1992).

Em relação ao diagnóstico positivo de DA, obtivemos a evidência de CDT com pontuação igual a 4, o Pfeffer, com pontuação igual a zero, o TMT, com pontuação na faixa de zero a 16, e o IQCode, com pontuação na faixa de zero a 3,55. O IQCode é um questionário aplicado para detecção do declínio cognitivo com base no relato do informante, desenvolvido na Austrália, originalmente no idioma inglês, composto em sua versão original por 26 itens, com perguntas organizadas em uma escala Likert, com cinco opções: (1) muito melhor, (2) um pouco melhor, (3) não houve mudança, (4) um pouco pior, e (5) muito pior. O resultado final é obtido pela soma ponderada dos itens, dividindo-os pelo total de itens da escala (SANCHEZ e LOURENÇO, 2009).

O TCL pode ser considerado como um estágio pré-clínico de DA, justificando então, a sua importância no diagnóstico desta doença (DUBOIS *et al.*, 2007). Para o diagnóstico negativo de TCL, obtivemos a evidência do teste Berg com pontuação na faixa de zero a 54. A escala de equilíbrio funcional de Berg é utilizada para avaliar o equilíbrio e demais processos motores. Pesquisas revelam que os processos motores regulados por mecanismos corticais complexos podem estar afetados no estágio leve e estágio moderado da DA (CHRISTOFOLETTI *et al.*, 2006). Em seguida, aparecem no conjunto de evidências para diagnóstico negativo de TCL, o Lawton, com pontuação acima de 15, o CDR, com pontuação 2 (demência moderada), e o IQCode, com pontuação acima de 3,33.

Em relação ao diagnóstico positivo de TCL, aparecem o TMT, com pontuação de 0 a 36, o CDT, com pontuação zero, o Berg, com pontuação acima de 36, e o Stroop, com pontuação entre zero e 17.

Foi realizada uma consulta a especialistas médicos do CDA para conhecer a relevância dos testes psiconeurológicos para o diagnóstico de cada doença de interesse deste trabalho. Na Tabela 31, está apresentada a classificação feita por especialistas médicos em graus de relevância: “vazio” para pouca relevância, “+” para média relevância, e “++” para muita relevância.

Tabela 31. Relação de avaliações neuropsicológicas e o grau de relevância para o diagnóstico positivo das doenças de interesse classificado pelo especialista.

Instrumento de avaliação	Diagnóstico ⁽¹⁾		
	Demência	DA	TCL
MMSE	++	++	+
CDR	++	++	++
VFT	++		+
Pfeffer	++	++	++
CDT	+		+
TMT	+	+	+
Stroop	+	+	+
Lawton	++	++	++
IQCode	+	+	+
Berg	+	+	+
Depressão	+	+	+
CAMCOG	++	++	++
<i>Rey Complex Figure</i>	+	+	+
<i>Digit Symbol</i>	+	+	+
NPI	+	+	+
Cornell			
TUG			
POMA			
STS			
<i>Digit Span</i>	+	+	+
RAVL	+	+	+

(1) “vazio” = pouca relevância; “+” = média relevância; “++” = muita relevância para o diagnóstico clínico da doença de interesse.

Comparando com os resultados apresentados na Tabela 30, notamos que, no diagnóstico positivo de Demência, houve concordância do CDR e VFT como instrumentos de avaliação com muita relevância. O CDT e TMT foram classificados com média relevância pelo especialista. No diagnóstico positivo da DA, houve concordância no Pfeffer como instrumento de avaliação como muita relevância, e os demais itens listados na análise de sensibilidade (TMT e IQCode) aparecem classificados com média relevância. No diagnóstico positivo de TCL, todos os itens (TMT, CDT, Berg e Stroop) aparecem classificados como instrumentos de média relevância para o diagnóstico em questão. Esta comparação valida a modelagem Bayesiana dos critérios de diagnóstico das doenças de interesse realizada neste trabalho, pois não há discordância entre as variáveis aleatórias de maior relevância das três

redes propostas para a base do CDA e o critério utilizado na prática por especialistas para diagnosticar cada uma das doenças.

Na Tabela 31, é possível ainda notar que, o CAMCOG, instrumento de avaliação classificado com muita relevância para o diagnóstico, foi excluído pelo alto volume de dados incompletos (77% do total de instâncias), conforme análise descrita na Seção 4.5.3 (Seleção dos Atributos). Da mesma forma, os instrumentos NPI, *Digit Symbol*, *Digit Span*, *Rey Complex Figure*, e o RAVL, classificados com média relevância pelo especialista, foram excluídos da modelagem pelo mesmo motivo. Para incluir estes instrumentos de avaliação e, ao mesmo tempo, manter ou melhorar o desempenho da classificação e robustez da rede Bayesiana, é necessária uma base de casos clínicos com um maior número de instâncias, ou dados mais completos, buscando reduzir os efeitos da polarização (*biased*) da base de casos clínicos e, por consequência, melhorar sua representatividade no domínio destas doenças.

Neste capítulo, avaliamos os resultados de redes Bayesianas propostas, envolvendo medidas de avaliação de desempenho e robustez. O próximo capítulo irá descrever o sistema de apoio à decisão proposto, envolvendo a modelagem de informações aplicadas à área de saúde, e arquitetura baseada em serviços web.

CAPÍTULO 6 - CDSS PARA DIAGNÓSTICO DE DEMÊNCIA, DA E TCL

Este capítulo descreve a especificação do sistema de apoio à decisão clínica (ou CDSS) proposto nesta tese para o diagnóstico de Demência, Doença de Alzheimer (DA) e Transtorno Cognitivo Leve (TCL). Além da especificação, o capítulo descreve também a implementação de um protótipo do sistema. Está organizado nas seguintes seções: a Seção 6.1 mostra uma visão geral do sistema e seus principais componentes. A Seção 6.2 descreve o modelo de representação computacional de redes Bayesianas. A Seção 6.3 mostra a especificação das evidências utilizadas pelo CDSS, aplicando os princípios da modelagem multinível da informação. Como modelo de informações médicas, adotamos o modelo de referência MLHIM (*Multilevel Health Information Modeling*) (CAVALINI e COOK, 2011), descrito também nesta seção. A Seção 6.4 descreve as tecnologias utilizadas na implementação de um protótipo do sistema, descrevendo os módulos de serviço hospedados no servidor CDSS. Os protótipos das aplicações cliente são apresentados na Seção 6.5.

6.1 ARQUITETURA DO SISTEMA

A Figura 27 mostra os principais componentes do sistema de apoio à decisão proposto nesta tese, representado no padrão UML (*Unified Modeling Language*). O modelo de decisão, considerando os critérios de diagnóstico da doença de interesse, é representado por um diagrama de influência. O diagrama de influência contém a rede Bayesiana, composta por nodos ou variáveis aleatórias, estados discretos, relacionamentos causais e a distribuição de probabilidades. A função utilidade e a caixa de decisão completam o diagrama de influência. No CDSS proposto, foram especificados cinco diagramas de influência, conforme a doença e a base de casos clínicos, utilizada para o aprendizado da distribuição de probabilidades: (1) Demência e (2) DA utilizando a base do CERAD, (3) Demência, (4) DA e (5) TCL, utilizando a base do CDA. Novos diagramas de influência podem ser incluídos, indicando a possibilidade de extensão do sistema. Na Seção 3.2, descrevemos alguns mecanismos de inferência aplicados em modelos baseados em rede Bayesiana e diagrama de influência, cujo objetivo é estimar a distribuição de probabilidades *a posteriori*. Pela diversidade dos algoritmos de inferência, representamos o mecanismo de inferência como um módulo separado do diagrama de influência. O módulo de aprendizagem tem a função de preparar os dados presentes na base de casos clínicos, selecionar e discretizar os atributos, e executar o algoritmo de aprendizagem.

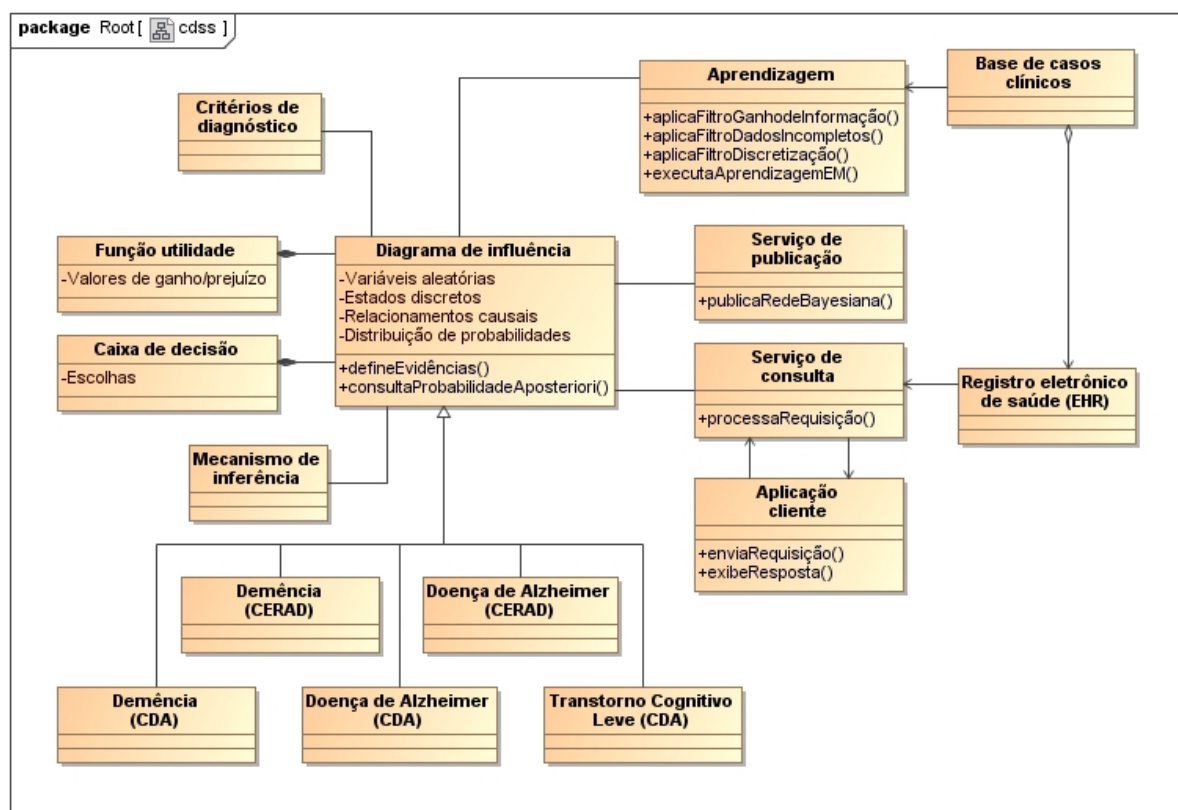


Figura 27. Componentes do sistema de suporte à decisão.

A base de casos clínicos é composta por registros eletrônicos de saúde, que seguem padrões interoperáveis para facilitar a integração com outros sistemas médicos. A interoperabilidade é definida como a capacidade de sistemas e organizações trocarem informações (NAMLI *et al.*, 2009). É comum um indivíduo buscar diferentes instituições de saúde ao longo do curso do atendimento ou tratamento da doença (CAVALINI e COOK, 2012). A capacidade de enviar e receber dados de outros sistemas de informação, e interpretá-los corretamente, é uma característica importante dos sistemas de saúde. Na Seção 6.3 apresentaremos os conceitos clínicos utilizados pela rede Bayesiana, aplicando uma modelagem multinível das informações de saúde. A modelagem multinível facilita e viabiliza a interoperação de dados de saúde (CAVALINI e COOK, 2011).

O sistema segue uma arquitetura distribuída baseada no modelo cliente-servidor, facilitando o acesso remoto e ubíquo. A interface de comunicação é composta por dois componentes: (1) o serviço de publicação, responsável pelo recebimento, armazenamento e publicação do modelo de decisão no servidor do CDSS, e (2) o serviço de consulta, responsável por processar as instâncias clínicas, executar a entrada das evidências na rede

Bayesiana, e a encaminhar ao requisitante a mensagem de resposta, contendo o resultado do mecanismo de inferência e outros processamentos no servidor do CDSS. O serviço de consulta também utiliza padrões interoperáveis para registros eletrônicos de saúde.

A Figura 28 apresenta um diagrama de sequência do CDSS. A estrutura do modelo de decisão (rede Bayesiana) é definida por especialistas do domínio do conhecimento, baseada em pesquisas e publicações científicas mais recentes dos critérios de diagnóstico de uma determinada doença (etapa 1). Em seguida, o analista de sistemas ou profissional da área da tecnologia da informação, com ajuda de especialistas, deve relacionar as variáveis aleatórias mapeadas na rede Bayesiana com os atributos da base de casos clínicos (etapa 2). Conforme mencionado anteriormente (Seção 4.1), a base de casos clínicos é utilizada no aprendizado dos parâmetros da rede Bayesiana, representados pela distribuição de probabilidades. Para definição da estrutura e aprendizagem de parâmetros, pode ser utilizada alguma ferramenta de autoria e inferência Bayesiana.

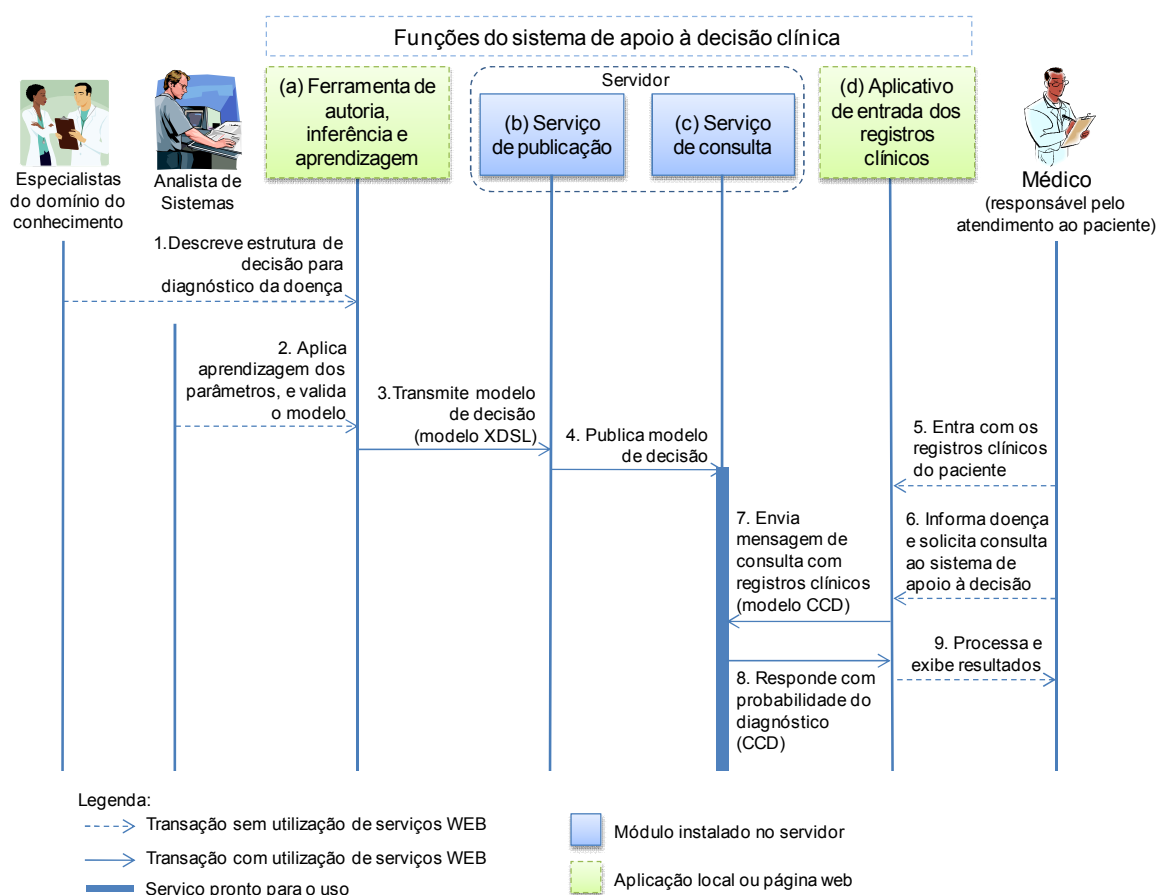


Figura 28. Diagrama de sequência mostrando a implantação do Sistema de Apoio à Decisão em uma arquitetura orientada a serviços.

Após a validação do modelo de decisão, este deve ser publicado no servidor do CDSS (etapas 3 e 4). O modelo de decisão completo (estrutura e parâmetros da rede Bayesiana, função utilidade e caixa de decisão) deve ser preferencialmente descrito em formato interoperável, como, por exemplo, em XML (*Extensible Markup Language*) (BRAY *et al.*, 1997), facilitando sua transmissão e utilização em diferentes sistemas. A Seção 6.2 comenta alguns formatos para representação de redes Bayesianas. Com o modelo de decisão publicado no servidor, este está pronto para responder requisições de consulta.

O módulo cliente disponibiliza uma interface onde o médico registra as observações clínicas. Os registros clínicos e a informação sobre a doença de interesse são transmitidos através do módulo cliente (etapas 5, 6 e 7) para o serviço de consulta. O servidor CDSS deverá combinar os registros clínicos informados pelo cliente com as variáveis aleatórias da rede Bayesiana. Esta correspondência de dados é possível devido à utilização de um modelo de dados multinível (CAVALINI e COOK, 2011) para representação de registros eletrônicos de saúde. A representação das evidências consideradas no modelo de decisão deve utilizar as mesmas informações de metadados a que o arquivo de registros clínicos faz referência. Quando é identificada uma correspondência, uma evidência é definida na rede Bayesiana, sendo propagada pela rede através do mecanismo de inferência.

O servidor retorna a mensagem ao requisitante com a probabilidade de diagnóstico positivo da doença de interesse (etapa 8). A resposta do serviço também segue os conceitos de modelagem multinível, considerando as mesmas informações de metadados dos registros clínicos. O servidor CDSS irá também retornar quais evidências foram mais relevantes para o diagnóstico em questão, como forma de o médico avaliar a robustez da resposta, e retornar também quais informações ainda não observadas seriam mais importantes para a identificação do diagnóstico. Esta mensagem de retorno é processada e exibida ao responsável pelo atendimento do paciente (etapa 9).

A próxima seção apresenta o modelo computacional utilizado nesta tese para representar o modelo de decisão do CDSS baseado em redes Bayesianas.

6.2 MODELO DE REPRESENTAÇÃO DE REDES BAYESIANAS

Um modelo de representação do conhecimento é definido por um conjunto de convenções sintáticas e semânticas que torna possível descrever objetos e propriedades, que sejam, ao mesmo tempo, legíveis por pessoas e processáveis por aplicações computacionais ou sistemas inteligentes (BRACHMAN e LEVESQUE, 2004). O conhecimento clínico

compreende a aplicação do método científico, em que evidências são avaliadas, publicadas e reguladas através da pesquisa científica, objetivando a formulação de um diagnóstico. O conhecimento clínico também pode estar relacionado com a experiência do médico ou especialista do domínio da doença (SICILIA *et al.*, 2009). Estas duas facetas encontradas na prática clínica demandam um modelo de representação do conhecimento de fácil interpretação e rápida atualização pelo especialista que domina o conhecimento, com o mínimo de envolvimento do analista de sistemas, ou profissional da área de engenharia do conhecimento.

Um modelo de representação de redes Bayesianas tem por objetivo formalizar um conjunto de regras para documentação do diagrama causal, permitindo seu armazenamento em meios eletrônicos, compartilhamento, transmissão e futuras revisões ou atualizações. No Capítulo 4, modelamos três redes Bayesianas para o diagnóstico de algumas doenças de transtornos mentais. Nesta seção, iremos mostrar alguns modelos de documentação da rede Bayesiana. Vimos que esta prática é fundamental, pois irá possibilitar revisões e atualizações pelos especialistas do domínio do conhecimento, em um ambiente altamente dinâmico e em constante evolução, como é o da prática clínica.

A Tabela 32 relaciona diferentes formatos utilizados no armazenamento de redes Bayesianas e/ou diagramas de influência. Estes formatos, na sua maioria, foram desenvolvidos para uma única ferramenta ou software de autoria e inferência Bayesiana.

Alguns formatos foram desenvolvidos exclusivamente para servir de intercâmbio entre ferramentas de autoria e inferência Bayesiana, produzidas por diferentes grupos de pesquisa e desenvolvimento. É o caso do BNIF (*Bayesian Networks Interchange Format*), formato de representação de redes Bayesianas adotado por diversas Instituições de pesquisa e Organizações comerciais, tais como, Ergo (Noetic Inc.), MSBN (Microsoft Corporation), Hugin BN (Hugin A.G.), Netica (Norsys Software Corp.), DExpress (Knowledge Industries, Inc.) e GeNIe/SMILE (Universidade de Pittsburgh). Contudo, não houve consenso entre os grupos de desenvolvimento a respeito de determinados elementos e características presentes em modelos Bayesianos, deixando o BNIF restrito a redes Bayesianas de variáveis aleatórias discretas, e um conjunto bastante limitado de características (DODIER, 1998). O formato BNIF foi um dos primeiros utilizados para documentação de redes Bayesianas baseado em XML. O XDSL é muito semelhante ao formato BNIF, com a diferença de abranger também a possibilidade de representação de diagramas de influência e estruturas locais.

Tabela 32. Levantamento de formatos utilizados para representação de redes Bayesianas e/ou diagramas de influência.

Formato	Sistema e Desenvolvedor	Formato codificado	Ano base ¹	Referência
DNET	Netica, Norsys Software Corp.	Texto estruturado	1995	Especificação disponível na web ⁴
Elvira	Consórcio entre Universidades da Espanha	Texto estruturado	1997	Especificação disponível na web ⁵
BNIF	Universidade de São Paulo	XMI ²	1998	Especificação disponível na web ⁶
XDSL	SMILE e GeNIe, Universidade de Pittsburgh	XMI	1999	Druzdzet (1999). Especificação disponível na web ⁷
XBN	MSBNx	XMI	2001	Kadie <i>et al.</i> (2001). Especificação disponível na web ⁸
BayesOWL	Universidade de Maryland	OWL ³	2006	Ding <i>et al.</i> (2006). Especificação da ontologia disponível na web ⁹
PR-OWL2	UNBBayes, Universidade de Brasília	OWL	2010	Carvalho <i>et al.</i> (2010). Especificação da ontologia disponível na web ¹⁰
ProbModelXML	UNED	XMI	2011	Arias <i>et al.</i> (2011). Especificação disponível na web ¹¹

1 Ano que o formato foi inicialmente publicado

2 XMI (*XML Metadata Interchange*), modelo intercambiável de dados, baseado no formato XML (*Extensible Markup Language*)

3 OWL (*Ontology Web Language*), formato de representação de ontologias

4 http://www.norsys.com/downloads/DNET_File_Format.txt

5 <http://leo.ugr.es/elvira/devel/Formato/formato.html>

6 <http://www.poli.usp.br/p/fabio.cozman/Research/InterchangeFormat>

7 genie.sis.pitt.edu/SMILEHelp/Appendices/XDSL_File_Format_-_XML_Schema_Definitions.htm

8 <http://research.microsoft.com/en-us/um/redmond/groups/adapt/msbnx/>

9 http://semanticweb.org/wiki/Bayes_OWL

10 <http://www.pr-owl.org/pr-owl2.owl>

11 <http://www.cisiad.uned.es/ProbModelXML/>

O formato XBN foi proposto como padrão de documentação de redes Bayesianas pela Microsoft. As primeiras contribuições deste padrão foram definidas pela comunidade científica da UAI (*Uncertainty in Artificial Intelligence*), e suas especificações estão disponíveis na Internet (KADIE *et al.*, 2001).

O formato Bayes-OWL (DING *et al.*, 2006) utiliza a notação OWL (*Ontology Web Language*) (ANTONIOU e HARMELEN, 2009), linguagem baseada em XML, para especificação de redes Bayesianas. Esse formato fornece um conjunto de regras e procedimentos, utilizados para o mapeamento direto de componentes OWL (classes, propriedades e instâncias), para elementos presentes em uma rede Bayesiana.

O formato PR-OWL, desenvolvido pela Universidade George Mason (Virgínia, Estados Unidos) e Universidade de Brasília, representa em uma extensão do OWL, e tem por objetivo a documentação de redes Bayesianas através de um formalismo híbrido, utilizando

lógica de primeira ordem e conhecimento probabilístico (COSTA e LASKEY, 2006). A base foi o modelo representacional denominado redes Bayesianas multientidades, ou MEBN (*Multi-Entity Bayesian Networks*) (LASKEY, 2008). No MEBN, a rede Bayesiana é instanciada em tempo de execução pela combinação de fragmentos ou *MFrag*s, conforme a verificação de variáveis booleanas (verdadeiro), organizadas em expressões em lógica de primeira ordem. Um *MFrag* representa um relacionamento probabilístico de um grupo de variáveis aleatórias, ou variáveis residentes, representando algum conceito, hipótese ou contexto. Cada *MFrag* pode ser instanciado múltiplas vezes, de forma recursiva, gerando redes Bayesianas mais complexas. As mesmas instituições propuseram revisões na versão inicial deste formato, sendo formalizadas no PR-OWL2 (CARVALHO *et al.*, 2010).

O formato ProbModelXML, desenvolvido pela UNED (Universidade Nacional de Educação à Distância de Madri), visa codificar modelos gráficos probabilísticos, com base em XML (ARIAS *et al.*, 2011). Neste formato é possível representar vários tipos de modelos gráficos probabilísticos, como, por exemplo, redes Bayesianas, redes de Markov, redes dinâmicas e redes de decisão, incluindo diagramas de influência.

A Tabela 33 mostra algumas características desejadas em um modelo de representação do conhecimento (CHEN, R., 2009). Os modelos baseados em OWL (Bayes-OWL, PR-OWL) atendem a todas as características citadas. Contudo, Bayes-OWL ainda está em desenvolvimento, e sua versão 1.0 permite apenas variáveis aleatórias binárias (dois estados: verdadeiro e falso). PR-OWL, disponível junto com uma ferramenta de autoria, inferência e aprendizagem Bayesiana desenvolvida pela Universidade de Brasília em código aberto, denominada UnBBayes (<http://unbbayes.sourceforge.net>), segue a representação multientidade de redes Bayesianas, utilizando uma estratégia de modelagem com objetivos diferentes do desenvolvido nesta tese (ver Capítulo 4). Os modelos baseados em XML atendem com plenitude a legibilidade, e parcialmente às demais características.

Neste trabalho, adotamos um modelo baseado em XML como forma de representação, em particular, o XDSL, padrão XML utilizado pela ferramenta de autoria GeNIe/SMILE 2.0, desenvolvida e disponibilizada para o domínio público pelo Laboratório de Sistemas de Decisão da Universidade de Pittsburg (<http://genie.sis.pitt.edu>). Este aplicativo pode ser instalado nos principais sistemas operacionais (ex., Windows, Linux e MacOS), possui uma interface gráfica altamente intuitiva, disponibiliza diferentes métodos de inferência (exata e aproximada) e aprendizagem Bayesiana, e ainda recursos que facilitam a correspondência das variáveis aleatórias com atributos da base de treinamento. A ferramenta GeNIe/SMILE possui ainda uma especificação de interface de aplicativo ou API (*Application Programming*

Interface), desenvolvida pela mesma instituição e disponível em várias plataformas de desenvolvimento, o que possibilita sua integração com outras soluções ou arquiteturas de desenvolvimento.

Tabela 33. Características desejadas em um modelo de representação do conhecimento

Característica	Descrição
Legibilidade	Ser inteligível, tanto para pessoas, como sistemas computacionais. As pessoas incluem os especialistas do domínio do conhecimento que podem não ser analistas de sistemas ou profissionais da área da computação.
Flexibilidade	Ser flexível, visando possibilitar o acréscimo de novos conhecimentos, ou revisão dos conceitos mapeados, sem prejuízo ao legado. Esta característica inclui também a possibilidade de reúso, ou aplicação do modelo para outros propósitos.
Dimensionamento	Possibilitar sua extensão e aumento da complexidade, ou redução da granularidade da informação ou artefatos do conhecimento, sem prejuízos a sua legibilidade e clareza.
Clareza	Não deve permitir que seja gerada ambiguidade nos artefatos do conhecimento mapeados.
Expressividade	Possibilitar o mapeamento do conhecimento de forma completa e abrangente, permitindo a expressão de todas as particularidades ou facetas relevantes para o modelo em questão.

Na próxima seção, iremos apresentar o modelo de representação de conceitos clínicos baseado em uma modelagem multinível. O uso de um modelo multinível facilita a interoperação do CDSS proposto com outros sistemas na área de saúde. O conhecimento do modelo de representação de registros de saúde é importante na indicação das evidências do modelo de decisão.

6.3 MODELO DE REPRESENTAÇÃO DE REGISTROS DA SAÚDE

O modelo de representação de registros da saúde é um importante componente da construção de sistemas de informações aplicados à área de saúde, também conhecidos pela sigla HIS (*Health Information System*). Antes de tratar o modelo de representação, convém enumerarmos alguns desafios da implantação de HIS na área de saúde: (1) transmissão e comunicação de dados de saúde de pacientes entre múltiplas instituições (ex., postos de saúde, laboratórios clínicos, clínicas de exames de imagens, hospitais, unidades de atendimento emergencial), (2) redução dos custos dos testes de integração e aceitação de sistemas, (3)

redução de erros médicos, melhoria da qualidade de atendimento e segurança do paciente, (4) redução da documentação física (papel) e melhoria da sua legibilidade (CANTRILL e STEPHEN, 2010). Os fatores que mais impactam a adoção de HIS podem ser analisados sobre três aspectos: (1) organizacional, (2) ambiental, (3) tecnológico. No aspecto tecnológico, alguns fatores de impacto são: (1) alto custo de desenvolvimento, (2) falta de padronizações abertas ou de domínio público, (3) falta de orientação tecnológica, (4) dificuldade de manutenção (VEST, 2010). Os desafios são decorrentes do fato da área de saúde representar um sistema complexo e dinâmico (HYMAN, 2010).

Analisando o processo de atendimento à saúde, o paciente normalmente é levado a percorrer múltiplas unidades de atendimento. O acesso do paciente ao serviço de saúde geralmente acontece através de um posto de saúde ou unidade emergencial. Em casos mais complexos, este paciente é conduzido a outras unidades mais especializadas. Sobre o ponto de vista de um HIS, espera-se que os dados de saúde sejam transmitidos entre as instituições, sejam elas Governamentais, ou não, mantendo seus registros longitudinais, e todas as informações detalhadas que possam servir para o planejamento, diagnóstico ou tratamento do paciente. Estas informações devem ainda estar disponíveis a qualquer momento durante seu atendimento. Adicionalmente, é comum as instituições de saúde adotarem procedimentos específicos. Em consequência, os registros de saúde do paciente têm significado limitado ao HIS onde são aplicados (ZUSMAN, 2010). Isto representa um problema para o intercâmbio de dados entre outras instituições (HYMAN, 2010). Para estes cenários, o modelo de representação de dados relacional, comumente encontrado em sistemas de informação, provou não ser adequado, devido à necessidade da interoperação de diferentes modelos de dados (METAXIOTIS *et al.*, 2004).

Com objetivo de assegurar a interoperabilidade semântica e a segurança da informação entre diferentes HISes, foram propostas padronizações. Estas padronizações podem ser classificadas em duas categorias: padronização (1) de terminologia e (2) de especificação. As terminologias mais abrangentes aplicadas na área de saúde são: (1) SNOMED-CT (*Systematized Nomenclature of Medicine – Clinical Terms*) com 311.000 conceitos organizados de forma hierárquica, mantido pela IHTSDO (*International Health Terminology Standards Development Organization*), (2) ICD (*International Classification of Diseases*), formado por termos de diagnóstico clínico, utilizado em pesquisas epidemiológicas, gerenciamento de saúde e propósitos clínicos, mantido pela WHO (*World Health Organization*).

As padronizações de especificações definem um conjunto de modelos de dados, estrutura de relacionamentos e procedimentos de modelagem para representação de conceitos da área de saúde, utilizando uma abordagem de engenharia de sistemas orientada a modelos. Algumas padronizações aplicadas à área de saúde são: (1) HL7 (DOLIN *et al.*, 2001), (2) ISO 13606 (SANTOS *et al.*, 2010), (3) openEHR (KALRA, D. *et al.*, 2005) e (4) MLHIM (CAVALINI e COOK, 2011).

O padrão HL7 é mantido por uma organização sem fins lucrativos e atualmente está na versão 3. No HL7, os dados são modelados com base em um modelo de referência, denominado RIM (*Reference Information Model*). Os registros de saúde são então apresentados como uma instância deste modelo, o que deixa este padrão muito semelhante ao modelo de base relacional (CHEN, R. S. *et al.*, 2000; DOLIN *et al.*, 2001).

Os padrões ISO 13606 e openEHR e o formato MLHIM seguem uma modelagem multinível (ALVERNAZ *et al.*, 2012). O princípio da modelagem multinível é baseado nos princípios da análise orientada a objetos, sendo composto, no mínimo, por dois níveis. No primeiro nível está o modelo de referência (RM – *Reference Model*), o qual descreve classes e tipos de dados genéricos. Os tipos de dados aplicados à área de saúde estão especificados pela ISO 21090. O segundo nível é o modelo de domínio (DM – *Domain Model*). Todo o conceito de um determinado domínio de HIS é modelado pela definição de restrições às classes RM. Na modelagem multinível, não deve haver necessidade de modificação estrutural das classes RM, sendo estáveis ao longo do tempo. Já as restrições definidas no DM fornecem uma interpretação semântica para os objetos armazenados. Os artefatos DM são criados e modificados por especialistas do domínio da área de saúde, sem a necessidade de interação com analistas de sistema, ou cientistas da computação. A ideia da modelagem multinível é que qualquer modificação de regra de negócio do HIS seja refletido na segunda camada (DM), e não na primeira camada (RM) (BEALE, T., 2003). A Figura 29 ilustra os componentes envolvidos na modelagem multinível.

Apesar das especificações abrangentes e detalhadas, na prática, há poucas implementações da modelagem multinível em HISes (KALRA, D. *et al.*, 2005). As principais razões são: (1) complexidade do modelo (ATALAG *et al.*, 2011), (2) limitações dos dispositivos eletrônicos (Lang, 2011), (3) diversidade da área de saúde. No modelo de domínio, os conceitos são descritos utilizando um modelo de restrição às classes descritas no modelo de referência. Nos padrões ISO 13606 e openEHR, as descrições dos conceitos são denominadas arquétipos, e são expressos em uma linguagem computacional baseada em texto estruturado denominada ADL (*Archetype Definition Language*) (MARTÍNEZ-COSTA *et al.*,

2010). Esse modelo exige a correspondência unívoca entre o conceito clínico e o arquétipo (KALRA, D. *et al.*, 2005). Desta forma, antes da sua utilização em HISes, a modelagem exige o consenso de toda a comunidade quanto ao conceito representado pelo arquétipo, o que onera e dificulta sua implementação e implantação na rotina clínica em larga escala. Outro ponto a ser relevado é a própria linguagem de representação dos arquétipos. A linguagem ADL não é de domínio da maioria da comunidade científica, nem reconhecida como padrão pelo W3C (*World Wide Web Consortium*), o que aumenta o tempo de aprendizagem.

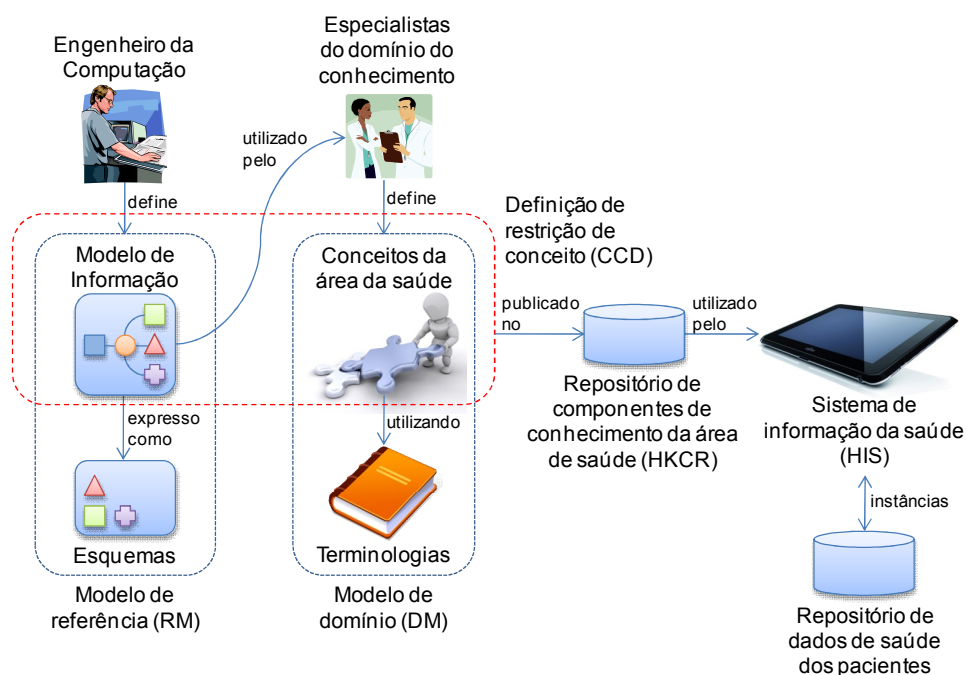


Figura 29. Componentes da modelagem multinível para sistemas de informação da saúde (HIS).

Neste trabalho, adotamos o formato MLHIM para modelagem dos conceitos clínicos utilizados pelo CDSS. No MLHIM, o modelo de referência e modelo de domínio possuem classes análogas às definidas pela ISO 13606 e openEHR, contudo, suas especificações são descritas em XSD (*XML Schema Definition*) e XML. Em alguns HISes, o XML já é utilizado para a estrutura de informações (SOKOLOWSKI e DUDECK, 1999). A adoção destes padrões tem por objetivo assegurar a sustentabilidade do modelo e melhorar a relação de custos da sua implementação em HISes. Assim, um arquivo XSD é utilizado para descrever restrições ao modelo de referência. Este XSD é denominado Definição de Restrição de Conceito (CCD - *Concept Constraint Definition*) e define todas as combinações e restrições das classes do modelo de referência do MLHIM para representar um conceito da área de

saúde, assim, de forma análoga, aos arquétipos especificados no padrão ISO 13606 e openEHR (CAVALINI e COOK, 2011).

Enquanto os padrões ISO 13606 e openEHR adotam um modelo de governança centralizado (de cima para baixo), o MLHIM adota um modelo descentralizado (de baixo para cima) (CAVALINI e COOK, 2011). Para cada CCD, é atribuída uma identificação universal única, ou UUID (*Unique Universal Identifier*), permitindo que duas ou mais implementações de CCDs representem conceitos similares. A interoperabilidade semântica é assegurada pela validação do CCD através da sua distribuição baseada na web, denominada repositório de componentes de conhecimento na área de saúde, ou HKCR (*Healthcare Knowledge Component Repository*). Apesar da possibilidade de existir CCDs duplicados, com a maturidade do modelo, espera-se que os melhores CCDs se sobressaiam e sejam naturalmente selecionados para o conceito clínico correspondente (SHAH e KESAN, 2008).

O formato MLHIM está em conformidade com as seguintes padronizações da ISO: ISO/TS 18308:2004 (Informática em Saúde – requisitos para uma arquitetura de registros eletrônicos de saúde), ISO/TR 20514:2005 (Registro Eletrônico de Saúde – definição, escopo e contexto), ISO 13606-1:2008 (Informática em Saúde – comunicação de registros eletrônicos de saúde, modelo de referência), ISO 13606-2:2008 (Informática em Saúde – comunicação de registros eletrônicos de saúde, especificação do intercâmbio de arquétipos), ISO 13606-3:2009 (Informática em Saúde – comunicação de registros eletrônicos de saúde, lista de termos e arquétipos de referência), ISO 13606-4:2009 (Informática em Saúde – comunicação de registros eletrônicos de saúde, segurança), ISO 13606-5:2010 (Informática em Saúde – comunicação de registros eletrônicos de saúde, especificação da interface) e ISO/FDIS 21090:2011 (Informática em Saúde – harmonização de tipos de dados para intercâmbio de informações).

Para a implementação de um protótipo do CDSS proposto, selecionamos quatorze atributos da base de treinamento do CDA, listados na Tabela 10, descrita no Capítulo 4. Desenvolvemos um CCD com o conjunto de especificações dos atributos. Os atributos incluem avaliações neuropsicológicas, dados demográficos do paciente, e a observação de sintomas ou sinais, relatados pelo paciente, acompanhante, ou identificados pelo médico no momento do atendimento. Foi especificado apenas o modelo clínico para abrigar o resultado ou pontuação final de cada avaliação neuropsicológica. A Figura 30 mostra a seção de metadados do CCD. No elemento *subject*, estão especificados os conceitos clínicos citados no CCD. Os conceitos aparecem separados por ponto-e-vírgula, e seus nomes devem ser

fornecidos por algum dicionário ou vocabulário controlado. Após a especificação do CCD, o próximo passo é sua publicação em algum repositório de componentes de conhecimento da área da saúde, ou HKCR. Para que o CCD possa ser identificado e resgatado após sua publicação, é fundamental a utilização de um vocabulário controlado na especificação dos conceitos citados no CCD, e sua devida descrição na seção de metadados.

```

...
01 <xs:annotation>
02   <xs:documentation source="http://www.mlhim.org" xml:lang="en-US">
03     This is a Concept Constraint Definition schema file (CCD).
04     <rdf:RDF>
05       <rdf:Description rdf:about=
06         "http://www.hkcr.net/ccd/ccd-ce2a1c62-37a3-4ff6-a7f3-6f1fef408d22.xsd">
07         <dc:title>Assessment Scales</dc:title>
08         <dc:creator>Flavio L. Seixas</dc:creator>
09         <dc:contributor/>
10         <dc:subject>Intellectual Disability; Memory Disorders; MeSH Heading
11           Neuropsychological Tests;Assessment scales;Mini-mental state
12           score;Clinical dementia rating scale; Verbal fluency test score;
13           Pfeffer questionnaire score;Clock drawing test scale;
14           Trial making test;Education;Age;Stroop color word test;Lawton
15           scale;IQCode score;Berg balance scale;Depression;Gender
16         </dc:subject>
17         <dc:source>See please the Reference Section in comments at the end
18           of this file.</dc:source>
19         <dc:rights/>
20         <dc:relation>None</dc:relation>
21         <dc:coverage>Universal</dc:coverage>
22         <dc:type>MLHIM Concept Constraint Definition (CCD)</dc:type>
23         <dc:identifier>
24           ccd-ce2a1c62-37a3-4ff6-a7f3-6f1fef408d22</dc:identifier>
25         <dc:description>This is a CCD to record results from assessment
26           scales.</dc:description>
27         <dc:publisher>MLHIM</dc:publisher>
28         <dc:date>2012-09-01</dc:date>
29         <dc:format>text/xml</dc:format>
30         <dc:language>en-US</dc:language>
31       </rdf:Description>
32     </rdf:RDF>
33   </xs:documentation>
34 </xs:annotation>
...

```

Figura 30. Seção de metadados do CCD utilizado para especificação dos instrumentos de avaliação psiconeurológica. Código em XSD e RDF.

Os conceitos associados ao CCD aparecem listados na Tabela 34, bem como sua referência terminológica, quando identificada. Os termos foram extraídos do SNOMED-CT, versão de Janeiro/2012, em pesquisa pelo navegador (SNOMED *browser*), disponibilizado no endereço <https://uts.nlm.nih.gov/snomedctBrowser.html>, e do dicionário médico *Medical Subject Headings*, ou *MeSH Terms*, disponibilizado no endereço http://www.nlm.nih.gov/mesh/2012/mesh_browser/MBrowser.html, ambas as terminologias

administradas pela UMLS (*Unified Medical Language System*), serviço fornecido pela Biblioteca Nacional de Medicina dos Estados Unidos. Não conseguimos encontrar referências para todos os termos, por isso, alguns aparecem com Identificação igual a “Não encontrado”.

Tabela 34. Relação de conceitos associados ao CCD, identificando sua referência terminológica.

Instrumento de avaliação	Termo (Original)	Tipo	Identificação ⁽¹⁾	Fonte
-	<i>Assessment scales</i>	Escala	273249006	SNOMED-CT
MMSE	<i>Mini mental states score</i>	Escala	414750000	SNOMED-CT
CDR	<i>Clinical dementia rating scale</i>	Escala	273367002	SNOMED-CT
VFT	<i>Verbal fluency test score</i>	Escala	Não encontrado	
Pfeffer	<i>Pfeffer questionnaire score</i>	Escala	Não encontrado	
CDT	<i>Clock drawing test scale</i>	Escala	Não encontrado	
TMT	<i>Trial making test</i>	Escala	273882000	SNOMED-CT
Educação	<i>Education</i>	Entidade observável	276031006	SNOMED-CT
Idade	<i>Age</i>	Entidade observável	424144002	SNOMED-CT
Stroop	<i>Stroop color word test</i>	Escala	304776005	SNOMED-CT
Lawton	<i>Lawton scale</i>	Escala	Não encontrado	
IQCode	<i>IQCode score</i>	Escala	Não encontrado	
Berg	<i>Berg balance scale</i>	Escala	428203000	SNOMED-CT
Depressão	<i>Depression</i>	Achado clínico	35489007	SNOMED-CT
Gênero	<i>Gender</i>	Entidade observável	263495000	SNOMED-CT
Demência	<i>Dementia</i>	Achado clínico	52448006	SNOMED-CT
Doença de Alzheimer	<i>Alzheimer's disease</i>	Achado clínico	26929004	SNOMED-CT
Transtorno cognitivo leve	<i>Mild cognitive disorder</i>	Achado clínico	386805003	SNOMED-CT

(1) Identificação única numérica do conceito na base terminológica.

As Figuras 31 e 32 apresentam um trecho do código descrevendo dois elementos CCD: o primeiro (Figura 31) utilizado para os resultados da avaliação do MMSE e o segundo (Figura 32) utilizado para os resultados do CDR. Nota-se que, em ambos os elementos, foi utilizado o tipo de dado ordinal (*DvOrdinalType*), especificado no nível de modelo de referência (RM). O MMSE apresenta resultados de 0 (zero) a 30 (trinta). Nota-se que, para este exame, subdividimos seu domínio numérico em 4 subfaixas mutuamente exclusivas e sequenciais, representadas na enumeração com a sintaxe *de – para*. Isto se deve ao método de

discretização supervisionada, aplicado para melhorar o desempenho do classificador Bayesiano, descrito no Capítulo 4. No CDR, estão representados 5 níveis escalares indicando o resultado final do exame, expressos na enumeração.

Finalmente, a Figura 33 mostra um exemplo de instância em XML para o CDR. A interpretação do registro clínico será possível se o CCD estiver disponibilizado (por exemplo, via HKCR), bem como o modelo de referência (RM), conforme especifica o atributo "*xsi:schemaLocation*", definido no cabeçalho da instância XML.

```

...
01 <xs:complexType name="ct-1b5fb55f-035c-4590-a1e2-4b271e9289ba">
02   <xs:complexContent>
03     <xs:restriction base="mlhim2:ElementType">
04       <xs:sequence>
05         <xs:element maxOccurs="unbounded" minOccurs="0"
06           ref="mlhim2:links"/>
07         <xs:element maxOccurs="1" minOccurs="1" name="annotation"
08           type="xs:string" fixed="Mini mental state score"
09           rdf:resource="https://uts.nlm.nih.gov/snomedctBrowser.html?conceptId=414750000"/>
10         <xs:element minOccurs="1" maxOccurs="1"
11           ref="mlhim2:el-82277514-3bdf-4579-95f0-0ba3ea912fe4"/>
12       </xs:sequence>
13     </xs:restriction>
14   </xs:complexContent>
15 </xs:complexType>
16 <xs:element name='el-82277514-3bdf-4579-95f0-0ba3ea912fe4'
17   substitutionGroup='mlhim2:Element-dv'
18   type='mlhim2:ct-82277514-3bdf-4579-95f0-0ba3ea912fe4' />
19 <xs:complexType name="ct-82277514-3bdf-4579-95f0-0ba3ea912fe4">
20   <xs:complexContent>
21     <xs:restriction base="mlhim2:DvOrdinalType">
22       <xs:sequence>
23         <xs:element name="data-name" type="xs:string" maxOccurs="1"
24           minOccurs="1" fixed="Mini mental state score" />
25         <xs:element maxOccurs="1" minOccurs="1" name="DvOrdinal-dv">
26           <xs:simpleType>
27             <xs:restriction base="xs:decimal">
28               <xs:enumeration value="1"/>
29               <xs:enumeration value="2"/>
30               <xs:enumeration value="3"/>
31             </xs:restriction>
32           </xs:simpleType>
33         </xs:element>
34         <xs:element maxOccurs="1" minOccurs="1" name="symbol">
35           <xs:simpleType>
36             <xs:restriction base="xs:string">
37               <xs:enumeration value="0-17"/>
38               <xs:enumeration value="18-26"/>
39               <xs:enumeration value="26-30"/>
40             </xs:restriction>
41           </xs:simpleType>
42         </xs:element>
43       </xs:sequence>
44     </xs:restriction>
45   </xs:complexContent>
46 </xs:complexType>
...

```

Figura 31. Trecho do CCD utilizado para especificação da pontuação do instrumento de avaliação MMSE. Código em XSD.

```

...
01 <xs:complexType name="ct-7cb5d29a-cd6b-4508-8296-abe889d8baa3">
02   <xs:complexContent>
03     <xs:restriction base="mlhim2:ElementType">
04       <xs:sequence>
05         <xs:element maxOccurs="unbounded" minOccurs="0"
06           ref="mlhim2:links"/>
07         <xs:element maxOccurs="1" minOccurs="1" name="annotation"
08           type="xs:string" fixed="Clinical dementia rating scale"
09           rdf:resource="https://uts.nlm.nih.gov/snomedctBrowser.html?conceptId=273367002"/>
10         <xs:element minOccurs="1" maxOccurs="1"
11           ref="mlhim2:el-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0"/>
12       </xs:sequence>
13     </xs:restriction>
14   </xs:complexContent>
15 </xs:complexType>
16 <xs:element name='el-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0'
17   substitutionGroup='mlhim2:Element-dv'
18   type='mlhim2:ct-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0' />
19 <xs:complexType name="ct-1e4b3a59-2f1d-485d-b15b-0e9c9f7397c0">
20   <xs:complexContent>
21     <xs:restriction base="mlhim2:DvOrdinalType">
22       <xs:sequence>
23         <xs:element name="data-name" type="xs:string" maxOccurs="1"
24           minOccurs="1" fixed="Clinical dementia rating scale"/>
25         <xs:element maxOccurs="1" minOccurs="1" name="DvOrdinal-dv">
26           <xs:simpleType>
27             <xs:restriction base="xs:decimal">
28               <xs:enumeration value="0"/>
29               <xs:enumeration value="0.5"/>
30               <xs:enumeration value="1"/>
31               <xs:enumeration value="2"/>
32               <xs:enumeration value="3"/>
33             </xs:restriction>
34           </xs:simpleType>
35         </xs:element>
36         <xs:element maxOccurs="1" minOccurs="1" name="symbol">
37           <xs:simpleType>
38             <xs:restriction base="xs:string">
39               <xs:enumeration value="0-normal control"/>
40               <xs:enumeration value="0.5-very mild"/>
41               <xs:enumeration value="1-mild"/>
42               <xs:enumeration value="2-moderate"/>
43               <xs:enumeration value="3-severe"/>
44             </xs:restriction>
45           </xs:simpleType>
46         </xs:element>
47       </xs:sequence>
48     </xs:restriction>
49   </xs:complexContent>
50 </xs:complexType>
...

```

Figura 32. Trecho do CCD utilizado para especificação da pontuação do instrumento de avaliação CDR. Código em XSD.

```

01 <?xml version="1.0" encoding="UTF-8"?>
02 <mlhim2:ccd-ce2a1c62-37a3-4ff6-a7f3-6f1fef408d22
    xmlns:mlhim2="http://www.mlhim.org/xmlns/mlhim2/2_4_0"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:data-view="http://www.w3.org/2003/g/data-view#"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.mlhim.org/xmlns/mlhim2/2_4_0
file:/home/MLHIM/Flavio_CDS/ccd_ce2a1c62_37a3_4ff6_a7f3_6f1fef408d22.xsd">
03 <mlhim2:e1-852cd734-b124-462b-8e33-9384e53bfa49>
04   <mlhim2:language>en-US</mlhim2:language>
05   <mlhim2:encoding>utf-8</mlhim2:encoding>
06   <mlhim2:e1-e168a6b0-e797-4020-bc68-acda2a351aac>
07     <mlhim2:party-name>Self</mlhim2:party-name>
08   </mlhim2:e1-e168a6b0-e797-4020-bc68-acda2a351aac>
09   <mlhim2:e1-5d097c67-58cf-4a66-a7dd-cc8023d5b808>
10     <mlhim2:cluster-subject>Care Cluster</mlhim2:cluster-subject>
11     <mlhim2:e1-0c71fe4c-8dd2-4d0f-af05-15b5b7b9de24>
12       <mlhim2:e1-1b5fb55f-035c-4590-ale2-4b271e9289ba>
13         <mlhim2:annotation>Mini mental state score</mlhim2:annotation>
14         <mlhim2:e1-82277514-3bdf-4579-95f0-0ba3ea912fe4>
15           <mlhim2:data-name>Mini mental state score</mlhim2:data-name>
16           <mlhim2:DvOrdinal-dv>1</mlhim2:DvOrdinal-dv>
17           <mlhim2:symbol>0-17</mlhim2:symbol>
18         </mlhim2:e1-82277514-3bdf-4579-95f0-0ba3ea912fe4>
...

```

Figura 33. Exemplo de trecho de código em XML, mostrando uma instância do MMSE.

Nesta tese, modelamos cinco modelos de decisão ou redes Bayesianas. Para cada modelo de decisão, foi necessário desenvolver um CCD, descrevendo os conceitos clínicos empregados na rede Bayesianas correspondente, bem como os valores nominais, ou intervalos numéricos, das variáveis aleatórias, ou nodos da rede Bayesianas.

6.4 DESCRIÇÃO DO PROTÓTIPO DO CDSS E DOS MÓDULOS DE SERVIÇO

Para a implementação de um protótipo do CDSS proposto, foi utilizada a arquitetura orientada a serviços, ou SOA (*Service-Oriented Architecture*) (KRAFZIG *et al.*, 2005). O SOA é uma plataforma de serviços baseada na web, usualmente denominada plataforma de serviços web, ou simplesmente *Web Services* (CURBERA *et al.*, 2005). Os componentes do *Web Services* podem ser classificados em três funções: o componente (1) Servidor (ou Fornecedor), responsável pelo processamento da requisição e entrega dos dados, o (2) Cliente (ou Consumidor), responsável por efetuar o pedido, montar o arquivo XML da requisição e tratar a mensagem de resposta do servidor. Um terceiro componente, (3) Agente de Serviços é responsável pelo processo de publicação, pesquisa e descoberta de serviços web. Os serviços web são especificados utilizando uma linguagem baseada em XML, em uma estrutura denominada Linguagem de Descrição de Serviços Web, ou WSDL (*Web Services Description*

Language), seguindo a recomendação do W3C (CHINNICI *et al.*, 2007). Os serviços são identificados pelo protocolo UDDI (*Universal Description, Discovery and Integration*) (CERAMI, 2002). Entre as principais vantagens da arquitetura orientada a serviços, podemos destacar: (1) os serviços são identificados por uma URI (*Uniform Resource Identifier*), descritos e definidos utilizando XML, distribuídos utilizando HTTP (*HyperText Transfer Protocol*), todas estas tecnologias já amplamente utilizadas em aplicativos da Internet; (2) a comunicação entre sistemas de informação acontece através de mensagens codificadas em texto através da Internet, facilitando a interoperabilidade entre diferentes sistemas, independente do local ou computador onde os sistemas estão hospedados, independente do sistema operacional (ex., Windows, Unix, MacOS, *etc.*), independente da interface e tecnologia de acesso (Ethernet, Wireless, V35, *etc.*); (3) a plataforma de serviços web possui também recursos para segurança da informação, como o protocolo SSL (*Secure Socket Layer*) no HTTP, encriptação dos dados (*XML Encryption*), e outras soluções de autenticação, integridade de dados e privacidade de serviços.

O serviço de publicação (Figura 28 – item b), construído sobre a plataforma de serviços web, é responsável pela transferência da rede de decisão ao servidor onde estará hospedado o serviço de consulta do CDSS. A rede Bayesiana é representada no formato XDSL, como descrito na Seção 6.2. No servidor, o arquivo XDSL é renomeado para um nome que identifique univocamente o diagnóstico o qual será aplicado, e a base de treinamento sobre a qual foi desenvolvido. Uma mensagem é retornada ao requisitante, indicando o sucesso, ou não, da operação. Os parâmetros de entrada são: (1) a identificação da doença e base de treinamento, (2) o conteúdo do arquivo XDSL com o modelo de decisão. O parâmetro de saída é a uma mensagem indicando de sucesso ou fracasso da operação. A Figura 34 mostra a especificação dos parâmetros de entrada (linhas 8 e 9), chamados *disease* e *xdslinstance*, e saída (linha 16), chamado *publishXDSLResult*, utilizando formato WSDL. Nota-se que os parâmetros, tanto de entrada como de saída, foram modelados com tipo *string* (*type="s:string"*).

O serviço de consulta (Figura 28 – item c) é responsável pelo atendimento às requisições de consulta a rede de decisão. Os parâmetros de entrada são: (1) a identificação da doença (parâmetro *disease*), (2) a identificação do nodo central (o nodo que representa a probabilidade de diagnóstico da doença (parâmetro *target_node*), e (3) a instância XML (parâmetro *ccdInstance*), contendo os registros clínicos do paciente modelados conforme CCD, como descrito na Seção 6.3. Os parâmetros de saída são: a mensagem XML com a probabilidade de diagnóstico positivo, dadas as evidências, compostas pelos itens clínicos do

paciente (parâmetro *queryCDSSresult*), (2) a relação de itens que foram mais relevantes para o diagnóstico apresentado (parâmetro *relevantItems*), e (3) a relação de itens ainda não observados mais importantes para identificação do diagnóstico (parâmetro *nextRelevantItems*). A Figura 35 mostra a especificação dos parâmetros de entrada (linhas 4, 5 e 6) e saída (linhas 13, 14 e 15) no formato WSDL. Para codificar a probabilidade no modelo CCD, utilizamos o tipo de dado Proporção (*DvProportionType*). A Figura 36 mostra o código XSD que especifica a resposta do CDSS. A probabilidade é registrada no rótulo *numerator* (Figura 36, linha 10). Os rótulos *magnitude* (Figura 36, linha 6) e *denominator* (Figura 36, linha 11) contêm valores unitários, facilitando algum processamento numérico na aplicação cliente, caso seja necessário. A Figura 37 mostra uma instância XML da resposta do CDSS.

O aplicativo de entrada dos registros clínicos (Figura 28 – item d) fornece uma interface gráfica amigável ao médico para entrada das observações clínicas do paciente. Após a entrada das observações clínicas, o aplicativo irá gerar uma instância XML, válida pelo modelo de restrição (CCD) e referência MLHIM. O aplicativo irá também tratar a mensagem de retorno do CDSS, podendo exibir na mesma interface gráfica, os resultados para avaliação do clínico.

```

01 <?xml version="1.0" encoding="utf-8"?>
02 <wSDL:definitions xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
   xmlns:mime="http://schemas.xmlsoap.org/wsdl/mime/"
   xmlns:tns="http://localhost/"
   xmlns:s="http://www.w3.org/2001/XMLSchema"
   xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
   xmlns:tm="http://microsoft.com/wsdl/mime/textMatching/"
   xmlns:http="http://schemas.xmlsoap.org/wsdl/http/"
   xmlns:wSDL="http://schemas.xmlsoap.org/wsdl/">
03 <wSDL:types>
04   <s:schema elementFormDefault="qualified" targetNamespace="http://localhost/">
05     <s:element name="PublishXDSL">
06       <s:complexType>
07         <s:sequence>
08           <s:element minOccurs="1" maxOccurs="1" name="disease" type="s:string" />
09           <s:element minOccurs="1" maxOccurs="1" name="xSDLInstance" type="s:string"/>
10         </s:sequence>
11       </s:complexType>
12     </s:element>
13     <s:element name="PublishXDSLResponse">
14       <s:complexType>
15         <s:sequence>
16           <s:element minOccurs="1" maxOccurs="1" name="publishXDSLResult"
17             type="s:string" />
18         </s:sequence>
19       </s:complexType>
20     </s:element>
21   </s:schema>
22 </wSDL:types>
23 </wSDL:definitions>

```

Figura 34. Trecho do arquivo WSDL, com a descrição do serviço de publicação.

```

...
01   <s:element name="QueryCDSS">
02     <s:complexType>
03       <s:sequence>
04         <s:element minOccurs="1" maxOccurs="1" name="disease" type="s:string" />
05         <s:element minOccurs="1" maxOccurs="1" name="target_node" type="s:string" />
06         <s:element minOccurs="1" maxOccurs="1" name="ccdInstance" type="s:string" />
07       </s:sequence>
08     </s:complexType>
09   </s:element>
10   <s:element name="QueryCDSSResponse">
11     <s:complexType>
12       <s:sequence>
13         <s:element minOccurs="1" maxOccurs="1" name="queryCDSSResult"
14           type="s:string" />
15         <s:element minOccurs="1" maxOccurs="1" name="relevantItems"
16           type="s:string" />
17         <s:element minOccurs="1" maxOccurs="1" name="nextRelevantItems"
18           type="s:string" />
19       </s:sequence>
20     </s:complexType>
21   </s:element>
...

```

Figura 35. Trecho do arquivo WSDL, contendo a descrição do serviço de consulta.

```

...
01   <xs:complexType name="ct-82277514-3bdf-4579-95f0-0ba3ea912fe5">
02     <xs:complexContent>
03       <xs:restriction base="mlhim2:DvRatioType">
04         <xs:sequence>
05           <xs:element maxOccurs="1" minOccurs="1" name="data-name"
06             type="xs:string" fixed="Clinical diagnosis"/>
07           <xs:element maxOccurs="1" minOccurs="1" name="magnitude"
08             type="xs:decimal"/>
09           <xs:element maxOccurs="1" minOccurs="1" name="error" default="0"
10             type="xs:int"/>
11           <xs:element maxOccurs="1" minOccurs="1" name="accuracy" default="0"
12             type="xs:decimal"/>
13           <xs:element maxOccurs="1" minOccurs="1" name="ratio-type"
14             type="xs:string" fixed="proportion"/>
15           <xs:element maxOccurs="1" minOccurs="1" name="numerator"
16             type="xs:decimal"/>
17           <xs:element maxOccurs="1" minOccurs="1" name="denominator"
18             type="xs:decimal"/>
19           <xs:element maxOccurs="1" minOccurs="1" name="numerator-units"
20             type="xs:string"/>
21           <xs:element maxOccurs="1" minOccurs="1" name="denominator-units"
22             type="xs:string"/>
23         </xs:sequence>
24       </xs:restriction>
25     </xs:complexContent>
26   </xs:complexType>
...

```

Figura 36. Trecho do CCD com especificação da resposta do CDSS. Código em XSD.

As evidências da rede Bayesiana são identificadas pela combinação dos registros clínicos com as características de diagnóstico da doença, modeladas na rede Bayesiana na forma de variáveis aleatórias. Utilizamos o atributo *data_name* do elemento CCD e o atributo *name* do nodo ou variável aleatória para fazer esta associação, conforme ilustrado na Figura 38. É possível notar também que o elemento raiz da instância XML é formado pelo UUID do CCD utilizado para modelar os registros clínicos, assegurando a validação dos objetos e correspondência semântica.

```

01 <?xml version="1.0" encoding="UTF-8"?>
02 <mlhim2:ccd-122ea196-3460-46fe-996b-c265eef7c98c
    xmlns:mlhim2="http://www.mlhim.org/xmlns/mlhim2/2_4_0"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:data-view="http://www.w3.org/2003/g/data-view#"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.mlhim.org/xmlns/mlhim2/2_4_0
    file:/home/MLHIM/Flavio_CDS/ccd-122ea196-3460-46fe-996b-c265eef7c98c.xsd">
03 <mlhim2:e1-f69b4b89-facd-40b8-a833-07ecae87256f>
04   <mlhim2:annotation>[Clinical Diagnosis Description] rdf:resource="[Some URI for
    the diagnosis.]"</mlhim2:annotation>
05   <mlhim2:e1-82277514-3bdf-4579-95f0-0ba3ea912fe5>
06     <mlhim2:data-name>Clinical diagnosis</mlhim2:data-name>
07     <mlhim2:magnitude>1</mlhim2:magnitude>
08     <mlhim2:error>0</mlhim2:error>
09     <mlhim2:accuracy>0.0</mlhim2:accuracy>
10     <mlhim2:ratio-type>proportion</mlhim2:ratio-type>
11     <mlhim2:numerator>0.7556</mlhim2:numerator>
12     <mlhim2:denominator>1</mlhim2:denominator>
13     <mlhim2:numerator-units>numerator-units0</mlhim2:numerator-units>
14     <mlhim2:denominator-units>denominator-units0</mlhim2:denominator-units>
15   </mlhim2:e1-82277514-3bdf-4579-95f0-0ba3ea912fe5>
16 </mlhim2:e1-f69b4b89-facd-40b8-a833-07ecae87256f>
17 </mlhim2:ccd-122ea196-3460-46fe-996b-c265eef7c98c>
...

```

Figura 37. Resposta do CDSS codificada em XML, modelada em CCD.

Especificação da rede Bayesiana (formato XDSL)	
⋮	
25 ▾	<cpt id="N14" distribution="observation">
26	<state id="S_0_19" label="N14_S1" />
27	<state id="S_20_27" label="N14_S2" />
28	<state id="S_28" label="N14_S3" />
29	<state id="S_29_30" label="N14_S4" />
30	<parents>N6</parents>
31	<probabilities>0.2203 0.3589 0.1015 0.3193 0.7124 0.2698 0.0089 0.0089</probabilities>
32	</cpt>
⋮	
146 ▾	<node id="N14">
147	<name>Mini mental state score</name>
148	<interior color="ffffff" />
149	<outline color="0000bb" />
150	
151	<position>99 343 201 397</position>
152	<barchart active="true" width="173" height="111" />
153	</node>
⋮	
⋮	
Instância clínica modelada em CCD (formato XML)	
⋮	
15 ▾	<mlhim2:el-1b5fb55f-035c-4590-a1e2-4b271e9289ba>
16	<mlhim2:annotation>Mini mental state score</mlhim2:annotation>
17 ▾	<mlhim2:el-82277514-3bdf-4579-95f0-0ba3ea912fe4>
18	<mlhim2:data-name>Mini mental state score</mlhim2:data-name>
19	<mlhim2:DvOrdinal-dv>1</mlhim2:DvOrdinal-dv>
20	<mlhim2:symbol>0-19</mlhim2:symbol>
21	</mlhim2:el-82277514-3bdf-4579-95f0-0ba3ea912fe4>
22	</mlhim2:el-1b5fb55f-035c-4590-a1e2-4b271e9289ba>
⋮	
⋮	

Figura 38. Exemplo de evidência da rede Bayesiana, obtida pela associação da instância clínica modelada em CCD, e a variável aleatória representada na rede Bayesiana.

6.5 DESCRIÇÃO DOS APLICATIVOS CLIENTE

Para a implementação do protótipo dos serviços de publicação e consulta em plataforma de serviços web, utilizamos a plataforma de desenvolvimento da Microsoft Framework, e compilador Microsoft Studio Express 2010. O código foi gerado em C-Sharp (C#). Dois protótipos de cliente foram desenvolvidos utilizando ASP.NET, um protótipo de cliente para o serviço de publicação e outro para o serviço de consulta.

Cada protótipo de aplicação cliente é responsável por enviar a mensagem SOAP de requisição ao serviço correspondente e processar a mensagem SOAP de resposta. A aplicação cliente do serviço de consulta é também responsável pela interface com o usuário do CDSS, ou seja, o médico.

A Figura 39 mostra um desenho esquemático contendo as mensagens de requisição e resposta ao serviço de consulta do CDSS. A aplicação cliente utiliza os dados clínicos

fornechos pelo médico para montar uma instância XML. A instância XML é validada com consulta ao respectivo CCD (Figura 39 – 1). A mensagem SOAP de requisição segue então com os seguintes parâmetros: Diagnóstico, representando a doença e a base de treinamento que se deseja consultar, e a instância XML (Figura 39 – 2). Estas informações chegam ao serviço de consulta, que registra as evidências na rede Bayesiana, e executa o mecanismo de inferência, obtendo o diagnóstico (negativo ou positivo), a probabilidade associada ao diagnóstico, os itens relevantes observados para a hipótese de diagnóstico atual, e os itens ainda não observados que ajudariam na identificação do diagnóstico. Estas informações são encaminhadas nos parâmetros da mensagem SOAP de resposta (Figura 39 – 3). Por fim, estas informações são recebidas no aplicativo requisitante, que pode processá-las e exibi-las em uma interface gráfica (Figura 39 – 4).

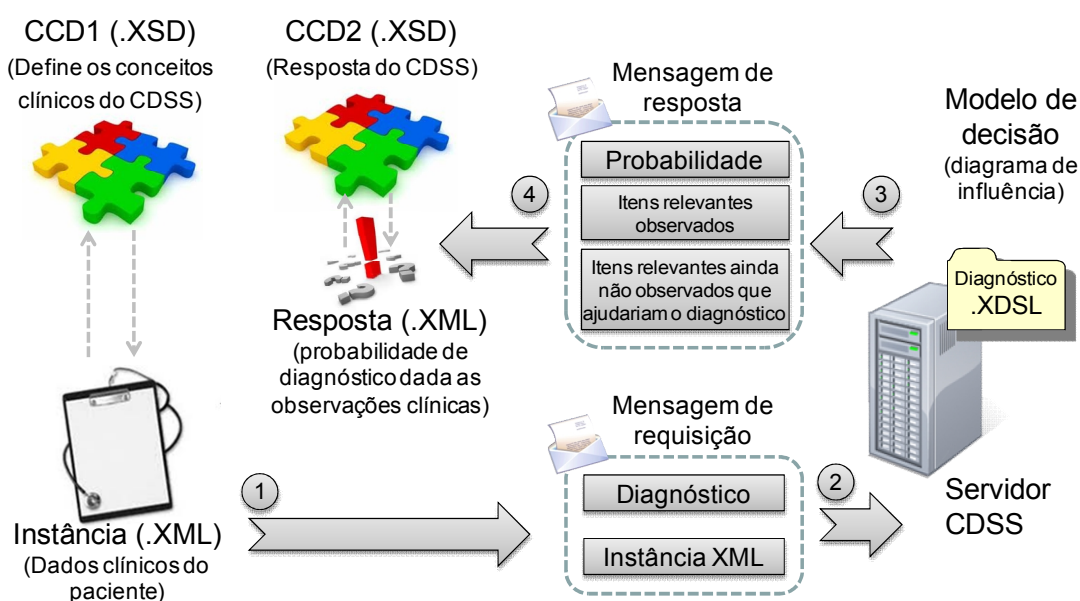


Figura 39. Esquema mostrando a sequência de mensagem de requisição e resposta, e os parâmetros encaminhados para o serviço de consulta, hospedado no servidor CDSS.

A Figura 40 mostra a tela do protótipo de cliente do serviço de publicação. Para publicar uma rede de decisão no servidor do CDSS, as seguintes tarefas são necessárias na aplicação: (1) selecionar a doença de interesse (ex., demência, doença de Alzheimer e transtorno cognitivo leve), (2) buscar o arquivo XDSL com o modelo de decisão (rede Bayesiana), (3) carregar o conteúdo do arquivo (pressionar botão *Load file*), (4) enviar o conteúdo ao servidor (pressionar botão *Send to Server*), (5) aguardar a resposta do servidor. Se a operação tiver sucesso, deverá aparecer a mensagem *Saved* na caixa de texto contendo a

resposta do servidor. O arquivo XDSL ficará armazenado em um local pré-definido no servidor para utilização do serviço de consulta.

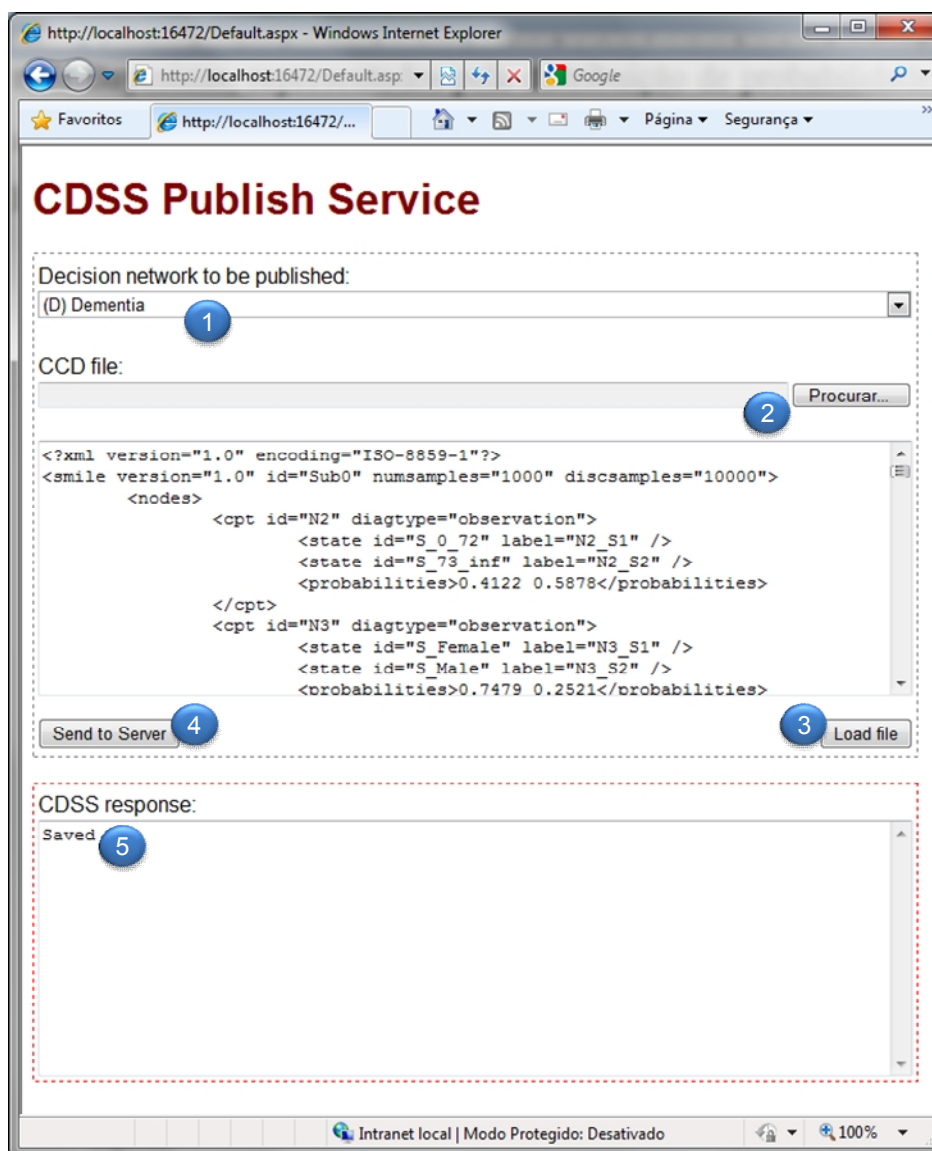


Figura 40. Cópia de tela com o aplicativo cliente do serviço de publicação da rede Bayesiana no servidor CDSS.

Para o serviço de consulta ao CDSS, foi desenvolvida uma aplicação web que disponibiliza um formulário eletrônico a ser preenchido pelo usuário (médico). O formulário disponibilizado pela aplicação para entrada das observações clínicas é construído em tempo de execução, a partir da seleção da doença de interesse para o diagnóstico. O formulário é especificado com base nos conceitos clínicos representados pelas variáveis aleatórias da rede Bayesiana. Cada controle do formulário corresponde a uma variável aleatória. Os controles são do tipo caixa de combinação do formulário (*DropDownList*). Cada item de seleção do

controle (*ListItem*) é associado a um nível discreto da variável aleatória correspondente. O nome (*label*) do controle deve ser igual ao da variável aleatória correspondente, conforme mostra a Figura 41. Os nomes dos controles devem ser formados apenas por caracteres alfanuméricos, sem diferenciação de maiúsculas e minúsculas, com os espaços entre os caracteres substituídos por sublinhado ("_").

Especificação da rede Bayesiana (formato XDSL)

```

:
25 <cpt id="N14" dist="observation">
26   <state id="S_0_19" label="N14_S1" />
27   <state id="S_20_27" label="N14_S2" />
28   <state id="S_28" label="N14_S3" />
29   <state id="S_29_30" label="N14_S4" />
30   <parents>N6</parents>
31   <probabilities>0.2203 0.3589 0.1015 0.3193 0.7124 0.2698 0.0089 0.0089</probabilities>
32 </cpt>
:
146 <node id="N14">
147   <name>Mini mental state score</name>
148   <interior color="ffff99" />
149   <outline color="0000bb" />
150   <font color="000000" name="Arial" size="8" />
151   <position>99 343 201 397</position>
152   <barchart active="true" width="173" height="111" />
153 </node>
:

```

Especificação do formulário dinâmico da aplicação de entrada dos dados clínicos (formato XML)

```

:
6 <FIELD type="DropDownList" label="Mini-mental state score:">
7   <PROPERTIES>
8     <PROPERTY name="ID">mini_mental_state_score</PROPERTY>
9     <PROPERTY name="WIDTH">200px</PROPERTY>
10  </PROPERTIES>
11  <LISTITEMS>
12    <LISTITEM value="">no_evidence</LISTITEM>
13    <LISTITEM value="1">0-19</LISTITEM>
14    <LISTITEM value="2">20-27</LISTITEM>
15    <LISTITEM value="3">28</LISTITEM>
16    <LISTITEM value="4">29-30</LISTITEM>
17  </LISTITEMS>
18 </FIELD>
:

```

Figura 41. Exemplo da associação dos controles do formulário dinâmico, com a especificação dos conceitos clínicos no modelo CCD.

A Figura 42 mostra a interface da aplicação cliente. Inicialmente, o usuário (médico) deve selecionar o diagnóstico desejado. Conforme explicado anteriormente, modelamos cinco redes Bayesianas, uma para cada doença e base de treinamento (Figura 42 – 2). Por exemplo, a seleção de *Dementia (CDA)*, indica que o médico deseja consultar o sistema para

diagnóstico de Demência, utilizando o modelo orientado pela base do CDA. O formulário é então construído com base nas informações deste modelo, e os controles disponibilizados para entrada das observações clínicas pelo médico (Figura 43 – 1). Depois de entradas as observações clínicas (Figura 43 – 2), o usuário deve pressionar o botão *Send to the CDSS*. Então, a instância XML é gerada automaticamente a partir das observações clínicas registradas no formulário, e uma mensagem SOAP de requisição é encaminhada ao servidor de consulta do CDSS.

Conforme descrito na Seção 6.4, a mensagem SOAP contém dois parâmetros: o diagnóstico desejado e a instância XML, baseada no respectivo CCD. O servidor CDSS envia então a mensagem SOAP de resposta com três parâmetros: a probabilidade do diagnóstico, a lista de itens relevantes observados no paciente e a lista de itens relevantes que ainda não foram observados que ajudariam a indicação de diagnóstico. Estas informações são processadas pela aplicação cliente e apresentadas na interface na forma de texto para visualização e análise do médico (Figura 43 – 3).

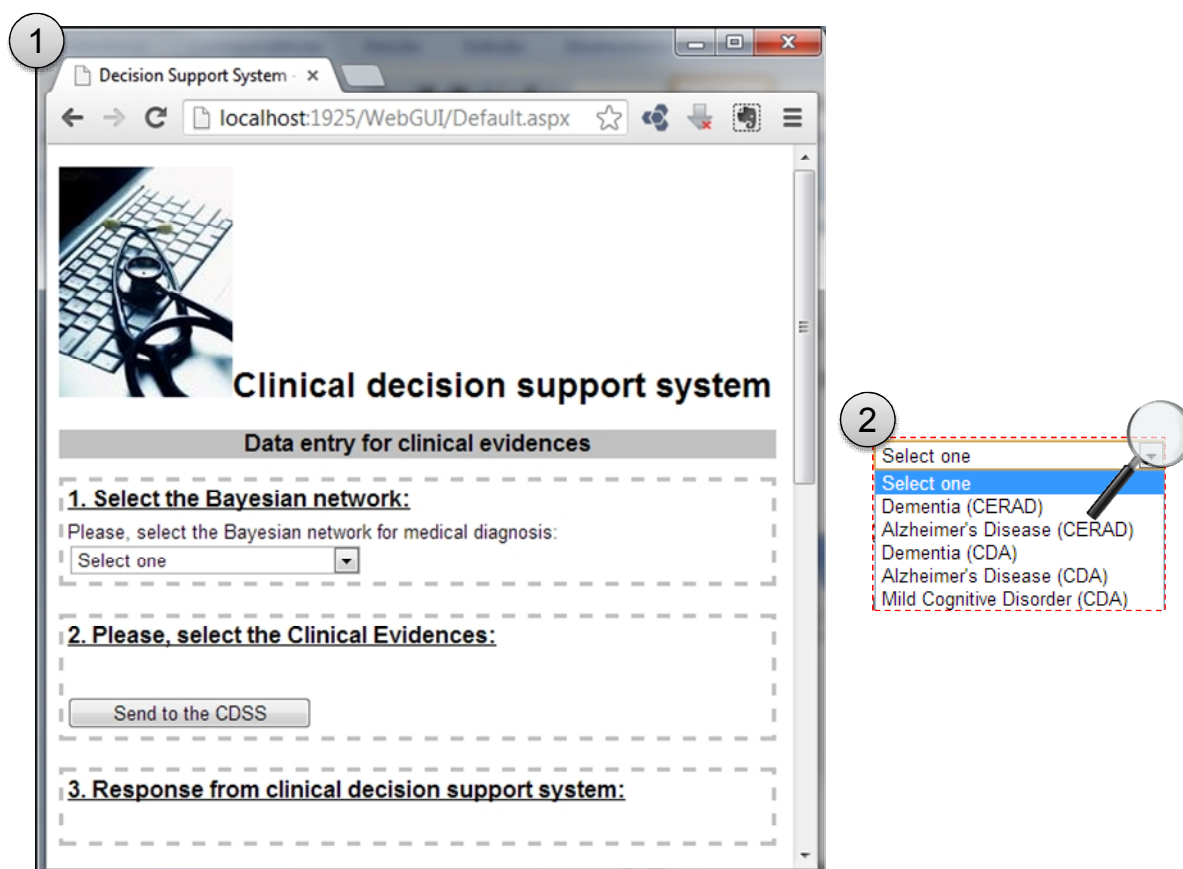


Figura 42. Cópia da tela do aplicativo cliente do serviço de consulta ao servidor CDSS.

1

Decision Support System - x
localhost:1925/WebGUI/Default.aspx?c

Clinical decision support system

Data entry for clinical evidences

1. Select the Bayesian network:
Please, select the Bayesian network for medical diagnosis:
Dementia (CDA)

2. Please, select the Clinical Evidences:

Age: no evidence
Gender: no evidence
Education: no evidence
Mini mental state score: no evidence
Clinical dementia rating scale: no evidence
Verbal fluency test score: no evidence
Pfeffer questionnaire score: no evidence
Clock drawing test scale: no evidence
Trial making test: no evidence
Stroop color word test: no evidence
Lawton scale: no evidence
IQCode score: no evidence
Berg balance scale: no evidence
Depression: no evidence

Send to the CDSS

2

Decision Support System - x
localhost:1925/WebGUI/Default.aspx?c

1. Select the Bayesian network:
Please, select the Bayesian network for medical diagnosis:
Dementia (CDA)

2. Please, select the Clinical Evidences:

Age: 0-72
Gender: female
Education: 0-13
✓ Mini mental state score: 0-17
★ Clinical dementia rating scale: no evidence
Verbal fluency test score: no evidence
Pfeffer questionnaire score: no evidence
Clock drawing test scale: no evidence
Trial making test: no evidence
Stroop color word test: no evidence
Lawton scale: no evidence
IQCode score: no evidence
Berg balance scale: no evidence
Depression: no evidence

✓ Items that led to the diagnosis hypothesis
★ Important items that would help indicating the diagnosis

Send to the CDSS

3

3. Response from clinical decision support system:

The probability of positive diagnosis of Dementia is about 75,56%, given the clinical context. Items that led to the diagnosis hypothesis were the following: Mini mental state score.

Important items that would help indicating the diagnosis: Clinical dementia rating scale.

Figura 43. Cópia das telas do aplicativo cliente do serviço de consulta ao servidor CDSS, mostrando três momentos: (1) seleção do diagnóstico ou rede Bayesiana desejada, (2) preenchimento do formulário com as observações clínicas, (3) visualização da resposta.

CAPÍTULO 7 - CONCLUSÃO

A segurança do paciente e o gerenciamento de riscos na área de saúde são os principais desafios da qualidade dos serviços assistenciais. Além disso, implantar uma rotina de registro das evidências médicas e gestão do conhecimento é fundamental para um gerenciamento eficiente do risco.

Com objetivo geral de melhorar a qualidade dos serviços na área de saúde e promover o registro das evidências médicas, este trabalho propôs um sistema de apoio à decisão, aplicado ao diagnóstico de demência, Doença de Alzheimer (DA) e Transtorno Cognitivo Leve (TCL). Estas doenças são relevantes, dada a sua elevada incidência entre os idosos (idade maior que 65 anos), ocupando a quinta posição entre as principais causas de morte desta população. O envelhecimento generalizado da população mundial devido às melhores condições ambientais e de higiene potencializa ainda mais o problema.

O CDSS proposto nesta tese é baseado em redes Bayesianas. As redes Bayesianas são apropriadas para modelos de decisão envolvendo incertezas e observações parciais da realidade, características presentes no domínio clínico. O mecanismo de inferência utilizado em redes Bayesianas é capaz de fornecer a relação de evidências mais relevantes para o resultado indicado. Esta informação pode ser utilizada pelo médico ou especialista do domínio do conhecimento para avaliação da resposta apresentada pelo CDSS, contribuindo para a confiabilidade do sistema. Além disso, a rede Bayesiana pode apresentar quais informações ainda não observadas seriam importantes para confirmação (ou refutação) da hipótese de diagnóstico atual.

Foi modelado um diagrama de influência, incluindo a rede Bayesiana, função utilidade e caixa de decisão para cada doença de interesse (Demência, DA e TCL). Cada caixa de decisão é relacionada a um ponto de decisão do processo de diagnóstico. Em uma simplificação do processo de diagnóstico destas doenças, assumimos decisões binárias, sequenciais, e não concorrentes. A DA é um subtipo da Demência e, assim, seu diagnóstico é subsequente ao diagnóstico de demência. A estrutura da rede Bayesiana foi construída orientada pelos dados da base de casos clínicos, posicionando os atributos no modelo causal conforme análise do seu significado, com ajuda do especialista do domínio do conhecimento. Para ajudar nesta tarefa, foi utilizada uma estrutura causal genérica. Os parâmetros da rede Bayesiana, compostos pela distribuição de probabilidades, foram estimados utilizando o algoritmo EM de aprendizagem supervisionada Bayesiana. Foram utilizadas duas bases de

casos clínicos: uma base do CERAD, um consórcio formado por instituições de pesquisa sobre padronizações de instrumentos de avaliação e diagnóstico de DA, e a outra base do CDA, vinculado ao Instituto de Psiquiatria da Universidade Federal do Rio de Janeiro (UFRJ). Foram modeladas cinco redes Bayesianas, uma para cada diagnóstico e base de treinamento.

As redes Bayesianas foram avaliadas utilizando medidas de desempenho e comparadas com outros classificadores conhecidos. Em geral, o modelo Bayesiano obteve um desempenho muito próximo ao melhor resultado apresentado pelos classificadores selecionados. No caso de TCL, a rede Bayesiana obteve as melhores medidas de AUC, F_1 e TPR na comparação realizada. Conclui-se, assim, que o modelo Bayesiano pode ser utilizado como classificador, obtendo resultados comparáveis aos demais modelos de classificação. Além disso, a rede Bayesiana reúne outras características interessantes ao contexto do diagnóstico clínico, como a análise de relevância das informações atuais e futuras, observações parciais, tratamento da incerteza, e simplicidade na representação dos critérios de diagnóstico.

Além da comparação com outros classificadores, foi realizada uma comparação das redes Bayesianas modeladas manualmente com redes obtidas de forma automática. Em geral, os resultados do método de descoberta automática da rede Bayesiana utilizando a base do CDA foram próximos aos obtidos com a estrutura da rede Bayesiana definida manualmente. No caso de DA, a rede descoberta obteve a medida de AUC superior ao da rede com estrutura pré-definida. A principal desvantagem do uso de métodos de descoberta automática da rede é a redução da legibilidade do modelo de decisão pelo aumento da sua complexidade.

É importante ressaltar que o foco principal da tese não foi apresentar a rede Bayesiana como um modelo de classificação com resultados que superariam outros classificadores, mas sim, representar critérios de diagnóstico médico em um modelo computacional inteligível com potencial de uso prático em sistemas e aplicações médicas, como o CDSS proposto. Sendo assim, conclui-se que o uso de redes Bayesianas é adequado para este fim.

Quanto aos testes neuropsicológicos relacionados pela análise de sensibilidade como relevantes para o diagnóstico das doenças em questão, eles apareceram na avaliação do especialista do domínio do conhecimento com graus de relevantes e muito relevantes, o que comprova a robustez do modelo Bayesiano proposto. Ainda, alguns testes neuropsicológicos classificados como relevantes pelo especialista foram desconsiderados na etapa de seleção dos atributos devido ao seu alto volume de dados incompletos ou ausentes.

Assim, é possível concluir que a modelagem de uma rede de decisão baseada em rede Bayesiana e diagrama de influência, aplicada a área de saúde como apoio ao diagnóstico clínico de Demência, Doença de Alzheimer, e Transtorno Cognitivo Leve, doenças pertencentes ao conjunto de doenças de transtorno mental, apresentou resultados promissores, com base nas medidas de avaliação de desempenho descritas nesta tese. O mapeamento dos critérios de diagnóstico destas doenças em uma estrutura computacional é viável e consistente, devido à capacidade do modelo de decisão de inferir um diagnóstico sobre condições de incerteza, utilizando a causalidade de características clínicas avaliadas durante o diagnóstico.

A relação dos itens mais relevantes para o resultado de diagnóstico apresentado é uma importante característica dos modelos Bayesianos. Além disso, através de uma rotina de avaliação de sensibilidade, o modelo Bayesiano fornece a relação de itens mais relevantes ainda não observados que ajudariam na identificação do diagnóstico. Estas características também foram exploradas no CDSS proposto nesta tese.

A modelagem dos dados manipulados pelo CDSS proposto foi baseada na linguagem XML. Este padrão apresenta inúmeras vantagens técnicas, por exemplo, é baseado em texto, permite representar estruturas simples e complexas de dados, e disponibiliza um formato de autodocumentação e validação dos valores. Estas características são importantes na manutenção e integração do CDSS a outros sistemas de informações médicas mais abrangentes. Na modelagem multinível, as informações são modeladas utilizando, pelo menos, dois níveis: um nível de referência, representando os dados básicos ou primários, e um nível de domínio, representando o conceito propriamente dito. Esta abordagem oferece também inúmeras vantagens técnicas, principalmente em relação à interoperabilidade semântica, característica importante na troca de dados entre diferentes sistemas.

7.1 CONTRIBUIÇÕES

As principais contribuições desta tese de doutorado são:

1. Desenvolvimento de um modelo causal aplicado ao apoio à decisão ao diagnóstico clínico, incluindo:
 - (a) Desenho do processo de diagnóstico;
 - (b) Especificação das características observadas no diagnóstico, em uma abordagem orientada por dados;

- (c) Tratamento da incerteza gerada pelas observações parciais utilizando a teoria geral das probabilidades e o teorema de Bayes;
 - (d) Modelo de decisão representado por diagrama de influência;
 - (e) Integração do modelo de decisão ao processo de diagnóstico;
 - (f) Avaliação do desempenho e robustez do modelo proposto por medidas quantitativas e análises qualitativas.
2. Aplicação do modelo causal proposto ao diagnóstico de Demência, DA e TCL.
 3. Modelagem das informações clínicas e pontuações dos instrumentos de avaliação neuropsicológicos, relacionados ao diagnóstico de Demência, DA e TCL, aplicando abordagem multinível e o modelo de referência MLHIM (*Multilevel Healthcare Information Modeling*). Este modelo está publicado em um repositório de componentes de conhecimento da área de saúde, denominado HKCR (*Healthcare Knowledge Component Repository*), disponível na web⁴.
 4. Desenvolvimento de um protótipo do sistema de apoio à decisão clínica proposto utilizando a arquitetura orientada a serviços web SOAP.

7.2 TRABALHOS FUTUROS

Neste trabalho, propusemos um processo de modelagem da rede de decisão genérico e extensível a outros domínios ou doenças. Visando validar a generalidade do processo de modelagem e identificar possíveis refinamentos, é importante sua aplicação em outros domínios de diagnóstico clínico, o que será realizado como trabalho futuro.

Conforme verificado, o desempenho e robustez da rede Bayesiana são influenciados pela qualidade e abrangência dos dados ou casos clínicos presentes na base de treinamento. Assim, é fundamental uma base de casos clínicos, mais completa e abrangente possível, incluindo todos os atributos considerados relevantes no diagnóstico, e um percentual menor de dados incompletos. Outro trabalho futuro é a identificação de outras bases de dados que possam ser utilizadas para testar o sistema desenvolvido.

Outra atividade é instalar o CDSS em uma rotina clínica, realizar testes de usabilidade, identificar possíveis refinamentos na modelagem e arquitetura do sistema, e medir a ocorrência de erros clínicos antes e depois da sua implantação, para realizar possíveis refinamentos futuros.

⁴ <http://www.hkcr.net/ccd/ccd-122ea196-3460-46fe-996b-c265eef7c98c.xsd/view>

O relacionamento causal entre variáveis aleatórias, representado graficamente por arcos direcionados entre nodos da rede Bayesiana, pode significar também características temporais a serem observadas durante o processo de tratamento. Assim, um item a ser considerado como trabalho futuro é a exploração desta importante característica expressa nas redes Bayesianas, considerando também o momento que um determinado sintoma é observado no paciente, ou um teste de avaliação neuropsicológica é aplicado, ao longo do curso da doença.

A notação BPMN foi empregada para representação gráfica do processo de diagnóstico das doenças de interesse. Este padrão oferece também uma representação de processos baseada em XML, denominado XPDL (*XML Process Definition Language*), formato padronizado pela WfMC (*Workflow Management Coalition*), uma organização formada por Universidades e instituições de pesquisa visando promover padronização de modelos de representação de processos ou fluxos de trabalho. Como trabalho futuro, é possível estudar uma forma de relacionar os pontos de decisão do processo com as caixas de decisão do diagrama de influência, promovendo maior integração entre estes domínios.

Além disso, vimos que um mesmo conceito clínico é abordado de três formas diferentes, representando momentos distintos na modelagem do CDSS: (1) modelagem do diagrama de influência (formato XDSL), (2) modelagem dos conceitos clínicos (CCD) para identificação das evidências (formato XSD), e (3) especificação do formulário dinâmico da aplicação web (formato HTML). Os formatos das três especificações são baseados em XML. Nesta tese, aplicamos a abordagem multinível para modelagem dos conceitos clínicos utilizados pelo CDSS. Uma proposta para trabalho futuro é aplicar a abordagem multinível também na modelagem do diagrama de influência e formulário dinâmico da aplicação web, visando melhorar a integração destas três especificações em um único *framework*, assegurando a semântica e integridade dos dados.

Em relação à modelagem dos conceitos clínicos, uma proposta de extensão e pesquisa é completar as avaliações neuropsicológicas com os detalhes e observações clínicas utilizadas internamente em cada avaliação, utilizando uma modelagem multinível, como o MLHIM.

Algumas melhorias operacionais podem ser enumeradas em relação ao CDSS, tais como, (1) melhorar a rotina de publicação do diagrama de influência, incluindo um mecanismo de gerenciamento de mudanças e liberações, e (2) integrar o protótipo de interface de comunicação do CDSS a um sistema de informações médicas.

A integração com um sistema de informações médicas pode ajudar a implantação do CDSS na rotina clínica. Ainda, é importante o desenvolvimento de um mecanismo de

realimentação do CDSS, trazendo informações do diagnóstico do médico. Em adição aos ciclos de reaprendizagem do modelo, estas funcionalidades podem melhorar a acuidade e robustez do modelo de decisão, fornecendo mais confiabilidade ao CDSS.

REFERÊNCIAS

ALBERT, M. S.; DEKOSKY, S. T.; DICKSON, D.; DUBOIS, B.; FELDMAN, H. H.; FOX, N. C.; GAMST, A.; HOLTZMAN, D. M.; JAGUST, W. J.; PETERSEN, R. C.; SNYDER, P. J.; CARRILLO, M. C.; THIES, B.; PHELPS, C. H. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. **Alzheimer's & dementia : the journal of the Alzheimer's Association**, v. 7, n. 3, p. 270-279, 2011. ISSN 1552-5260.

ALMEIDA, O. P. Mini exame do estado mental e o diagnóstico de demência no Brasil. **Arq Neuropsiquiatr**, v. 56, n. 3B, p. 605-12, 1998.

ALVERNAZ, C.; COOK, T. W.; CAVALINI, L. T. Migration of a pre-hospital cardiology emergency system from data model to multilevel modeling. **ACM SIGHT Record**, v. 2, n. 1, p. 9-9, 2012.

ANDERSON, J. R.; MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. **Machine learning: An artificial intelligence approach**. Morgan Kaufmann, 1986. ISBN 0934613001.

ANTONIOU, G.; HARMELEN, F. Web ontology language: Owl. **Handbook on ontologies**, p. 91-110, 2009.

APA, A. P. A. **Diagnostic and statistical manual of mental disorders (DSM-IV-TR)**. 4th edition, text revision. Washington-DC 2000.

ARAÚJO DE CASTRO, A.; PINHEIRO, P.; PINHEIRO, M. A hybrid model for aiding in decision making for the neuropsychological diagnosis of Alzheimer's disease. **Rough Sets and Current Trends in Computing**, p. 495-504, 2008.

ARIAS, M.; DIEZ, F.; PALACIOS, M. **ProbModelXML: a format for encoding probabilistic graphical models**. Technical Report CISIAD-11-02, UNED, Madrid, Spain. 2011

ASSOCIATION, A. S. **Alzheimer's Disease Facts and Figures**. Alzheimer's Association. Washington. 2012

ATALAG, K.; YANG, H.; TEMPERO, E.; WARREN, J. Model Driven Development of Clinical Information Systems using openEHR. **Studies in health technology and informatics**, v. 169, p. 849, 2011. ISSN 0926-9630.

BALL, L.; BISHOP, G.; BIRGE, S. A simple test of central processing speed: an extension of the Short Blessed Test. **Journal of the American Geriatrics Society**, v. 47, n. 11, p. 1359, 1999. ISSN 0002-8614.

BARZOTTI, T.; GARGIULO, A.; MAROTTA, M.; TEDESCHI, G.; ZANNINO, G.; GUGLIELMI, S.; DELL'ARMI, A.; ETTORRE, E.; MARIGLIANO, V. Correlation between

cognitive impairment and the Rey Auditory-Verbal Learning Test in a population with Alzheimer disease. **Archives of gerontology and geriatrics. Supplement**, n. 9, p. 57-62, 2004. ISSN 0924-7947.

BAUM, L. E.; PETRIE, T. Statistical inference for probabilistic functions of finite state Markov chains. **The Annals of Mathematical Statistics**, v. 37, n. 6, p. 1554-1563, 1966. ISSN 0003-4851.

BEALE, T. Archetypes and the EHR. **Studies in health technology and informatics**, p. 238-246, 2003. ISSN 0926-9630.

BEALE, T.; HEARD, S. An ontology-based model of clinical information. **Studies in health technology and informatics**, v. 129, n. 1, p. 760-764, 2007.

BECK, I. R.; GAGNEUX-ZURBRIGGEN, A.; BERRES, M.; TAYLOR, K. I.; MONSCH, A. U. Comparison of Verbal Episodic Memory Measures: Consortium to Establish a Registry for Alzheimer's Disease—Neuropsychological Assessment Battery (CERAD-NAB) versus California Verbal Learning Test (CVLT). **Archives of Clinical Neuropsychology**, v. 27, n. 5, p. 510-519, August 1, 2012 2012.

BELIAKOV, G.; WARREN, J. Fuzzy logic for decision support in chronic care. **Artificial Intelligence in Medicine**, v. 21, n. 1, p. 209-213, 2001.

BELLAZZI, R.; ZUPAN, B. Predictive data mining in clinical medicine: current issues and guidelines. **International Journal of Medical Informatics**, v. 77, n. 2, p. 81-97, 2008. ISSN 1386-5056.

BENJAMINI, Y. Opening the Box of a Boxplot. **The American Statistician**, v. 42, n. 4, p. 257-262, 1988. ISSN 0003-1305.

BERGER, J. O. **Statistical decision theory and Bayesian analysis**. Springer, 1985. ISBN 0387960988.

BERLIN, A.; SORANI, M.; SIM, I. A taxonomic description of computer-based clinical decision support systems. **J Biomed Inform**, v. 39, n. 6, p. 656-67, 2006. ISSN 1532-0480 (Electronic).

BERNER, E. S. **Clinical decision support systems: theory and practice**. Birmingham: Springer, 2007.

BERNER, E. S.; GRABER, M. L. Overconfidence as a cause of diagnostic error in medicine. **The American Journal of Medicine**, v. 121, n. 5 Suppl, p. S2, 2008. ISSN 0002-9343.

BERNER, E. S.; MAISIAK, R. S.; HEUDERBERT, G. R.; YOUNG, K. R., JR. Clinician performance and prominence of diagnoses displayed by a clinical diagnostic decision support system. **AMIA Annu Symp Proc**, p. 76-80, 2003. ISSN 1942-597X (Electronic).

BERTOLUCCI, P. H. F.; OKAMOTO, I. H.; BRUCKI, S. M. D.; SIVIERO, M. O.; TONIOLO NETO, J.; RAMOS, L. R. Applicability of the CERAD neuropsychological

battery to Brazilian elderly. **Arquivos de Neuro-Psiquiatria**, v. 59, n. 3A, p. 532-536, 2001. ISSN 0004-282X.

BIN, Z.; YUAN-HSIANG, C.; XIAO-HUI, W.; GOOD, W. F. Comparison of artificial neural network and Bayesian belief network in a computer-assisted diagnosis scheme for mammography. *Neural Networks*, 1999. IJCNN '99. International Joint Conference on, 1999. Jul 1999. p.4181-4185 vol.6.

BOTTINO, C. M. C.; AZEVEDO, D.; TATSCH, M.; HOTOTIAN, S. R.; MOSCOSO, M. A.; FOLQUITTO, J.; SCALCO, A. Z.; BAZZARELLA, M. C.; LOPES, M. A.; LITVOC, J. Estimate of dementia prevalence in a community sample from São Paulo. **Dementia and Geriatric Cognitive Disorders**, v. 26, n. 4, p. 291-299, 2008.

BRACHMAN, R. J.; LEVESQUE, H. J. **Knowledge representation and reasoning**. Morgan Kaufmann Pub, 2004. ISBN 1558609326.

BRAY, T.; PAOLI, J.; SPERBERG-MCQUEEN, C. M.; MALER, E.; YERGEAU, F. Extensible markup language (XML). **World Wide Web Journal**, v. 2, n. 4, p. 27-66, 1997.

BRIER, G. W. Verification of forecasts expressed in terms of probability. **Monthly weather review**, v. 78, n. 1, p. 1-3, 1950. ISSN 1520-0493.

BUCHANAN, B. G.; SHORTLIFFE, E. H. **Rule-based expert systems: the MYCIN experiments of the Stanford Heuristic Programming Project**. Addison-Wesley Reading, MA, 1984. ISBN 0201101726.

BUNTINE, W. Theory refinement on Bayesian networks. Proceedings of the seventh conference on Uncertainty in artificial intelligence, 1991. Los Angeles. Morgan Kaufmann. p.52-60.

CAMPOS, L. M. D. A Scoring Function for Learning Bayesian Networks based on Mutual Information and Conditional Independence Tests. **Journal of Machine Learning Research**, v. 7, p. 2149-2187, 2006. ISSN 1532-4435.

CANTRILL, M.; STEPHEN, V. Computers in patient care: the promise and the challenge. **Queue**, v. 8, n. 8, p. 20, 2010. ISSN 1542-7730.

CARTHERY-GOULART, M. T.; AREZA-FEGYVERES, R.; SCHULTZ, R. R.; OKAMOTO, I.; CARAMELLI, P.; BERTOLUCCI, P. H. F.; NITRINI, R. Cornell depression scale in dementia. **Arquivos de Neuro-Psiquiatria**, v. 65, p. 912-915, 2007. ISSN 0004-282X.

CARVALHO, R. N.; LASKEY, K. B.; COSTA, P. C. G. PR-OWL 2.0-bridging the gap to OWL semantics. 2010. p.73-84.

CAVALINI, L. T.; COOK, T. Health Informatics: The Relevance of Open Source and Multilevel Modeling. **Open Source Systems: Grounding Research**, v. 365, p. 338-347, 2011.

CAVALINI, L. T.; COOK, T. W. Knowledge engineering of healthcare applications based on minimalist multilevel models. *e-Health Networking, Applications and Services (Healthcom)*, 2012 IEEE 14th International Conference on, 2012. IEEE. p.431-434.

CECCARELLI, M.; DONATIELLO, A.; VITALE, D. KON³: A Clinical Decision Support System, in *Oncology Environment, Based on Knowledge Management. Tools with Artificial Intelligence*, 2008. ICTAI '08. 20th IEEE International Conference on, 2008. p.206-210.

CELSIS, P. Age-related cognitive decline, mild cognitive impairment or preclinical Alzheimer's disease? *Annals of medicine*, v. 32, n. 1, p. 6-14, 2000.

CERAMI, E. **Web services essentials: distributed applications with XML-RPC, SOAP, UDDI & WSDL**. O'Reilly Media, 2002. ISBN 0596002246.

CHANDRA, A.; NUNDY, S.; SEABURY, S. A. The growth of physician medical malpractice payments: evidence from the National Practitioner Data Bank. *HEALTH AFFAIRS-MILLWOOD VA THEN BETHESDA MA-*, v. 24, n. 4, p. 1186, 2005. ISSN 0278-2715.

CHAUDHRY, B.; WANG, J.; WU, S.; MAGLIONE, M.; MOJICA, W.; ROTH, E.; MORTON, S. C.; SHEKELLE, P. G. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med*, v. 144, n. 10, p. 742-52, 2006. ISSN 1539-3704 (Electronic).

CHAVES, M. L.; CAMOZZATO, A. L.; GODINHO, C.; PIAZENSKI, I.; KAYE, J. Incidence of mild cognitive impairment and Alzheimer disease in Southern Brazil. *J Geriatr Psychiatry Neurol*, v. 22, n. 3, p. 181-7, 2009. ISSN 0891-9887.

CHAWLA, N. V.; BOWYER, K. W.; HALL, L. O.; KEGELMEYER, W. P. SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of artificial intelligence research*, v. 16, p. 321-357, 2002.

CHEN, R. **Towards Interoperable and Knowledge-Based Electronic Health Records Using Archetype Methodology**. Linköping University Electronic Press, 2009. ISBN 9173935042.

CHEN, R. S.; NADKARNI, P.; MARENCO, L.; LEVIN, F.; ERDOS, J.; MILLER, P. L. Exploring performance issues for a clinical database organized using an entity-attribute-value representation. *Journal of the American Medical Informatics Association*, v. 7, n. 5, p. 475, 2000. ISSN 1527-974X.

CHIN-JEN, K.; FINE, T. L. A Bayesian Independence Test for Small Datasets. *Signal Processing, IEEE Transactions on*, v. 54, n. 10, p. 4026-4031, 2006. ISSN 1053-587X.

CHINNICI, R.; MOREAU, J. J.; RYMAN, A.; WEERAWARANA, S. Web services description language (wsdl) version 2.0 part 1: Core language. **W3C recommendation**, v. 26, 2007.

CHOI, J.; CURRIE, L. M.; WANG, D.; BAKKEN, S. Encoding a clinical practice guideline using guideline interchange format: a case study of a depression screening and management

guideline. **Int J Med Inform**, v. 76 Suppl 2, n. 2, p. S302-7, 2007. ISSN 1872-8243 (Electronic).

CHRISTOFOLETTI, G.; OLIANI, M.; GOBBI, L.; GOBBI, S.; STELLA, F. Risco de quedas em idosos com doença de Parkinson e demência de Alzheimer: um estudo transversal. **Rev Bras Fisioter**, v. 10, n. 4, p. 429-33, 2006.

CONKLIN, H. M.; CURTIS, C. E.; KATSANIS, J.; IACONO, W. G. Verbal working memory impairment in schizophrenia patients and their first-degree relatives: evidence from the digit span task. **American Journal of Psychiatry**, v. 157, n. 2, p. 275-277, 2000. ISSN 0002-953X.

COOPER, G. F. The computational complexity of probabilistic inference using Bayesian belief networks. **Artificial intelligence**, v. 42, n. 2-3, p. 393-405, 1990. ISSN 0004-3702.

COOPER, G. F.; HERSKOVITS, E. A Bayesian method for the induction of probabilistic networks from data. **Machine learning**, v. 9, n. 4, p. 309-347, 1992. ISSN 0885-6125.

CORRIGAN, J.; DONALDSON, M.; KOHN, L.; MAGUIRE, S.; PIKE, K. Crossing the quality chasm: a new health system for the 21st century. In: STATES, I. O. M. O. T. U. (Ed.). Washington: National Academy Press, 2001.

COSTA, P. C. G.; LASKEY, K. B. PR-OWL: A framework for probabilistic ontologies. 2006. IOS Press. p.237-249.

COVER, T. H.; THOMAS, J. A. **Elements of Information Theory**. New York: John Wiley&Sons, 1991.

COWELL, R. G.; DAWID, P.; LAURITZEN, S. L.; SPIEGELHALTER, D. J. **Probabilistic Networks and Expert Systems: Exact Computational Methods for Bayesian Networks**. Springer, 2007. ISBN 9780387718231.

CUMMINGS, J. L.; MEGA, M.; GRAY, K.; ROSENBERG-THOMPSON, S.; CARUSI, D. A.; GORNBEIN, J. The Neuropsychiatric Inventory. **Neurology**, v. 44, n. 12, p. 2308, December 1, 1994 1994.

CURBERA, F.; LEYMANN, F.; STOREY, T.; FERGUSON, D.; WEERAWARANA, S. **Web services platform architecture: SOAP, WSDL, WS-policy, WS-addressing, WS-BPEL, WS-reliable messaging and more**. Prentice Hall PTR, 2005.

CZIBULA, G.; CZIBULA, I. G.; COJOCAR, G. S.; GURAN, A. M. IMASC - An Intelligent MultiAgent System for Clinical Decision Support. Complexity and Intelligence of the Artificial and Natural Complex Systems, Medical Applications of the Complex Systems, Biomedical Computing, 2008. CANS '08. First International Conference on, 2008. p.185-190.

DASH, D.; DRUZDZEL, M. J. **Robust independence testing for constraint-based learning of causal structure**. Proceedings of the Nineteenth conference on Uncertainty in Artificial Intelligence. Acapulco, Mexico: Morgan Kaufmann Publishers Inc.: 167-174 p. 2003.

DE CARVALHO, J. A.; GARCIA, R. A. The aging process in the Brazilian population: a demographic approach. **Cad Saude Publica**, v. 19, n. 3, p. 725-33, May-Jun 2003. ISSN 0102-311X (Print).

DEAN, T.; KANAZAWA, K. A model for reasoning about persistence and causation. **Computational intelligence**, v. 5, n. 2, p. 142-150, 1989. ISSN 1467-8640.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 39, n. 1, p. 1-38, 1977. ISSN 0035-9246.

DEN HEIJER, T.; VAN DER LIJN, F.; KOUDSTAAL, P. J.; HOFMAN, A.; VAN DER LUGT, A.; KRESTIN, G. P.; NIESSEN, W. J.; BRETELER, M. M. B. A 10-year follow-up of hippocampal volume on magnetic resonance imaging in early dementia and cognitive decline. **Brain**, v. 133, n. 4, p. 1163-1172, 2010. ISSN 0006-8950.

DENEKAMP, Y.; PELEG, M. TiMeDDx - a multi-phase anchor-based diagnostic decision-support model. **Journal of Biomedical Informatics**, v. 43, p. 111-124, 2010.

DEVANAND, D.; PRADHABAN, G.; LIU, X.; KHANDJI, A.; DE SANTI, S.; SEGAL, S.; RUSINEK, H.; PELTON, G.; HONIG, L.; MAYEUX, R. Hippocampal and entorhinal atrophy in mild cognitive impairment: prediction of Alzheimer disease. **Neurology**, v. 68, n. 11, p. 828, 2007.

DIEZ, F. J. Parameter adjustment in bayes networks. the generalized noisy or-gate. 1993. Morgan Kaufmann Publishers Inc. p.99-105.

DIEZ, F. J.; MIRA, J.; ITURRALDE, E.; ZUBILLAGA, S. DIAVAL: a Bayesian expert system for echocardiography. **Artificial Intelligence in Medicine**, v. 10, n. 1, p. 59-73, 1997.

DING, Z.; PENG, Y.; PAN, R. BayesOWL: Uncertainty modeling in semantic web ontologies. **Soft Computing in Ontologies and Semantic Web**, p. 3-29, 2006.

DO AMARAL, M. B.; SATOMURA, Y.; HONDA, M.; SATO, T. A psychiatric diagnostic system integrating probabilistic and categorical reasoning. **Methods Inf Med**, v. 34, n. 3, p. 232-43, Jun 1995. ISSN 0026-1270 (Print).

DODIER, R. A revision of the Bayesian network interchange format. **Technical report, JCEM-TR-1998-9**, 1998.

DOLIN, R. H.; ALSCHULER, L.; BEEBE, C.; BIRON, P. V.; BOYER, S. L.; ESSIN, D.; KIMBER, E.; LINCOLN, T.; MATTISON, J. E. The HL7 clinical document architecture. **Journal of the American Medical Informatics Association**, v. 8, n. 6, p. 552-569, 2001. ISSN 1527-974X.

DOUGHERTY, J.; KOHAVI, R.; SAHAMI, M. Supervised and unsupervised discretization of continuous features. MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-, 1995. Morgan Kaufmann Publishers, Inc. p.194-202.

DRUZDZEL, M. J. SMILE: Structural Modeling, Inference, and Learning Engine and GeNie: a development environment for graphical decision-theoretic models. 1999. JOHN WILEY & SONS LTD. p.902-903.

DUBOIS, B.; FELDMAN, H. H.; JACOVA, C.; DEKOSKY, S. T.; BARBERGER-GATEAU, P.; CUMMINGS, J.; DELACOURTE, A.; GALASKO, D.; GAUTHIER, S.; MEGURO, K.; O'BRIEN, J.; PASQUIER, F.; ROBERT, P.; ROSSOR, M.; SALLOWAY, S.; STERN, Y.; VISSER, P. J.; SCHELTENS, P. Research criteria for the diagnosis of Alzheimer's Disease: revising the NINCDS-ADRDA criteria. **Lancet Neurology**, v. 6, p. 734-746, 2007.

ELOMAA, T.; ROUSU, J. Efficient multisplitting revisited: Optima-preserving elimination of partition candidates. **Data Mining and Knowledge Discovery**, v. 8, n. 2, p. 97-126, 2004. ISSN 1384-5810.

ELSTEIN, A. S.; SCHWARZ, A. Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. **BMJ**, v. 324, n. 7339, p. 729-732, 2002. ISSN 0959-8138.

FAYYAD, U. M.; IRANI, K. B. Multi-interval discretisation of continuous-valued attributes for classification. Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, 1993. San Francisco CA. Morgan Kaufmann. p.1022-1027.

FIAT, A.; KAPLAN, H. **Making data structures confluent persistent**. Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms. Washington, D.C., United States: Society for Industrial and Applied Mathematics: 537-546 p. 2001.

FILLENBAUM, G. G.; VAN BELLE, G.; MORRIS, J. C.; MOHS, R. C.; MIRRA, S. S.; DAVIS, P. C.; TARIOT, P. N.; SILVERMAN, J. M.; CLARK, C. M.; WELSH-BOHMER, K. A. Consortium to Establish a Registry for Alzheimer's Disease (CERAD): The first twenty years. **Alzheimer's & dementia: the journal of the Alzheimer's Association**, v. 4, n. 2, p. 96-109, 2008. ISSN 1552-5260.

FOLSTEIN, M. F.; FOLSTEIN, S. E.; MCHUGH, P. R. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. **J Psychiatr Res**, v. 12, n. 3, p. 189-98, 1975. ISSN 0022-3956 (Print).

FOX, J.; PATKAR, V.; THOMSON, R. Decision support for health care: the PROforma evidence base. **Informatics in Primary Care**, v. 14, 2006.

FREUND, Y.; SCHAPIRE, R. A decision-theoretic generalization of on-line learning and an application to boosting. Computational learning theory, 1995. Springer. p.23-37.

FREUND, Y.; SCHAPIRE, R. E. Experiments with a new boosting algorithm. 1996. MORGAN KAUFMANN PUBLISHERS, INC. p.148-156.

FRIEDMAN, N. Learning belief networks in the presence of missing values and hidden variables. Proceedings of the Fourteenth International Conference on Machine Learning, 1997. San Francisco CA. Morgan Kaufmann. p.125-133.

FRIEDMAN, N.; GEIGER, D.; GOLDSZMIDT, M. Bayesian network classifiers. **Machine learning**, v. 29, n. 2, p. 131-163, 1997. ISSN 0885-6125.

FUNG, R.; CHANG, K. Weighing and integrating evidence for stochastic simulation in Bayesian networks. 1990. North-Holland Publishing Co. p.220.

GALANTER, W. L.; DIDOMENICO, R. J.; POLIKAITIS, A. Preventing exacerbation of an ADE with automated decision support. **J Healthc Inf Manag**, v. 16, n. 4, p. 44-9, 2002. ISSN 1099-811X (Print).

GALASKO, D.; BENNETT, D.; SANO, M.; ERNESTO, C.; THOMAS, R.; GRUNDMAN, M.; FERRIS, S. An inventory to assess activities of daily living for clinical trials in Alzheimer's disease. The Alzheimer's Disease Cooperative Study. **Alzheimer disease and associated disorders**, v. 11, p. S33, 1997. ISSN 0893-0341.

GARG, A. X.; ADHIKARI, N. K. J.; MCDONALD, H.; ROSAS-ARELLANO, M. P.; DEVEREAUX, P. J.; BEYENE, J.; SAM, J.; HAYNES, R. B. Effects of Computerized Clinical Decision Support Systems on Practitioner Performance and Patient Outcomes: A Systematic Review. **Jama**, v. 293, n. 10, p. 1223-1238, 2005.

GEIGER, D.; HECKERMAN, D.; MEEK, C. Asymptotic model selection for directed networks with hidden variables. **NATO ASI SERIES D BEHAVIOURAL AND SOCIAL SCIENCES**, v. 89, p. 461-478, 1998. ISSN 0258-123X.

GEIGER, D.; VERMA, T.; PEARL, J. Identifying independence in Bayesian networks. **Networks**, v. 20, n. 5, p. 507-534, 1990. ISSN 1097-0037.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. CRC press, 2004. ISBN 158488388X.

GLYMOUR, C.; MADIGAN, D.; PREGIBON, D.; SMYTH, P. Statistical inference and data mining. **Communications of the ACM**, v. 39, n. 11, p. 35-41, 1996. ISSN 0001-0782.

GRABER, M. L.; FRANKLIN, N.; GORDON, R. Diagnostic error in internal medicine. **Archives of Internal Medicine**, v. 165, n. 13, p. 1493, 2005.

GRIMSHAW, J. M.; RUSSELL, I. T. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. **The Lancet**, v. 342, n. 8883, p. 1317-1322, 1993. ISSN 0140-6736.

GRÜNWARD, P. D.; MYUNG, I. J.; PITT, M. A. **Advances in minimum description length: Theory and applications**. MIT press, 2005. ISBN 0262072629.

HAMDAN, A. C.; BUENO, O. F. A. Relações entre controle executivo e memória episódica verbal no comprometimento cognitivo leve e na demência tipo Alzheimer. **Estudos de Psicologia**, v. 10, n. 1, p. 63-71, 2005.

HANLEY, J. A. Characteristic (ROC) Curve. **Radiology**, v. 743, p. 29-36, 1982.

HART, A.; WYATT, J. Evaluating black-boxes as medical decision aids: issues arising from a study of neural networks. **Informatics for Health and Social Care**, v. 15, n. 3, p. 229-236, 1990.

HART, R. P.; KWENTUS, J. A.; WADE, J. B.; HAMER, R. M. Digit symbol performance in mild dementia and depression. **Journal of Consulting and Clinical Psychology**, v. 55, n. 2, p. 236, 1987. ISSN 1939-2117.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**: Springer New York: 2001.

HAYNES, R. B.; WILCZYNSKI, N. L. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: methods of a decision-maker-researcher partnership systematic review. **Implement Sci**, v. 5, n. 12, p. 12, 2010. ISSN 1748-5908.

HENRION, M. Propagating uncertainty by logic sampling in Bayes' networks. 1988. p.317-324.

HERRERA, E., JR.; CARAMELLI, P.; SILVEIRA, A. S.; NITRINI, R. Epidemiologic survey of dementia in a community-dwelling Brazilian population. **Alzheimer Dis Assoc Disord**, v. 16, n. 2, p. 103-8, 2002. ISSN 0893-0341.

HOWARD, R. A.; MATHESON, J. E. Influence diagrams. **Decision Analysis**, v. 2, n. 3, p. 127-143, 2005. ISSN 1545-8490.

HSIEH, T. C.; KUPERMAN, G. J.; JAGGI, T.; HOJNOWSKI-DIAZ, P.; FISKIO, J.; WILLIAMS, D. H.; BATES, D. W.; GANDHI, T. K. Characteristics and Consequences of Drug Allergy Alert Overrides in a Computerized Physician Order Entry System. **Journal of the American Medical Informatics Association**, v. 11, n. 6, p. 482-491, 2004.

HUNINK, M. G. M.; GLASZIOU, P. P.; SIEGEL, J. E.; WEEKS, J. C.; PLISKIN, J. S.; ELSTEIN, A. S.; WEINSTEIN, M. C. **Decision making in health and medicine: integrating evidence and values**. Cambridge University Press, 2001. ISBN 0521770297.

HYMAN, W. When medical devices talk to each other: the promise and challenges of interoperability. **Biomedical instrumentation & technology/Association for the Advancement of Medical Instrumentation**, p. 28, 2010. ISSN 0899-8205.

IBA, W.; LANGLEY, P. Induction of one-level decision trees. Proceedings of the Ninth International Conference on Machine Learning, 1992. p.233-240.

JACK, C. R.; ALBERT, M. S.; KNOPMAN, D. S.; MCKHANN, G. M.; SPERLING, R. A.; CARRILLO, M. C.; THIES, B.; PHELPS, C. H. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. **Alzheimer's & dementia : the journal of the Alzheimer's Association**, v. 7, n. 3, p. 257-262, 2011. ISSN 1552-5260.

JALBERT, J. J.; DAIELLO, L. A.; LAPANE, K. L. Dementia of the Alzheimer Type. **Epidemiologic reviews**, v. 30, n. 1, p. 15-34, 2008.

JAMUS, D. R.; MÄDER, M. J. The Rey Complex Figure and its importance a method of neuropsychological assessment. **Journal of Epilepsy and Clinical Neurophysiology**, v. 11, p. 193-198, 2005. ISSN 1676-2649.

JANSSEN, W. G. M.; BUSSMANN, H. B. J.; STAM, H. J. Determinants of the sit-to-stand movement: a review. **Physical Therapy**, v. 82, n. 9, p. 866-879, 2002. ISSN 0031-9023.

JENSEN, F. V.; NIELSEN, T. D. **Bayesian networks and decision graphs**. Springer Verlag, 2007. ISBN 0387682813.

JOHN, G. H.; LANGLEY, P. Estimating continuous distributions in Bayesian classifiers. Eleventh Conference on Uncertainty in Artificial Intelligence, 1995. San Mateo. Citeseer. p.338-345.

JOHN, R. I.; INNOCENT, P. R. Modeling uncertainty in clinical diagnosis using fuzzy logic. **Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on**, v. 35, n. 6, p. 1340-1350, 2005. ISSN 1083-4419.

JOHNSON, P. D.; TU, S.; BOOTH, N.; SUGDEN, B.; PURVES, I. N. Using scenarios in chronic disease management guidelines for primary care. **Proc AMIA Symp**, p. 389-393, 2000.

JONGSAWAT, N.; PREMCHAIWADI, W. A SMILE web-based interface for learning the causal structure and performing a diagnosis of a Bayesian network. Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on, 2009. IEEE. p.376-382.

JORM, A. F. The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): a review. **International Psychogeriatrics**, v. 16, n. 03, p. 275-293, 2004. ISSN 1041-6102.

JUNG, C. Y.; SWARD, K. A.; HAUG, P. J. Executing medical logic modules expressed in ArdenML using Drools. **J Am Med Inform Assoc**, v. 16, p. 16, Dec 16 2011. ISSN 1527-974X (Electronic).

KADIE, C. M.; HOVEL, D.; HORVITZ, E. MSBNx: A component-centric toolkit for modeling and inference with bayesian networks. **Microsoft Research, Richmond, WA, Technical Report MSR-TR-2001-67**, v. 28, 2001.

KALRA, D.; BEALE, T.; HEARD, S. The openEHR foundation. **Studies in health technology and informatics**, v. 115, p. 153-173, 2005. ISSN 0926-9630.

KALRA, D.; INGRAM, D. Electronic Health Records Information Technology Solutions for Healthcare. In: (Ed.): Springer London, 2006. p.135-181. (Health Informatics). ISBN 978-1-84628-141-9.

KAUSHAL R, S. K. G. B. D. W. Effects of computerized physician order entry and clinical decision support systems on medication safety: A systematic review. **Archives of Internal Medicine**, v. 163, n. 12, p. 1409-1416, 2003. ISSN 0003-9926.

KAWAMOTO, K.; HOULIHAN, C. A.; BALAS, E. A.; LOBACH, D. F. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. **BMJ**, v. 330, n. 7494, p. 765, 2005.

KERBER, R. Chimerge: Discretization of numeric attributes. Proceedings of the tenth national conference on Artificial intelligence, 1992. AAAI Press. p.123-128.

KHACHATURIAN, Z. S. Diagnosis of Alzheimer's disease. **Archives of neurology**, v. 42, n. 11, p. 1097, 1985.

KOHAVI, R. The power of decision tables. **Machine Learning: ECML-95**, p. 174-189, 1995a.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. International joint Conference on artificial intelligence, 1995b. Lawrence Erlbaum Associates Ltd. p.1137-1145.

KOHN, L. T.; CORRIGAN, J. M.; DONALDSON, M. S. **To Err is Human: Building a Safer Health System**. Washington: National Academy Press, 2000.

KONG, G.; XU, D.; YANG, J. Clinical decision support systems: a review on knowledge representation and inference under uncertainties. **International Journal of Computational Intelligence Systems**, v. 1, n. 2, p. 159–167, 2008.

KONONENKO, I. On biases in estimating multi-valued attributes. International Joint Conference on Artificial Intelligence, 1995. LAWRENCE ERLBAUM ASSOCIATES LTD. p.1034-1040.

KORB, K.; NICHOLSON, A. **Bayesian artificial intelligence**. Clayton, Victoria, Australia: Chapman & Hall/CRC, 2004. ISBN 9781584883876.

KOTSIANTIS, S.; KANELLOPOULOS, D. Discretization techniques: A recent survey. **GESTS International Transactions on Computer Science and Engineering**, v. 32, n. 1, p. 47-58, 2006.

KRAFZIG, D.; BANKE, K.; SLAMA, D. **Enterprise SOA: service-oriented architecture best practices**. Prentice Hall PTR, 2005. ISBN 0131465759.

KUBAT, M.; MATWIN, S. Addressing the curse of imbalanced training sets: one-sided selection. MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-, 1997. MORGAN KAUFMANN PUBLISHERS, INC. p.179-186.

KULLBACK, S.; LEIBLER, R. A. On information and sufficiency. **The Annals of Mathematical Statistics**, v. 22, n. 1, p. 79-86, 1951. ISSN 0003-4851.

LASKEY, K. B. MEBN: A language for first-order Bayesian knowledge bases. **Artificial intelligence**, v. 172, n. 2-3, p. 140-178, 2008. ISSN 0004-3702.

- LAURIA, E. J. M.; DUCHESSI, P. J. A methodology for developing Bayesian networks: An application to information technology (IT) implementation. **European Journal of operational research**, v. 179, n. 1, p. 234-252, 2007. ISSN 0377-2217.
- LAURITZEN, S. L.; SPIEGELHALTER, D. J. Local computations with probabilities on graphical structures and their application to expert systems. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 157-224, 1988. ISSN 0035-9246.
- LAWTON, M. P. Assessing quality of life in Alzheimer disease research. **Alzheimer disease and associated disorders**, 1997. ISSN 0893-0341.
- LE CESSIE, S.; VAN HOUWELINGEN, J. Ridge estimators in logistic regression. **Applied Statistics**, p. 191-201, 1992. ISSN 0035-9254.
- LINDGREN, H. **Decision support in dementia care: developing systems for interactive reasoning**. 2007. Umeå University, Faculty of Science and Technology, Computing Science
- LING, C. X.; LI, C. Data mining for direct marketing: Problems and solutions. Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining, 1998. p.73-79.
- LIU, H.; HUSSAIN, F.; TAN, C. L.; DASH, M. Discretization: An enabling technique. **Data Mining and Knowledge Discovery**, v. 6, n. 4, p. 393-423, 2002. ISSN 1384-5810.
- LIU, T. F.; SUNG, W. K.; MITTAL, A. Model gene network by semi-fixed Bayesian network. **Expert Systems with Applications**, v. 30, n. 1, p. 42-49, 2006. ISSN 0957-4174.
- LUCIANI, D.; CAVUTO, S.; ANTIGA, L.; MINIATI, M.; MONTI, S.; PISTOLESI, M.; BERTOLINI, G. Bayes pulmonary embolism assisted diagnosis: a new expert system for clinical use. **Emergency Medicine Journal**, v. 24, n. 3, p. 157-164, 2007.
- MACKAY, D. J. C. Information-based objective functions for active data selection. **Neural computation**, v. 4, n. 4, p. 590-604, 1992. ISSN 0899-7667.
- MANI, S.; MCDERMOTT, S.; VALTORTA, M. MENTOR: a Bayesian model for prediction of mental retardation in newborns. **Res. Dev. Disabil.**, v. 18, n. 5, p. 303-318, 1997.
- MARTÍNEZ-COSTA, C.; MENÁRGUEZ-TORTOSA, M.; FERNÁNDEZ-BREIS, J. T. An approach for the semantic interoperability of ISO EN 13606 and OpenEHR archetypes. **Journal of Biomedical Informatics**, v. 43, n. 5, p. 736-746, 2010. ISSN 1532-0464.
- MCDONALD, M.; HERTZ, R. P.; UNGER, A. N.; LUSTIK, M. B. Prevalence, awareness, and management of hypertension, dyslipidemia, and diabetes among United States adults aged 65 and older. **The Journals of Gerontology Series A: Biological Sciences and Medical Sciences**, v. 64, n. 2, p. 256, 2009. ISSN 1079-5006.
- MCKHANN, G.; DRACHMAN, D.; FOLSTEIN, M.; KATZMAN, R.; PRICE, D.; STADLAN, E. M. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. **Neurology**, v. 34, n. 7, p. 939-944, July 1, 1984 1984.

MCKHANN, G. M.; KNOPMAN, D. S.; CHERTKOW, H.; HYMAN, B. T.; JACK JR, C. R.; KAWAS, C. H.; KLUNK, W. E.; KOROSHETZ, W. J.; MANLY, J. J. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. **Alzheimer's and Dementia**, 2011. ISSN 1552-5260.

METAXIOTIS, K.; PTOCHOS, D.; PSARRAS, J. E-health in the new millennium: a research and practice agenda. **International Journal of Electronic Healthcare**, v. 1, n. 2, p. 165-175, 2004. ISSN 1741-8453.

MILLER, R. A. Computer-assisted diagnostic decision support: history, challenges, and possible paths forward. **Advances in health sciences education**, v. 14, p. 89-106, 2009. ISSN 1382-4996.

MILLER, R. A.; MCNEIL, M. A.; CHALLINOR, S. M.; MASARIE, F. E., JR.; MYERS, J. D. The INTERNIST-1/QUICK MEDICAL REFERENCE project--status report. **West J Med**, v. 145, n. 6, p. 816-22, 1986. ISSN 0093-0415.

MILLER, R. A.; POPLER JR, H. E.; MYERS, J. D. INTERNIST-I, an experimental computer-based diagnostic consultant for general internal medicine. **New England Journal of Medicine**, v. 307, n. 8, p. 468-476, 1982. ISSN 0028-4793.

MIRRA, S. S.; HEYMAN, A.; MCKEEL, D.; SUMI, S. M.; CRAIN, B. J.; BROWNLEE, L. M.; VOGEL, F. S.; HUGHES, J. P.; VAN BELLE, G.; BERG, L. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. **Neurology**, v. 41, n. 4, p. 479-86, Apr 1991. ISSN 0028-3878 (Print) 0028-3878 (Linking).

MONTANARI, A.; RIZZO, T. How to compute loop corrections to the Bethe approximation. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2005, n. 10, p. P10011, 2005. ISSN 1742-5468.

MONTAÑO, M.; RAMOS, L. R. Validade da versão em português da Clinical Dementia Rating. **Rev Saúde Pública**, v. 39, n. 6, p. 912-7, 2005.

MORRIS, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. **Neurology**, v. 43, n. 11, p. 2412-4, 1993. ISSN 0028-3878.

MORRIS, J. C. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. **International Psychogeriatrics**, v. 9, n. S1, p. 173-176, 1997. ISSN 1741-203X.

MORRIS, J. C.; HEYMAN, A.; MOHS, R. C.; HUGHES, J. P.; VAN BELLE, G.; FILLENBAUM, G.; MELLITS, E. D.; CLARK, C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. **Neurology**, v. 39, n. 9, p. 1159-65, 1989. ISSN 0028-3878.

MORRISON, I.; LEWIS, B.; NUGRAHANTO, S. Modelling in clinical practice with Web Services and BPEL. **International Journal of E-Business Research**, v. 2, n. 1, p. 45-57, 2006. ISSN 1548-1131.

MURPHY, K. The bayes net toolbox for matlab. **Computing science and statistics**, v. 33, n. 2, p. 1024-1034, 2001.

MURPHY, K.; MYORS, B.; WOLACH, A. **Statistical power analysis: A simple and general model for traditional and modern hypothesis tests**. Routledge Academic, 2012. ISBN 0415965551.

MUSEN, M.; TU, S. W.; DAS, A. K.; SHAHAR, Y. EON: a component-based approach to automation of protocol directed therapy. **J Am Med Inform Assoc**, v. 3, n. 6, p. 367-388, 1996.

NAMLI, T.; ALUC, G.; DOGAC, A. An interoperability test framework for HL7-based systems. **Information Technology in Biomedicine, IEEE Transactions on**, v. 13, n. 3, p. 389-399, 2009. ISSN 1089-7771.

NEAPOLITAN, R. E. **Learning bayesian networks**. Pearson Prentice Hall Upper Saddle River, NJ, 2004. ISBN 0130125342.

NEASE, R. F.; OWENS, D. K. Use of influence diagrams to structure medical decisions. **Medical Decision Making**, v. 17, n. 3, p. 263-275, 1997. ISSN 0272-989X.

NEWMAN-TOKER, D. E.; PRONOVOST, P. J. Diagnostic errors: the next frontier for patient safety. **JAMA: the journal of the American Medical Association**, v. 301, n. 10, p. 1060-1062, 2009. ISSN 0098-7484.

NITRINI, R.; BOTTINO, C. M.; ALBALA, C.; CUSTODIO CAPUNAY, N. S.; KETZOIAN, C.; LLIBRE RODRIGUEZ, J. J.; MAESTRE, G. E.; RAMOS-CERQUEIRA, A. T.; CARAMELLI, P. Prevalence of dementia in Latin America: a collaborative study of population-based cohorts. **Int Psychogeriatr**, v. 21, n. 4, p. 622-30, Aug 2009. ISSN 1041-6102.

NITRINI, R.; CARAMELLI, P.; HERRERA, E., JR.; BAHIA, V. S.; CAIXETA, L. F.; RADANOVIC, M.; ANGHINAH, R.; CHARCHAT-FICHMAN, H.; PORTO, C. S.; CARTHERY, M. T.; HARTMANN, A. P.; HUANG, N.; SMID, J.; LIMA, E. P.; TAKADA, L. T.; TAKAHASHI, D. Y. Incidence of dementia in a community-dwelling Brazilian population. **Alzheimer Dis Assoc Disord**, v. 18, n. 4, p. 241-6, Oct-Dec 2004. ISSN 0893-0341.

OLAFSDOTTIR, M.; MARCUSSON, J.; SKOOG, I. Mental disorders among elderly people in primary care: the Linkoping study. **Acta Psychiatr Scand**, v. 104, n. 1, p. 12-8, Jul 2001. ISSN 0001-690X.

OLSON, D. L.; DELEN, D. **Advanced data mining techniques**. Springer Verlag, 2008. ISBN 3540769161.

OSHEROFF, J. A.; TEICH, J. M.; MIDDLETON, B.; STEEN, E. B.; WRIGHT, A.; DETMER, D. E. A roadmap for national action on clinical decision support. **J Am Med Inform Assoc**, v. 14, n. 2, p. 141-5, 2007. ISSN 1067-5027.

PAAJANEN, T.; HÄNNINEN, T.; TUNNARD, C.; MECOCCHI, P.; SOBOW, T.; TSOLAKI, M.; VELLAS, B.; LOVESTONE, S.; SOININEN, H.; FOR THE ADDNEUROMED, C. CERAD Neuropsychological Battery Total Score in Multinational Mild Cognitive Impairment and Control Populations: The AddNeuroMed Study. **Journal of Alzheimer's Disease**, v. 22, n. 4, p. 1089-1097, 2010.

PASSOS, V. M. D. A.; GIATTI, L.; BARRETO, S. M.; FIGUEIREDO, R. C.; CAMELLI, P.; BENSEÑOR, I.; JESUS MENDES DA FONSECA, M. D.; CADE, N. V.; GOULART, A. C.; NUNES, M. A.; ALVES, M. G. D. M.; TRINDADE, A. A. M. D. Verbal fluency tests reliability in a Brazilian multicentric study, ELSA-Brasil. **Arquivos de Neuro-Psiquiatria**, v. 69, p. 814-816, 2011. ISSN 0004-282X.

PEARL, J. **Probabilistic reasoning in intelligent systems: networks of plausible inference**. Morgan Kaufmann, 1988. ISBN 1558604790.

PELEG, M.; BOXWALA, A. A.; OGUNYEMI, O.; ZENG, Q.; TU, S.; LACSON, R.; BERNSTAM, E.; ASH, N.; MORK, P.; OHNO-MACHADO, L.; SHORTLIFFE, E. H.; GREENES, R. A. GLIF3: the evolution of a guideline representation format. **Proc AMIA Symp**, p. 645-9, 2000. ISSN 1531-605X.

PELEG, M.; TU, S.; BURY, J.; CICCARESE, P.; FOX, J.; GREENES, R. A.; HALL, R.; JOHNSON, P. D.; JONES, N.; SEYFANG, A.; SHORTLIFFE, E. H.; STEFANELLI, M. Comparing computer-interpretable guideline models: a case-study approach. **Journal of the American Medical Informatics Association**, v. 10, n. 1, p. 52-68, 2003.

PETERSEN, R. C. **Mild cognitive impairment: aging to Alzheimer's disease**. Oxford University Press, USA, 2003. ISBN 0195123425.

PFEFFER, R. I.; KUROSAKI, T. T.; HARRAH, C. H.; CHANCE, J. M.; FILOS, S. Measurement of Functional Activities in Older Adults in the Community. **Journal of Gerontology**, v. 37, n. 3, p. 323-329, May 1, 1982 1982.

PHILLIPS, R. L.; BARTHOLOMEW, L. A.; DOVEY, S. M.; FRYER, G.; MIYOSHI, T.; GREEN, L. Learning from malpractice claims about negligent, adverse events in primary care in the United States. **Quality and Safety in Health Care**, v. 13, n. 2, p. 121-126, 2004. ISSN 2044-5423.

PINHEIRO, P. R.; CASTRO, A.; PINHEIRO, M. A Multicriteria Model Applied in the Diagnosis of Alzheimer's Disease: A Bayesian Network. Computational Science and Engineering, 2008. CSE '08. 11th IEEE International Conference on, 2008. 16-18 July 2008. p.15-22.

PODSIADLO, D.; RICHARDSON, S. The timed" Up & Go": a test of basic functional mobility for frail elderly persons. **Journal of the American Geriatrics Society**, v. 39, n. 2, p. 142, 1991. ISSN 0002-8614.

PRADHAN, M.; PROVAN, G.; MIDDLETON, B.; HENRION, M. Knowledge engineering for large belief networks. *Uncertainty in Artificial Intelligence*, 1994. San Francisco. Proceedings of the Seventh Conference. p.484–490.

PYPER, C.; FRIZE, M.; LINDGAARD, G. Bayesian-based diagnostic decision-support for pediatrics. *Engineering in Medicine and Biology Society*, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, 2008. p.4318-4321.

QUINLAN, J. R. **C4. 5: programs for machine learning**. Morgan Kaufmann, 1993. ISBN 1558602380.

QUINLAN, J. R. Learning decision tree classifiers. **ACM Computing Surveys (CSUR)**, v. 28, n. 1, p. 71-72, 1996. ISSN 0360-0300.

RAMONI, M.; SEBASTIANI, P. Robust learning with missing data. **Machine learning**, v. 45, n. 2, p. 147-170, 2001. ISSN 0885-6125.

RISSANEN, J. Modeling by shortest data description. **Automatica**, v. 14, n. 5, p. 465-471, 1978. ISSN 0005-1098.

RODRIGUES, A. B.; YAMASHITA, É. T.; CHIAPPETTA, A. L. D. M. L. Teste de fluência verbal no adulto e no idoso: verificação da aprendizagem verbal. **Revista CEFAC**, v. 10, p. 443-451, 2008. ISSN 1516-1846.

RODRIGUEZ, J. J. L.; FERRI, C. P.; ACOSTA, D.; GUERRA, M.; HUANG, Y.; JACOB, K. S.; KRISHNAMOORTHY, E. S.; SALAS, A.; SOSA, A. L.; ACOSTA, I.; DEWEY, M. E.; GAONA, C.; JOTHEESWARAN, A. T.; LI, S.; RODRIGUEZ, D.; RODRIGUEZ, G.; KUMAR, P. S.; VALHUERDI, A.; PRINCE, M. Prevalence of dementia in Latin America, India, and China: a population-based cross-sectional survey. **The Lancet**, v. 372, n. 9637, p. 464-474, 2008. ISSN 0140-6736.

ROJO, M. G.; ROLON, E.; CALAHORRA, L.; GARCIA, F. O.; SANCHEZ, R. P.; RUIZ, F.; BALLESTER, N.; ARMENTEROS, M.; RODRIGUEZ, T.; ESPARTERO, R. M. Implementation of the Business Process Modelling Notation (BPMN) in the modelling of anatomic pathology processes. **Diagn Pathol**, v. 3 Suppl 1, n. 3, p. S22, 2008. ISSN 1746-1596.

ROSSETTI, H. C.; CULLUM, C. M.; HYNAN, L. S.; LACRITZ, L. The CERAD Neuropsychological Battery Total Score and the Progression of Alzheimer's Disease. **Alzheimer disease and associated disorders**, v. 24, n. 2, p. 138, 2010.

RUBIN, D. B. Inference and missing data. **Biometrika**, v. 63, n. 3, p. 581-592, 1976. ISSN 0006-3444.

RUSSELL, S.; NORVIG, P. **Artificial intelligence: a modern approach**. Prentice hall, 2009.

SAAR-TSECHANSKY, M.; PROVOST, F. Handling missing values when applying classification models. **Journal of machine learning research**, 2007.

SALAS-GONZALEZ, D.; GÓRRIZ, J. M.; RAMÍREZ, J.; LÓPEZ, M.; ÁLVAREZ, I.; SEGOVIA, F.; CHAVES, R.; PUNTONET, C. G. Computer-aided diagnosis of Alzheimer's disease using support vector machines and classification trees. **Physics in Medicine and Biology**, v. 55, p. 2807-2817, 2010.

SANCHEZ, M. A. S.; LOURENÇO, R. A. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): cross-cultural adaptation for use in Brazil. **Cadernos de Saúde Pública**, v. 25, n. 7, p. 1455-1465, 2009. ISSN 0102-311X.

SANTOS, M. R.; BAX, M. P.; KALRA, D. Building a logical EHR architecture based on ISO 13606 standard and semantic web technologies. **Studies in health technology and informatics**, v. 160, n. Pt 1, p. 161, 2010. ISSN 0926-9630.

SCHIFF, G. D.; HASAN, O.; KIM, S.; ABRAMS, R.; COSBY, K.; LAMBERT, B. L.; ELSTEIN, A. S.; HASLER, S.; KABONGO, M. L.; KROSNJAR, N. Diagnostic error in medicine: analysis of 583 physician-reported errors. **Archives of Internal Medicine**, v. 169, n. 20, p. 1881, 2009.

SCHMIDT, H.; NORMAN, G.; BOSCHUIZEN, H. A cognitive perspective on medical expertise: theory and implication [published erratum appears in Acad Med 1992 Apr; 67 (4): 287. **Academic medicine**, v. 65, n. 10, p. 611, 1990. ISSN 1040-2446.

SEIFFERT, C.; KHOSHGOFTAAR, T. M.; VAN HULSE, J.; NAPOLITANO, A. Resampling or Reweighting: A Comparison of Boosting Implementations. Tools with Artificial Intelligence, 2008. ICTAI'08. 20th IEEE International Conference on, 2008. IEEE. p.445-451.

SEIXAS, F. L.; ZADROZNY, B.; LAKS, J.; CONCI, A.; MUCHALUAT-SAADE, D. C. A decision support system for Alzheimer's disease diagnosis. In: RAÚL FEIJÓO, A. Z., PABLO J. BLANCO (Ed.). **Scientific Computing Applied to Medicine and Healthcare: current state and future trends at the INCT-MACC the Brazilian National Institute of Science and Technology in Medicine Assisted by Scientific Computing**. São Paulo: INCT-MACC/LNCC, 2012. p.355-372.

SEIXAS, F. L.; ZADROZNY, B.; LAKS, J.; SAADE, D. C. M.; CONCI, A. Sistema de Apoio à Decisão Aplicado ao Diagnóstico da Doença de Alzheimer. XI Workshop de Informática Médica. XXXI Congresso da Sociedade Brasileira de Computação. Disponível em: < http://www.dimap.ufrn.br/csbc2011/anais/eventos/contents/WIM/WIM_Sessao_3_Artigo_7_Seixas.pdf>. Acessado em Dezembro de, 2011.

SHAH, R.; KESAN, J. P. An empirical examination of open standards development. 2008. IEEE. p.212-212.

SHAHAR, Y.; MIKSCH, S.; JOHNSON, P. The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. **Artificial Intelligence in Medicine**, v. 14, n. 1, p. 29-51, 1998. ISSN 0933-3657.

SHWE, M. A.; MIDDLETON, B.; HECKERMAN, D.; HENRION, M.; HORVITZ, E.; LEHMANN, H.; COOPER, G. Probabilistic diagnosis using a reformulation of the

INTERNIST-1/QMR knowledge base. **Methods of information in medicine**, v. 30, n. 4, p. 241-255, 1991.

SICILIA, J. J.; SICILIA, M. A.; SANCHEZ-ALONSO, S.; GARCIA-BARRIOCANAL, E. Knowledge representation issues in ontology-based clinical Knowledge Management systems. **International Journal of Technology Management**, v. 47, n. 1, p. 191-206, 2009. ISSN 0267-5730.

SIM, I.; GORMAN, P.; GREENES, R. A.; HAYNES, R. B.; KAPLAN, B.; LEHMANN, H.; TANG, P. C. Clinical decision support systems for the practice of evidence-based medicine. **J Am Med Inform Assoc**, v. 8, n. 6, p. 527-34, Nov-Dec 2001. ISSN 1067-5027.

SINGH, H.; GIARDINA, T. D.; PETERSEN, L. A.; SMITH, M. W.; PAUL, L. W.; DISMUKES, K.; BHAGWATH, G.; THOMAS, E. J. Exploring situational awareness in diagnostic errors in primary care. **BMJ Quality & Safety**, v. 21, n. 1, p. 30-38, 2012. ISSN 2044-5423.

SINGH, M.; VALTORTA, M. Construction of Bayesian network structures from data: a brief survey and an efficient algorithm. **International journal of approximate reasoning**, v. 12, n. 2, p. 111-131, 1995. ISSN 0888-613X.

SOKOLOWSKI, R.; DUDECK, J. XML and its impact on content and structure in electronic health care documents. 1999. American Medical Informatics Association. p.147.

SOTANIEMI, M.; PULLIAINEN, V.; HOKKANEN, L.; PIRTTILÄ, T.; HALLIKAINEN, I.; SOININEN, H.; HÄNNINEN, T. CERAD-neuropsychological battery in screening mild Alzheimer's disease. **Acta Neurologica Scandinavica**, v. 125, n. 1, p. 16-23, 2012. ISSN 1600-0404.

SPIEGELMAN, D.; MCDERMOTT, A.; ROSNER, B. Regression calibration method for correcting measurement-error bias in nutritional epidemiology. **The American journal of clinical nutrition**, v. 65, n. 4, p. 1179S-1186S, 1997. ISSN 0002-9165.

SPIRITES, P.; GLYMOUR, C. N.; SCHEINES, R. **Causation, prediction, and search**. The MIT Press, 2000. ISBN 0262194406.

SPREEN, O.; STRAUSS, E. **A compendium of neuropsychological tests: Administration, norms, and commentary**. Oxford University Press, USA, 1998. ISBN 0195100190.

STECK, H.; JAAKKOLA, T. Predictive discretization during model selection. **Pattern Recognition**, p. 1-8, 2004.

STROOP, J. R. Studies of interference in serial verbal reactions. **Journal of Experimental Psychology: General**, v. 121, n. 1, p. 15, 1992. ISSN 1939-2222.

SUTTON, D. R.; FOX, J. The syntax and semantics of the PROforma guideline modeling language. **J Am Med Inform Assoc**, v. 10, n. 5, p. 433-43, Sep-Oct 2003. ISSN 1067-5027.

TEICH, J. M.; OSHEROFF, J. A.; PIFER, E. A.; SITTING, D. F.; JENDERS, R. A. Clinical decision support in electronic prescribing: recommendations and an action plan: report of the

joint clinical decision support workgroup. **Journal of the American Medical Informatics Association**, v. 12, n. 4, p. 365-376, 2005. ISSN 1067-5027.

TEJERINA, E.; ESTEBAN, A.; FERNÁNDEZ-SEGOVIANO, P.; MARÍA RODRÍGUEZ-BARBERO, J.; GORDO, F.; FRUTOS-VIVAR, F.; ARAMBURU, J.; ALGABA, Á.; GONZALO SALCEDO GARCÍA, Ó.; LORENTE, J. A. Clinical diagnoses and autopsy findings: Discrepancies in critically ill patients*. **Critical care medicine**, v. 40, n. 3, p. 842, 2012. ISSN 0090-3493.

TRIPOLITI, E. E.; FOTIADIS, D. I.; ARGYROPOULOU, M. A supervised method to assist the diagnosis and monitor progression of Alzheimer's disease using data from an fMRI experiment. **Artif Intell Med**, v. 53, n. 1, p. 35-45, Sep 2011. ISSN 1873-2860 (Electronic).

TVERSKY, A.; KAHNEMAN, D. Judgment under uncertainty: Heuristics and biases. **science**, v. 185, n. 4157, p. 1124-1131, 1974. ISSN 0036-8075.

VEST, J. R. More than just a question of technology: Factors related to hospitals' adoption and implementation of health information exchange. **International Journal of Medical Informatics**, v. 79, n. 12, p. 797-806, 2010. ISSN 1386-5056.

VISWESWARAN, S. **Learning Patient-Specific Models from Clinical Data**. 2007. (PhD). Department of Biomedical Informatics and Intelligent Systems Studies, University of Pittsburgh, Pittsburgh.

WANG, D. Translating Arden MIMs into GLIF guidelines--a case study of hyperkalemia patient screening. **Stud Health Technol Inform**, v. 101, p. 177-81, 2004. ISSN 0926-9630.

WANG, D.; SHORTLIFFE, E. H. GLEE: a model-driven execution system for computer-based implementation of clinical practice guidelines. **Proc AMIA Symp**, p. 855-9, 2002. ISSN 1531-605X.

WANG, H. **Building Bayesian Networks: Elicitation, Evaluation, and Learning**. 2007. University of Pittsburgh

WANG, J. K.; SHABOT, M. M.; DUNCAN, R. G.; POLASCHEK, J. X.; JONES, D. T. A clinical rules taxonomy for the implementation of a computerized physician order entry (CPOE) system. 2002. American Medical Informatics Association. p.860.

WEISS, S. M.; KULIKOWSKI, C. A.; SAFIR, A. A model-based consultation system for the long-term management of glaucoma. **Specialized Systems**, v. 3, p. 826-832, 1976.

WELSH, K.; BUTTERS, N.; HUGHES, J.; MOHS, R.; HEYMAN, A. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. **Archives of neurology**, v. 48, n. 3, p. 278, 1991.

WEMMENHOVE, B.; MOOJI, J.; WIEGERINCK, W.; LEISINK, M.; HAPPEN, H.; NEIJT, J. Inference in the Promedas medical expert system. **Artificial Intelligence in Medicine**, p. 456-460, 2007.

WENK, G. L. Neuropathologic changes in Alzheimer's disease. **J Clin Psychiatry**, v. 64, n. 9, p. 7-10, 2003.

WERNER, P.; STEIN-SHVACHMAN, I.; KORCZYN, A. D. Early onset dementia: clinical and social aspects. **International Psychogeriatrics**, v. 21, n. 4, p. 631, 2009. ISSN 1041-6102.

WHITE, S. A.; MIERS, D. **BPMN modeling and reference guide**. Future Strategies Inc., 2008. ISBN 0977752720.

WILLIAMS, B. W.; MACK, W.; HENDERSON, V. W. Boston naming test in Alzheimer's disease. **Neuropsychologia**, v. 27, n. 8, p. 1073-1079, 1989. ISSN 0028-3932.

WINKLER, R. L. The assessment of prior distributions in Bayesian analysis. **Journal of the American Statistical Association**, p. 776-800, 1967. ISSN 0162-1459.

WITTEN, I. H.; FRANK, E. **Data Mining: Practical machine learning tools and techniques**. Morgan Kaufmann, 2005. ISBN 0120884070.

WRIGHT, A.; SITTING, D. F.; ASH, J. S.; FEBLOWITZ, J.; MELTZER, S.; MCMULLEN, C.; GUAPPONE, K.; CARPENTER, J.; RICHARDSON, J.; SIMONAITIS, L.; EVANS, R. S.; NICHOL, W. P.; MIDDLETON, B. Development and evaluation of a comprehensive clinical decision support taxonomy: comparison of front-end tools in commercial and internally developed electronic health record systems. **J Am Med Inform Assoc**, v. 18, n. 3, p. 232-42, May 1 2011. ISSN 1527-974X (Electronic).

YAN, S.; SHIPIN, L.; YIYUAN, T. Construction and Application of Bayesian Network in Early Diagnosis of Alzheimer Disease's System. *Complex Medical Engineering*, 2007. CME 2007. IEEE/ICME International Conference on, 2007. 23-27 May 2007. p.924-929.

YAO, W.; KUMAR, A. Integrating clinical pathways into CDSS using context and rules: a case study in heart disease. 2012. ACM. p.611-620.

ZUSMAN, E. E. Form facilitates function: innovations in architecture and design drive quality and efficiency in healthcare. **Neurosurgery**, v. 66, n. 6, p. N24, 2010. ISSN 0148-396X.