

## Sistema preditivo para a doença de Alzheimer na triagem clínica

System predictive for Alzheimer's disease in clinical trial

Predictivo Sistema para la enfermedad de Alzheimer en ensayo clínico

Leonard Barreto Moreira<sup>1</sup>, Anderson Amendoeira Namen<sup>2</sup>

### RESUMO

**Descritores:** Mineração de Dados; Técnicas de Apoio para a Decisão; Doença de Alzheimer

**Objetivo:** Descrever uma aplicação que, utilizando-se de técnicas de mineração de dados, visa auxiliar os especialistas no processo de diagnóstico de pacientes com suspeita clínica de Alzheimer atendidos pelo Centro de Alzheimer e Parkinson no município de Campos dos Goytacazes/RJ. **Método:** Aplicação de técnicas relacionadas à etapa de pré-processamento dos dados, de classificação (naïve bayes, redes bayesianas e árvores de decisão) com avaliação dos resultados a partir do uso da validação cruzada estratificada *k-fold*, cujas implementações estão disponíveis na ferramenta Weka. **Resultado:** Observou-se os resultados numéricos dos modelos de acordo com as métricas: acurácia, taxa de erro, sensibilidade, taxa de falsos positivos e taxa de falsos negativos, obtendo-se as taxas de 73,8%, 26,2%, 76,3%, 27,4%, 23,7%, respectivamente. **Conclusão:** Verificou-se que os classificadores bayesianos, em especial redes bayesianas, apresentaram os melhores resultados para o diagnóstico da doença de Alzheimer a partir das métricas supracitadas.

### ABSTRACT

**Keywords:** Data Mining; Decision Support Techniques; Alzheimer Disease

**Objective:** Describe an application that uses data mining techniques and aims to support those specialists in the Alzheimer diagnostic process which assist patients in the Alzheimer's and Parkinson's Center, located in the city of Campos dos Goytacazes, Brazil. **Methods:** Application of methods related with data pre-processing, classification (naïve bayes, bayesian networks and decision trees) using k-fold method for results evaluation, whose implementations are available in the Weka tool. **Results:** The obtained results related to the accuracy of the models were: accuracy - 73.8%; error rate - 26.2%; sensitivity - 76.3%; false positive rate - 27.4%; and false negative rate - 23.7%. **Conclusion:** The Bayesian classifiers, more specifically Bayesian networks, presented best results for the diagnosis of Alzheimer.

### RESUMEN

**Descriptores:** Minería de Datos; Técnicas de Apoyo para la Decisión; Enfermedad de Alzheimer

**Objetivo:** Describir una aplicación que utiliza técnicas de minería de datos, con objetivo de ayudar a los expertos en el diagnóstico de pacientes con sospecha de enfermedad de Alzheimer, que son atendidos en el Centro de Alzheimer y Parkinson en el municipio de Campos dos Goytacazes, Brasil. **Método:** Aplicación de técnicas para el procesamiento previo de los datos, la clasificación (naïve Bayes, redes bayesianas y rboles de decisión), aplicando el método k-fold en la evaluación de los resultados, cuyas implementaciones están disponibles en la herramienta Weka. **Resultado:** Los resultados obtenidos en relación con la exactitud de los modelos fueron: precisión - 73,8%; tasa de error - 26,2%; Sensibilidad - 76,3%; tasa de falsos positivos - 27,4%; y tasa de falsos negativos - 23,7%. **Conclusión:** Los clasificadores bayesianos, más concretamente redes bayesianas, presentaron mejores resultados para el diagnóstico de la enfermedad de Alzheimer.

<sup>1</sup> Doutorando do curso de Modelagem Computacional, Universidade Estadual do Rio de Janeiro - UERJ, Nova Friburgo (RJ), Brasil.

<sup>2</sup> Professor Adjunto do Departamento de Modelagem Computacional, Universidade Estadual do Rio de Janeiro - UERJ, Nova Friburgo (RJ), e Professor da Universidade Veiga de Almeida (RJ), Brasil.

## INTRODUÇÃO

O envelhecimento populacional é um fenômeno mundial, ocorrendo em diferentes ritmos nas diversas regiões do mundo. A partir da década de 1970 a expectativa de vida mundial aumentou em média 20 anos. No Brasil, neste mesmo período, esse aumento foi de 30 anos<sup>(1)</sup>. Segundo dados do Fundo de População das Nações Unidas, pessoas com 60 anos ou mais representavam 11,5% da população mundial no ano de 2012. Estima-se que em 2050 este percentual atinja 21,8%, valor ligeiramente menor do que os 22,5% previstos para o Brasil neste mesmo ano<sup>(1-2)</sup>.

O crescimento da população idosa e da expectativa de vida implica no aumento de problemas vinculados aos declínios no funcionamento fisiológico. Mais especificamente, no âmbito da saúde mental, o aumento da população nesta faixa etária tem como efeito a ampliação da proporção de adultos idosos com algum tipo de demência, com previsões de que o número de casos duplique a cada 20 anos. Considera-se esta uma das principais causas de incapacidade durante o envelhecimento. Dentre os diversos tipos, a Doença de Alzheimer (DA) é a mais prevalente das demências, respondendo por 60% a 70% dos casos<sup>(3)</sup>.

O custo estimado das demências no mundo representou algo em torno de 604 bilhões de dólares em 2010, ou seja, 1% do Produto Interno Bruto global. Nos Estados Unidos, por exemplo, os custos totais relacionados à DA somente são superados pelos relacionados ao câncer e às doenças coronarianas<sup>(3-4)</sup>.

Diante deste cenário, o envelhecimento foi reconhecido não apenas como uma questão de segurança social e de bem-estar, mas de desenvolvimento global e de política econômica<sup>(2)</sup>.

É crescente o avanço de pesquisas direcionadas ao diagnóstico de DA. Sabe-se que o processo de diagnóstico médico não é uma tarefa fácil, dada a diversidade de doenças, sintomas, exames, além da complexidade da fisiologia humana. Especificamente em relação ao diagnóstico de síndromes demenciais, tal dificuldade é ainda maior devido à inexistência de um exame específico para determinar o tipo de demência, demandando assim uma extensa investigação clínica através de uma série de consultas, avaliações e exames<sup>(5)</sup>. Dessa forma, a utilização de mecanismos que minimizem a complexidade do processo de diagnóstico é benéfica tanto para os especialistas do domínio quanto para a sociedade em geral<sup>(6)</sup>.

A criação de sistemas computacionais inteligentes na área de saúde tem sido um dos desafios tecnológicos mais ambiciosos de nossa era. Dentre as diversas aplicações desenvolvidas neste recente período, pode-se citar os sistemas de suporte à decisão clínica (*Clinical Decision Support Systems* - CDSS), que, em sua forma mais simples, visam auxiliar especialistas no processo de investigação médica, possibilitando assim diagnósticos mais precisos<sup>(7)</sup>.

Diversos são os métodos e modelos matemáticos aplicados ao diagnóstico médico em CDSSs e o uso desses geralmente depende de sua finalidade, o que impacta na

escolha de uma metodologia, ou de um conjunto de métodos<sup>(8)</sup>. A mineração de dados e de texto têm se apresentado como métodos promissores dentro das atividades de previsão diagnóstica, extraindo ou ajudando a evidenciar padrões e auxiliando na descoberta de conhecimento em um conjunto de dados<sup>(9)</sup>. O estudo conduzido por Mani et al<sup>(10)</sup>, por exemplo, consistiu na investigação de pacientes com suspeita de DA segundo critérios do CERAD. Para tal, os dados demográficos e de resultados de testes neuropsicológicos de 428 pacientes foram analisados e submetidos às técnicas de aprendizagem supervisionada, como indutores de árvores de decisão (C4.5 e CART) e regras (C4.5Rules), além do classificador Naive Bayes. Nesse trabalho, foram obtidas taxas de acurácia e de sensibilidade de 73,1% e 68,8%, respectivamente, considerando-se o classificador Naive Bayes, que obteve a maior precisão entre os modelos. O trabalho realizado por Seixas<sup>(11)</sup>, por sua vez, foi fundamentado na aplicação da técnica de redes bayesianas na investigação de indivíduos com suspeita de transtorno cognitivo leve e DA. Utilizando a base do CERAD, que contém informações relacionadas a dados demográficos e resultados de testes de 1094 pacientes com suspeita de DA, as taxas obtidas para predição da referida doença em relação as métricas Medida F, sensibilidade e taxa de verdadeiros negativos foram de 65,0%, 66,0% e 57,0%, respectivamente. Diversos outros trabalhos, também aplicam a mineração de dados para o diagnóstico de demências<sup>(12-14)</sup>; contudo, sua abordagem envolve a análise de imagens e de dados temporais, não tratados no presente trabalho.

Diante deste cenário, o presente trabalho propõe o uso da mineração de dados para auxiliar os especialistas no processo de diagnóstico de pacientes com suspeita clínica de DA na triagem médica.

## MÉTODOS

Trata-se de uma pesquisa descritiva quantitativa, que utilizou dados provenientes de informações contidas nos prontuários de 3.281 pacientes atendidos pelo Centro de Alzheimer e Parkinson (CDAP) em Campos dos Goytacazes/RJ, coletadas a partir da aplicação desenvolvida. As informações dos pacientes consistem em dados demográficos, clínicos e laboratoriais. Os dados relativos à identificação do paciente foram desconsiderados no trabalho, permitindo total anonimato.

O desenvolvimento do estudo foi realizado nas seguintes etapas: (i) coleta e integração dos dados, (ii) seleção dos dados, (iii) pré-processamento e transformação dos dados, (iv) construção dos modelos e (v) avaliação da performance.

Na etapa (i) foi realizada a coleta e integração dos dados contidos nos prontuários (em papel) com informações parciais dos pacientes existentes em uma planilha eletrônica. Além da importação dos dados das planilhas diretamente para um banco de dados Postgres<sup>(15)</sup>, foi feita a inserção da maioria dos dados existentes nos prontuários através dos formulários existentes em um sistema de informação integrado ao banco de dados

citado. A Figura 1 apresenta a tela parcial referente ao perfil do paciente.

Os dados coletados e integrados na primeira etapa foram analisados, sendo adotado como critério de seleção para o processo de mineração o quantitativo de pacientes diagnosticados com algum tipo de demência. Deste, a doença de Alzheimer correspondeu a 46,12 % de toda a amostra. Portanto, o conjunto final de dados que foram selecionados na segunda etapa é composto por 1.483 registros, cada um contendo 13 atributos relacionados à triagem médica.

Na etapa (iii), pré-processamento e transformação dos dados, foram realizadas ações para corrigir imperfeições, tais como: valores faltando, dados inconsistentes e duplicados. Dentre estas ações citam-se o preenchimento de valores faltantes, correções de informações errôneas ou inconsistentes e a remoção de registros duplicados. Além disso, os diversos atributos pertencentes à base de dados foram discretizados, sendo adotadas as escalas e pontos de corte indicados na literatura médica e por especialistas do domínio. Posteriormente, para a predição de DA, o atributo diagnóstico foi transformado para binário, onde 1 (um) indica que o paciente é acometido pela doença e 0 (zero) caso contrário. A Tabela 1 apresenta informações detalhadas sobre os atributos selecionados para etapa (iv).

Para a construção dos modelos preditivos, etapa (iv), primeiramente foi necessária a designação do atributo alvo (ou classe), sendo definido o atributo **diagnostico**.

A seguir, os dados foram particionados em  $k$  partições (*fold*s) de mesmo tamanho mutuamente exclusivas, onde um subconjunto foi escolhido para teste e os  $k-1$  restantes para treinamento do modelo. Cabe ressaltar que essa divisão foi realizada de forma estratificada, assegurando-se que cada uma das classes fosse representada com proporções aproximadamente iguais em ambos os subconjuntos. Uma das grandes características desse método de avaliação, denominado de validação cruzada estratificada *k-fold*, é que todos os dados são utilizados tanto para teste quanto para treinamento. Apesar do esquema *10-fold* ser considerado na literatura o método de avaliação padrão<sup>(16)</sup>, mais oito experimentos foram conduzidos, com  $k$  variando de 2 a 9, de forma a identificar o arranjo que produzia os melhores resultados, no caso do presente estudo,  $k = 6$ .

Foram aplicados alguns classificadores disponíveis na ferramenta Weka (*Waikato Environment for Knowledge Analysis*): Naive Bayes, Redes Bayesianas e Árvores de Decisão. Esta escolha justifica-se pela crescente adoção destas técnicas, observada por intermédio de realização de pesquisa ampla sobre trabalhos relacionados ao auxílio ao diagnóstico médico<sup>(17)</sup>.

A escolha dos parâmetros para a geração dos modelos preditivos foi resultado da análise, efetuada por um médico especialista, dos modelos gerados a partir de diferentes parâmetros, pelas técnicas de árvores de decisão e redes bayesianas. Especificamente para a técnica *Naive Bayes*, os parâmetros foram obtidos em função da acurácia

Co-morbidades	Presente	Ausente	Não informado
Insuficiência cardíaca grave	✓		
Arritmia cardíaca não controlada	✓		
Síndrome Parkinsoniana	✓		
Doença péptica sem resposta ao tratamento	✓		
Diarréia crônica ou persistente	✓		

Figura 1 - Tela (parcial) do módulo de apresentação dos dados relativos ao perfil do paciente da aplicação desenvolvida

dos modelos gerados, devido à dificuldade de interpretação dos modelos gráficos por parte do especialista do domínio.

Finalmente, após a construção dos modelos, foi necessário avaliar a performance dos classificadores - etapa (v). As métricas utilizadas foram: acurácia, taxa de erro, sensibilidade ou taxa de verdadeiros positivos, taxa de falsos positivos (TFP), taxa de falsos negativos (TFN) e a medida F, cujas definições formais são apresentadas nas equações de 1 a 7.

$$\text{acurácia} = \frac{(VP+VN)}{(P+N)} \quad (1)$$

$$\text{taxa de erro} = \frac{FP + FN}{P + N} \quad (2)$$

$$\text{sensibilidade} = \text{lembrança} = \frac{(VP + VN)}{(P + N)} \quad (3)$$

$$\text{TFP} = \frac{FP}{(FP + VN)} \quad (4)$$

$$\text{TFN} = 1 - \text{sensibilidade} \quad (5)$$

$$\text{precisão} = \frac{VP}{(VP + FP)} \quad (6)$$

$$\text{medida F} = \frac{2 \times \text{precisão} \times \text{lembrança}}{(\text{precisão} + \text{lembrança})} \quad (7)$$

Estas medidas são baseadas na contagem dos registros classificados como positivos (P), negativos (N), verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN), resultados estes que compõe a matriz de confusão.

## RESULTADOS E DISCUSSÃO

Foi realizada uma análise nos modelos de classificação obtidos pela ferramenta, de modo a identificar o conjunto

de características mais pertinentes para determinar, ou refutar, a possibilidade de um indivíduo ser acometido pela doença de interesse. Em um primeiro momento, esta investigação consistiu na observação das estruturas gráficas, probabilidades e/ou relacionamentos entre as variáveis.

### Naïve Bayes – Resultados obtidos

De abordagem simples e semântica clara, o classificador Naïve Bayes tem como principal característica a suposição de que os atributos são condicionalmente independentes, dado o rótulo da classe<sup>(16,18)</sup>. Cabe ressaltar que, mesmo aplicado em situações em que não há independência condicional dos atributos, caso do presente trabalho, esse método é empregado em diversas situações, com obtenção de resultados bastante satisfatórios<sup>(18-19)</sup>. As probabilidades de cada atributo e da classe ocorrer são estimadas baseadas nas suas frequências no conjunto de treinamento. A classificação de um novo registro é realizada a partir da multiplicação destas probabilidades,  $P(X_i|C_j)$ , sendo o registro rotulado com a classe com maior valor proveniente desta multiplicação<sup>(9)</sup>. A Tabela 2 apresenta os quantitativos e as respectivas probabilidades referentes aos valores dos atributos mais representativos para que um registro seja classificado na classe de interesse, que indica um diagnóstico de DA. Por limitações de espaço, foram ocultadas as probabilidades dos demais valores de cada atributo.

Vale salientar que grande parte dos atributos gds, camcog, cdr e fluencia\_verbal não estavam preenchidos, o que pode indicar problemas na inserção das informações ou que tais exames não eram indicados, dado o estágio clínico do paciente. Pode-se notar que os resultados obtidos refletem essa situação.

### Redes Bayesianas – Resultados obtidos

As Redes Bayesianas (RBs) se caracterizam pela possibilidade de representarem dependências entre um subconjunto de variáveis. As RBs geralmente são representadas por um grafo acíclico direcionado em que

**Tabela 1** - Relação de atributos selecionados da base de dados após a etapa de pré-processamento e suas respectivas características.

Atributo	Valores	Descrição
Idade	0: < 65 anos, 1: >= 65 anos	Idade do paciente baseada no dia da primeira consulta
Genero	1: masculino, 2: feminino	Gênero
Escolaridade	0: [0], 1: [1-4], 2: [5-8], 3: >=9	Educação (em anos)
Meem	0: (0-17), 1: [18-26], 2: [27-30]	Resultado do Mini Exame do Estado Mental (MEEM)
Avd	0: grande dependência, 1: dependência moderada, 2: independência	Resultado da avaliação de Atividades de Vida Diária (AVD)
Aivd	1: independência, 2: dependência parcial, 3: dependência total	Resultado da avaliação de Atividades Instrumentais de Vida Diária (AIVD)
Gds	0: normal, 1: depressão, 2: depressão grave	Resultado da avaliação da Escala de Depressão Geriátrica
Camcog	A: alto, M: moderado, B: baixo	Resultado do índice de comprometimento cognitivo
Cdr	0: saudável, 0,5: questionável, 1: leve, 2: moderada, 3: grave	Escala clínica de demência ( <i>clinical dementia rating</i> - CDR)
Fluencia_Verbal	0: <= 10 palavras, 1: > 10 palavras	Quantidade de palavras em 1 minuto
Diabetes	0: negativo, 1: positivo	Possui diabetes ?
Has	0: negativo, 1: positivo	Possui hipertensão arterial ?
Diagnostico	0: negativo, 1: positivo	Acometido pela DA ?

os nós representam as variáveis e as arestas os relacionamentos de dependência entre um conjunto de variáveis. Cada variável do modelo possui uma tabela de probabilidades condicionais (TPC), sendo que cada tabela especifica a distribuição condicional em relação aos seus nodos pais imediatos.

O processo de classificação baseado em redes bayesianas é realizado em dois estágios, nessa ordem: aprendizado da estrutura da rede e aprendizado das TPC's. Após esta etapa, pode-se realizar inferências acerca de um conjunto de dados que reúna as variáveis aleatórias representadas no grafo.

A topologia da rede bayesiana representa a relação causal entre as variáveis envolvidas no processo de investigação pré-clínica para DA, conforme Figura 2.

Observa-se no grafo que os fatores de risco (idade e escolaridade) e os resultados dos testes neuropsicológicos (meem, avd, aivd e gds) impactam no resultado da classe de interesse, ou seja, o diagnóstico de DA. Consequentemente, esta relação influencia em outros aspectos clínicos, evidenciados pelos vértices (camcog, diabetes, has e genero) alcançados diretamente pelo nó diagnóstico.

A Tabela 3 apresenta as probabilidades relacionadas ao

diagnóstico de DA (*diagnostico=1*) mais pertinentes. Sua apresentação é parcial, por motivos de espaço, uma vez que a TPC completa possui 1920 instâncias.

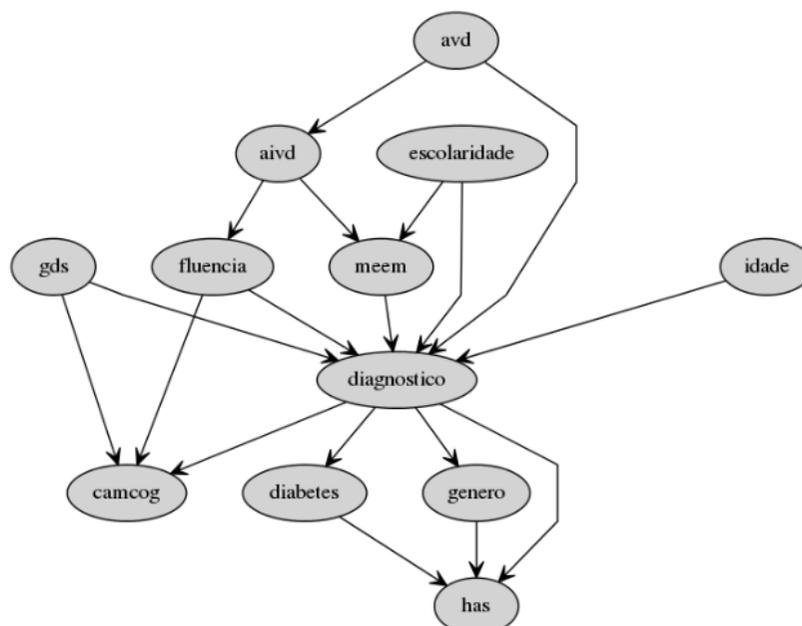
De acordo com a Tabela 3, é altamente provável (94,8 %) que indivíduos diagnosticados com DA sejam idosos com idade superior ou igual a 65 anos (*idade=1*), tenham desempenho ruim no Mini Exame do Estado Mental (*meem=0*), não tenham sido submetidos ao instrumento para detecção de depressão (*gds=NULO*) e sejam considerados independentes de acordo com a avaliação de Atividades de Vida Diária (*avd=2*). Deste percentual, os pacientes tendem a apresentar dependência em relação às Atividades Instrumentais da Vida Diária (AIVD), diferenciando-se somente pela intensidade avaliada, podendo ser considerados dependentes parciais (*aivd=2*) ou dependentes totais (*aivd=3*). Percebe-se na TPC que a diferenciação das AIVDs possui relação direta com a escolaridade, indicando que pacientes considerados analfabetos (*escolaridade=0*) têm como característica a dependência total (*aivd=3*), e os que possuem 9 anos ou mais de estudo (*escolaridade=3*) apresentam dependência parcial (*aivd=2*).

### Árvores de decisão – Resultados obtidos

A árvore de decisão foi a terceira técnica utilizada no

**Tabela 2 -** Modelo gerado pela técnica naïve bayes.

Atributo	Valor	CLASSE DA	
		Qtd	%
idade	1 65 anos ou mais de idade	463	0,905647912
genero	2 Feminino	337	0,667540979
escolaridade	1 1 a 4 anos de escolaridade	299	0,581984375
meem	0 0 a 17 pontos na avaliação	401	0,794598212
avd	2 Independência	266	0,523547945
aivd	3 Dependência total	207	0,402088063
gds		442	0,866970646
camcog		406	0,791520548
cdr		382	0,741639376
fluencia_verbal		442	0,866666667
diabetes	0 Negativo	362	0,706754222
has	1 Positivo	322	0,624678165



**Figura 2 -** Rede bayesiana para o problema de diagnóstico de DA

processo de mineração. Uma árvore de decisão é uma estrutura que divide de forma sucessiva um conjunto de registros em conjuntos cada vez menores, adotando a estratégia dividir-para-conquistar. Em uma árvore de decisão, o nó raiz e os nós internos contêm uma condição de teste acerca de um atributo, os ramos representam os valores deste atributo e os nós terminais (folhas) recebem as classes<sup>(18)</sup>. A construção da árvore de decisão com a estrutura descrita anteriormente é realizada por determinados métodos, chamados algoritmos de indução. Foi utilizado o algoritmo C4.5, que adota um critério denominado Razão do Ganho para selecionar os atributos para os nós. Detalhes do algoritmo podem ser encontrados em estudo realizado por Quinlan<sup>(20)</sup>. Mais especificamente, no presente trabalho, foi utilizado o algoritmo denominado J48, uma implementação do algoritmo C4.5 disponível na ferramenta Weka.

Dentre as diversas características do algoritmo, destacam-se as relacionadas à estrutura e ao processo de poda da árvore<sup>(18)</sup>. A forma da divisão dos nós seguiu uma abordagem binária, gerando um modelo mais representativo e de fácil entendimento para os especialistas do domínio. Visando à criação de uma árvore menos complexa, optou-se pela realização da poda, através dos parâmetros que representam o limite de confiança (*confidenceFactor*), ajustado para 0,35, e o número mínimo de instâncias por folha, cujo valor atribuído foi 10.

Conforme mencionado na seção MÉTODOS, após análise e comparação de diversas árvores geradas a partir da aplicação de diferentes parâmetros, um médico especialista identificou o modelo mais representativo. Por critérios de legibilidade, na Tabela 4 são apresentadas as regras (em formato SE ... E ... ENTÃO) para o diagnóstico de DA. Essas regras são resultantes dos diferentes percursos (do nó raiz à folha) da árvore considerada mais representativa pelo especialista. Vale salientar que na terceira coluna da Tabela 4

são apresentados dois números: o primeiro representa a quantidade de registros corretamente classificados para a classe de interesse, enquanto o segundo valor expressa o número de registros, dentre os classificados, que receberam o rótulo de classe erroneamente.

Percebe-se em todas as regras que o resultado do MEEM é o atributo mais relevante para configurar se o paciente é acometido por DA. Entre os pacientes que apresentaram valores de MEEM iguais a 1 ou a 2, os que obtiveram comprometimento cognitivo alto no camcog foram rotulados com DA (R1), ao passo que para as pessoas com desempenho baixo ou moderado, o resultado do teste de AIVD foi relevante na atribuição do diagnóstico. Deste subconjunto, os que se apresentaram totalmente dependentes (*aivd=3*) têm como fator para o diagnóstico a existência de hipertensão arterial (*bas=1*) (R2). Pacientes com 65 anos ou mais de idade com desempenho ruim no MEEM, não depressivos e incapazes de nomear 10 objetos em menos de 1 minuto tendem a serem acometidos pela doença (R3). Entretanto, deste conjunto, pacientes com menos de 65 anos de idade com boa escolaridade apresentam uma alta probabilidade de DA (R4). Deste subconjunto, os fatores determinantes para a doença são a baixa escolaridade, pertencer ao gênero feminino e possuir um quadro de hipertensão arterial positiva (R5).

#### Análise comparativa da precisão dos modelos

De acordo com a Tabela 5, observa-se que o classificador redes bayesianas apresentou não somente a melhor taxa de instâncias classificadas corretamente (acurácia), e conseqüentemente menor taxa de erro, como também obteve um alto percentual de registros positivos que foram corretamente classificados (sensibilidade), e uma alta porcentagem na medida F, além das taxas de falsos positivos (IFP) e negativos (IFN) menores.

Do ponto de vista da acurácia, a taxa obtida (73,8%)

**Tabela 3** - TPC parcial referente ao vértice diagnóstico da RB para diagnóstico de DA.

Variáveis pertinentes ao diagnóstico de DA						Diagnóstico	
idade	meem	gds	escolaridade	avd	aivd	0	1
1	0		3	2	2	5,2%	94,8%
1	0		0	2	3	5,2%	94,8%

**Tabela 4** - Regras extraídas a partir da árvore de decisão para o diagnóstico de DA.

Regra	Eventos	Totais
R1	SE resultado do meem com 18 ou mais pontos E resultado do camcog alto ENTÃO provável DA	20 / 7
R2	SE resultado do meem 18 ou mais pontos E resultado do camcog baixo ou moderado E paciente com dependência total no AIVD E possuir hipertensão arterial ENTÃO provável DA	17 / 7
R3	SE resultado do meem com 17 ou menos pontos E gds normal E fluência verbal com mais de 10 palavras/minuto E paciente com 65 ou mais anos de idade ENTÃO provável DA	479 / 135
R4	SE resultado do meem com 17 ou menos pontos E gds normal E fluência verbal com mais de 10 palavras por minuto E paciente com menos de 65 anos de idade E com pouca escolaridade ENTÃO provável DA	13 / 3
R5	SE resultado do meem com 17 ou menos pontos E gds normal E fluência verbal com mais de 10 palavras por minuto E paciente com menos de 65 anos de idade E com pouca escolaridade E gênero feminino E paciente possui hipertensão arterial ENTÃO provável DA	14 / 4

**Tabela 5** - Resultados obtidos dos métodos de classificação para DA.

Classificador	Acurácia	Taxa de erro	Sensibilidade	TFP	TFN	Medida F
naïve bayes	71,7%	28,3%	74,7%	30,1%	25,0%	73,1%
redes bayesianas	73,8%	26,2%	76,3%	28,7%	23,7%	73,0%
j48	72,0%	28,0%	72,5%	27,4%	27,8%	71,1%

mostra-se bem superior aos 68,5% observados em estudos semelhantes.<sup>(10-11)</sup> A sensibilidade por eles obtida foi de 68,8% e 66,0%, respectivamente, contra 76,3% do presente trabalho. Outra importante métrica, a Medida F, apresentou resultados melhores (73,0%) em relação aos 65,0% obtidos em um dos estudos.<sup>(11)</sup>

### CONCLUSÃO

Diante dos resultados numéricos ligados à precisão dos diferentes classificadores, concluiu-se que a acurácia dos modelos ficou aproximadamente entre 71,7% e 73,8%. Os modelos baseados nos classificadores bayesianos, especificamente redes bayesianas, apresentaram os melhores resultados para a classe DA em relação às demais técnicas de classificação, observação esta pautada nas métricas discutidas anteriormente. Em comparação com estudos relacionados ao uso de redes bayesianas para apoio ao diagnóstico de DA, os resultados mostraram-se coerentes e ligeiramente mais precisos.

Acredita-se que o presente estudo pode trazer algumas contribuições para a compreensão do processo de diagnóstico de pacientes com suspeita clínica de DA;

entretanto, algumas limitações devem ser ressaltadas. Apesar da unidade de saúde atender casos relacionados a síndromes demenciais, o escopo deste trabalho englobou somente a DA. Além disso, os dados utilizados representavam características dos pacientes relacionadas essencialmente a informações demográficas, fatores de risco e resultados de exames de rastreio, não sendo consideradas informações registradas em texto livre contidas no prontuário. Outro fato relevante é que, apesar de suficiente para o processo de mineração, o quantitativo de dados utilizado não engloba a totalidade de pacientes atendidos pelo programa. Finalmente, o alto índice de valores faltantes em alguns atributos seria outro fator complicador para obtenção de modelos mais precisos.

Estas limitações, no entanto, abrem perspectivas para o desenvolvimento de trabalhos futuros. Entre eles, pode-se destacar a inclusão de outras hipóteses diagnósticas; a descoberta de novas características descritas no prontuário médico, de forma a melhorar o poder preditivo dos modelos; a melhoria e ampliação das informações na base de dados; e a utilização de método de grupos de modo a aumentar a acurácia dos classificadores base.

### REFERÊNCIAS

1. The World Bank. Institute for Health Metrics and Evaluation, Human Development Network. The Global Burden of disease: generating evidence, Guiding Policy – Latin America and Caribbean Regional Edition. 2013 May [cited 2014 Jul 01]. Available from: [http://www.healthdata.org/sites/default/files/files/policy\\_report/2013/GBD\\_GeneratingEvidence/IHME\\_GBD\\_GeneratingEvidence\\_FullReport\\_PORTUGUESE.pdf](http://www.healthdata.org/sites/default/files/files/policy_report/2013/GBD_GeneratingEvidence/IHME_GBD_GeneratingEvidence_FullReport_PORTUGUESE.pdf)
2. United Nations Population Fund - UNFPA. Ageing in the Twenty-First Century: a celebration and a challenge. 2012 [cited 2014 Jul 03]. Available from: <http://www.unfpa.org/sites/default/files/pub-pdf/UNFPA-Exec-Summary.pdf>
3. Alzheimer Disease International - ADI. Policy brief for heads of government: the global impact of dementia 2013-2050. 2013 Dec [cited 2014 Jul 05]. Available from: <http://www.alz.co.uk/research/GlobalImpactDementia2013.pdf>
4. Gutierrez BAO, Silva HSS da, Guimarães C, Campino AC. Impacto econômico da doença de Alzheimer no Brasil: é possível melhorar a assistência e reduzir custos? *Cien Saúde Colet*. 2014;19(11):4479-86.
5. Machado JCB. Doença de Alzheimer. In: Freitas EV, Py L. Tratado de geriatria e gerontologia. 3a ed. Rio de Janeiro: Guanabara Koogan; 2011. p.178-210.
6. Passos URC, Matias IO de, Passos UR, Pereira MA, Carvalho VM. Uma ferramenta baseada em computação evolutiva para o apoio ao diagnóstico da cardiopatia isquêmica. *J. Health Inform*. 2014;6(4):153-60.
7. Shortliffe EH, Cimino JJ. Biomedical Informatics: computer applications in healthcare and biomedicine. 4a ed. New York: Springer; 2013.
8. Greenes RA. Clinical decision support: the road ahead. Elsevier; 2007.
9. Tan PN, Steinbach M, Kumar V. Introdução ao Data Mining: mineração de Dados. Rio de Janeiro: Ciência Moderna Ltda; 2009.
10. Mani S, Shankle WR, Pazzani MJ, Smyth P, Dick MB. Differential diagnosis of dementia: a knowledge discovery and data mining (KDD) approach. *Proc of AMIA Annu Fall Symp*. 1997;875(1):80-97.
11. Seixas LF. Sistema de apoio à decisão aplicado ao diagnóstico de demência, doença de Alzheimer e transtorno cognitivo leve [tese]. Niterói(RJ): Universidade Federal Fluminense - Programa de Pós-graduação em Computação; 2012.
12. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de Mendonça A. Data mining methods in the prediction of Dementia: a real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res Notes*. 2011;4:299.
13. Sun Y, Lv S, Tang Y. Construction and application of bayesian network in early diagnosis of Alzheimer Disease's system. In: International Conference on Complex Medical Engineering. CME 2007. IEEE/ICME; 2007 May 23-27; Beijing. IEEE; 2007. p.924-9.
14. Williams JA, Weakley A, Cook DJ, Edgecomb-Schmitter M. Machine learning techniques for diagnostic differentiation of mild cognitive impairment and dementia. In: First Workshop on Expanding the Boundaries of Health Informatics Using Artificial Intelligence; 2013 Jul 14-18; Bellevue, Washington, USA; 2013. p.71-6.
15. PostgreSQL. Documentação do Postgresql 9.2 [cited 2014 Jun 25]. Available from: <http://www.postgresql.org/docs/9.2/interactive/index.html>
16. Witten IH, Frank E, Hall MA. Data mining: practical machine learning tools and techniques. 3a ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc; 2011.
17. Moreira LB. Inteligência computacional para suporte à decisão na predição de diagnóstico relacionado a síndromes demenciais [Monografia]. Nova Friburgo(RJ): Universidade

- do Estado do Rio de Janeiro, Instituto Politécnico – Programa de Pós-graduação em Modelagem Computacional; 2014.
18. Han J, Kamber M, Pei J. Data mining: concepts and techniques. 3a ed. USA: Morgan Kaufmann Publishers; 2011.
  19. Abraham R, Simha JB, Iyengar SS. A comparative analysis of discretization methods for medical data mining with Naive Bayesian classifier. In: 9th International Conference on Information Technology, 2006. ICIT '06; 2006 18-21 Dez; Bhubaneswar; IEEE; 2006. p.18-21.
  20. Quinlan JR. C4.5: programs for machine learning. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc; 1993.