

“Uma Análise em possíveis Casos de Patologias Médicas utilizando a Curva ROC em Lógica Paraconsistente Anotada para Apoio a Decisão Médica em busca de melhor Precisão de Resposta na Web”

Maio 2008

Paulo Roberto Schroeder de Souza¹, UNIVERSIDADE SANTA CECÍLIA – prssouza@unisanta.br

Categoria F – Métodos e Tecnologia

Setor Educacional 3 – Educação Universitária

Natureza B – Descrição de Projeto em Andamento

Classe 1 – Investigação Científica

RESUMO

Este trabalho estabelece uma aplicação da Teoria de Bayes (TB) na Lógica Paraconsistente Anotada (LPA), chamamos de ParaBayes (PB), para diagnosticar a patologias de diferentes tipos que usam a mesma metodologia, como os dados de uma tabela 2x2, chamada matriz confusão ou de contingência, cruzando os resultados de um ou mais exame por exemplo, realizado para uma técnica considerada de referência (padrão ouro) e uma técnica com algoritmo será verificada.

A Teoria de Bayes permite delinear os parâmetros que ajudam a definir a área de ajuste de diagnóstico. Neste trabalho utilizamos nas análises o reticulado representativo aplicando o teorema de Bayes na Lógica Paraconsistente Anotada. Os valores de grau de certeza e de contradição foram participados com os valores calculados da sensibilidade e especificidade, valores clínicos da tabela Gold Standard, cálculos fornecendo os falso-positivos, falso-negativos, doentes e não doentes, portanto um quadro geral para a análise para duas ou mais propostas de mesma patologia. O Cálculo da Sensibilidade e do Falso Positivo (1-Especificidade) são valores dos eixos estabelecidos pela curva ROC, que pretendemos minimizar e maximizar através valores específicos dos grupos em estudo.

Palavras chave: Lógica Paraconsistente Anotada (LPA); Teorema de Bayes (TB); ParaBayes(PB); patologia; reticulado; Gold Standard; matriz confusão; matriz contingência, Falso Positivo, Falso Negativo, Sensibilidade, Especificidade, curva ROC.

¹ Orientador: Lincoln de Assis Moura Junior, PhD. Co-Orientador: João Inácio da Silva Filho, Dr.

1. Introdução

1.1 Idealização do Trabalho

Os resultados dos estudos apresentados neste trabalho são baseados numa classe de Lógica Não-Clássica denominada de Lógica Paraconsistente Anotada de anotação com dois valores (LPA2v). A análise de diagnóstico utilizando a LPA2v permite que vários problemas ocasionados por situações contraditórias e paracompletas possam ser tratados de uma maneira próxima da realidade, através da consideração de evidências e possibilitando ter uma proximidade melhor e mais exata de diagnóstico. A Lógica de Bayes vem como uma aplicação diagnóstica, com sua interação e adaptação na LPA, mostrar outro caminho diagnóstico tão bom ou melhor que a curva ROC como apoio e melhore a sua resposta pela LPA. Neste trabalho é demonstrada a funcionalidade de Bayes na Lógica Paraconsistente com a apresentação de formas inéditas de aplicações de uma classe de Lógica Não-Clássica originada dos estudos da Lógica Paraconsistente Anotada (LPA2v) [1].

Os métodos de aplicações obtidos, seguem os procedimentos estruturais teóricos da Lógica Paraconsistente Anotada a ser apresentada proporcionando uma forma prática e inovadora de fazer tratamento de dados contraditórios de modo não trivial fornecido especificamente pelo cálculo de Bayes na Sensibilidade e Especificidade.

1.2 Objetivos

Esta pesquisa objetiva apresentar métodos que vão proporcionar meios de aplicação da Teoria de Bayes na Lógica Paraconsistente, possibilitando o início de novas linhas de pesquisas mais aprofundadas na área de Apoio a decisão Médica. Os métodos colocados neste trabalho trazem resultados muito compensadores na obtenção de meios capazes de promover novas formas de tratamento do conhecimento incerto e/ou contraditório como o Erro de Bayes. Estes métodos podem ser aplicados em Sistemas de Diagnóstico. Portanto, os objetivos deste trabalho podem ser resumidos da seguinte forma, elaborar métodos de aplicações da Lógica Paraconsistente com base na sua estrutura teórica objetivando implementações práticas em Sistemas de Apoio a Decisão Médica. Gerar um Sistema Especialista.

2. Lógica Paraconsistente Anotada

2.1 Lógica Clássica

Tudo indica que foi com os trabalhos do filósofo grego Aristóteles, sábio que viveu entre 384 a 322 a.C. na cidade de Estagira, na Macedônia, que teve início os estudos da Lógica. O filósofo Aristóteles vivia na busca de um instrumento para descrever e a ajudar a compreensão de um mundo real e verdadeiro. Para a descrição do mundo a Lógica conhecida como Clássica foi criada e sustentada por rígidas leis, sem as quais não poderia existir. Era uma linha de raciocínio lógico baseado em premissas e conclusões do tipo: Se é observado que: "todo ser vivo é mortal" (premissa 1) e a seguir é constatado que "O leão é um ser vivo" (premissa 2) temos que: "O leão é mortal"(conclusão).

Dentro deste raciocínio a lógica Clássica tem sido binária, isto é, uma declaração é falsa ou verdadeira, não podendo ser ao mesmo tempo parcialmente verdadeira e parcialmente falsa. Com esta suposição, e a lei da não-contradição, onde uma declaração não pode contradizer a outra, todas as possibilidades foram cobertas pelas leis da Lógica Clássica formando assim a base do pensamento lógico Ocidental.

Na Lógica Clássica toda teoria que for inconsistente é trivial e vice-versa. Portanto, não há separação entre teorias inconsistentes e triviais [1].

2.2 Lógica Não-Clássica

Estudos mais aprofundados verificaram que no mundo real nem todas as situações podem ser classificadas simplesmente como verdadeiras ou falsas. Quando queremos precisão para descrever algo, fica difícil estabelecer limites, que nos permita afirmativas ou negativas a respeito de qualidade das coisas. Quase sempre os limites entre o “falso” e “verdadeiro” são indefinidos, incertos, ambíguos e até mesmo contraditórios. Para dar uma resposta mais satisfatória a problemas relacionados com situações não cobertas pela lógica clássica foram criadas as lógicas denominadas não-clássicas [1].

2.3 Lógica Paraconsistente

A Lógica Paraconsistente é uma lógica não-clássica criada para tratar contradições e assim permitir uma melhor adequação e análise de informações inconsistentes.

Os precursores da Lógica Paraconsistente foram o lógico polonês J. Lukasiewicz e o filósofo russo N.A. Vasilév, que simultaneamente, por volta de 1910 de maneira independente, ventilaram a possibilidade de uma Lógica diferente da clássica que restringiria, por exemplo, o princípio de contradição. Vasilév chegou mesmo a articular determinada Lógica, que ele batizou de imaginária, mas nenhum deles tinha, na época, uma vista ampla da Lógica Não-Clássica, tal como hoje a encaramos.

Uma Lógica Paraconsistente Anotada LPA é uma classe de Lógica Paraconsistente que possuem um reticulado associado que desempenha um papel sumamente importante na sua representação.

A Lógica Paraconsistente Anotada LPA é uma classe de lógica Paraconsistente Evidencial que faz tratamento de sinais representados por anotações permitindo uma descrição e equacionamento por meio de Algoritmos [2].

2.4 Representação da Lógica Paraconsistente Anotada LPA

A LPA pode ser estudada um Quadrado unitário no Plano Cartesiano (figura 1) onde são inseridos os graus de evidência favorável (crença) μ e graus de evidência desfavorável (descrença) λ onde através destes valores são calculados os graus de Certeza G_c e de Contradição G_{ct} conforme as equações abaixo [3]:

Sendo:

μ = grau de evidência favorável

onde:

$\mu \in [0,1]$

$P\mu$ = proposição anotada

Grau de evidência desfavorável = λ

$\lambda = 1 - \mu$

$\lambda \in [0,1]$

Na figura 1 verifica-se que podemos obter quatro situações possíveis:

1-Inconsistente; no ponto (1, 1) = T

2-Verdadeira; no ponto (1, 0) = V

1- Falsa; no ponto (0, 1) = F

2- Indeterminada; no ponto (0, 0) = \perp

A partir do Quadrado unitário podemos calcular os valores dos graus de contradição G_{ct} e dos graus de certeza G_c conforme as equações abaixo.

$G_c = \mu - \lambda$ e $G_{ct} = (\mu + \lambda) - 1$

Reticulado associado da Lógica Paraconsistente Anotada

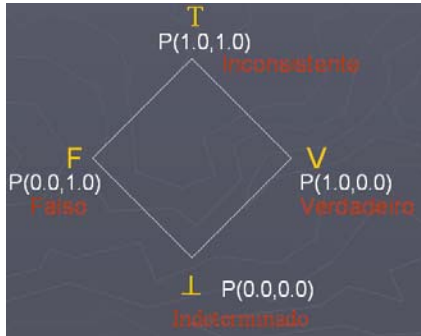


Figura 1. Reticulado representativo da LPA

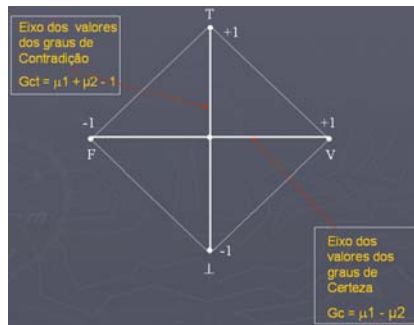


Figura 2. Eixos dos Graus de Contradição e de Certeza

Com o procedimento básico de interpretar os valores através de um Quadrado Unitário descrito no plano cartesiano podemos ter valores diversos de G_{ct} e G_c a partir dos graus evidência recebido todos normalizados e, portanto valorados na faixa entre 0 e 1 inclusive[4].

Dessa forma podemos agora saber, para qualquer anotação, com valores de graus evidência favorável (crença) e de evidência desfavorável (descrença), quais são os valores do Grau de Certeza e de Contradição, relacionado a determinada proposição. Fazendo a interpolação no reticulado isto indicará onde está localizado, no reticulado da LPA, o ponto resultante. A tomada de decisão é feita relacionando o resultado com proximidade, ou não, de algum dos estados lógicos extremos localizados nos vértices [4].

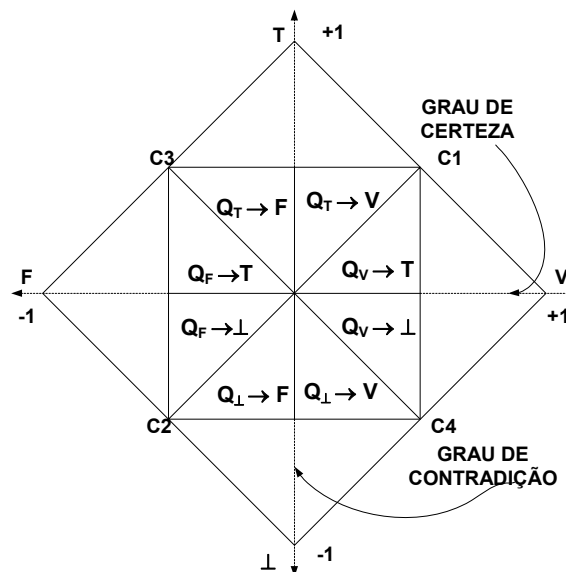


Figura 3. Representação de graus de Certeza e de Contradição

Estados não extremos	Símbolo
Quase-verdadeiro tendendo ao Inconsistente	$QV \rightarrow T$
Quase-verdadeiro tendendo ao Paracompleto	$QV \rightarrow \perp$
Quase-falso tendendo a Inconsistente	$QF \rightarrow T$
Quase-falso tendendo a Paracompleto	$QF \rightarrow \perp$
Quase-inconsistente tendendo ao Verdadeiro	$QT \rightarrow V$
Quase-inconsistente tendendo a Falso	$QT \rightarrow F$
Quase-paracompleto tendendo ao Verdadeiro	$Q\perp \rightarrow V$
Quase-paracompleto tendendo ao Verdadeiro	$Q\perp \rightarrow F$

Tabela 1. Controle Legenda Estados Não Extremos

Assim, tais valores de limite chamados Valores de Controle conforme Tabela 1.

$V_{cic} =$	Máximo valor de controle de incerteza	$= C_3$
$V_{cve} =$	Máximo valor de controle de certeza	$= C_1$
$V_{cpa} =$	Mínimo valor de controle de incerteza	$= C_4$
$V_{cfa} =$	Mínimo valor de controle de certeza	$= C_2$

Estados Extremos	Símbolo
Verdadeiro	V
Falso	F
Inconsistente	T
Indeterminado	\perp

Tabela 2. Estados da Lógica

3. Características da Curva ROC

Um dos primeiros usos dos gráficos ROC em aprendizado por máquina foi desenvolvido por [Spackman, 1989], que demonstrou a utilidade das curvas ROC na avaliação e comparação entre algoritmos.

Geralmente, a sensibilidade e a especificidade são características difíceis de conciliar, isto é, é complicado aumentar a sensibilidade e a especificidade de um teste ao mesmo tempo. As curvas ROC (Receiver Operating Characteristic) são uma forma de representar a relação, normalmente antagônica, entre a sensibilidade e a especificidade de um teste diagnóstico quantitativo ao longo de valores contínuos de ponto de corte.

Para construir uma curva ROC traça-se um diagrama que represente a sensibilidade em função da proporção de falsos positivos (1 - especificidade) para um conjunto de valores de ponto de corte [7].

Quando se tem uma variável contínua, resultado da aplicação de um teste diagnóstico quantitativo, e se pretende transformá-la numa variável dicotômica, do tipo doente/não doente, temos que utilizar um determinado valor na escala contínua que discrimine entre essas duas classes. A esse valor dá-se o nome de “ponto de corte” (cut off point).

O valor escolhido como ponto de corte vai influenciar as características do teste, como exemplificado na Figura 1. Neste exemplo, quanto maior o ponto de corte, maior a especificidade do teste, mas a sensibilidade será menor; e quanto menor o ponto de corte, maior a sensibilidade, mas a especificidade será menor. A Figura 2 representa graficamente a relação entre a sensibilidade e a especificidade para todos os possíveis pontos de corte da curva C_1 , C_2 e C_3 . Quanto maior for a sobreposição das curvas normais, menor será a área sob a curva ROC.

As curvas ROC descrevem a capacidade discriminativa de um teste diagnóstico para um determinado número de valores de ponto de corte. Isso permite colocar em evidência os valores para os quais existe uma maior otimização da sensibilidade em função da especificidade. O ponto numa curva ROC em que isso acontece é aquele que se encontra mais próximo do canto superior esquerdo do diagrama [8].

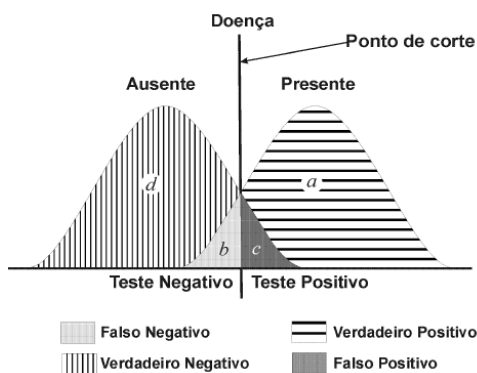


Figura 4. Ponto de Corte Figura [9]

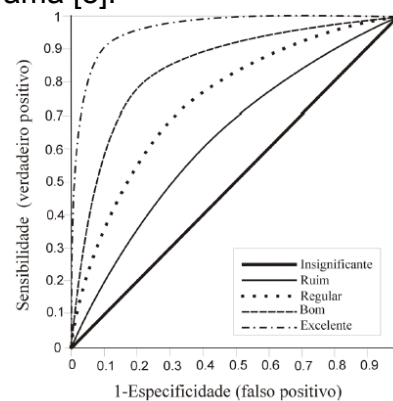


Figura 5. Curva ROC e a Qualidade no Diagnóstico [9]

Por outro lado, as curvas ROC permitem quantificar a exatidão de um teste diagnóstico, já que esta é proporcional a área sob a curva ROC (AUC), isto é, ela será tanto maior quanto mais a curva se aproximar do canto superior esquerdo do diagrama. Em virtude disso, a curva será útil também na comparação de testes diagnósticos, que terá uma exatidão tanto maior quanto maior for a área sob a curva ROC. O valor da área igual a 1 representa um teste perfeito; a área igual a 0.5 representa um valor sem importância. A Figura 3 exemplifica várias curvas ROC e a Tabela 3 associa à qualidade do diagnóstico a área da curva ROC [8].

Área (AUC)	Qualidade do diagnóstico
0.9 a 1.0	Excelente
0.8 a 0.9	Bom
0.7 a 0.8	Regular
0.6 a 0.7	Ruim
0.5 a 0.6	Insignificante

Tabela 3: Qualidade do diagnóstico em relação a área da curva ROC [8].

4. Metodologia de Aplicação

Ao fazermos a analogia de Bayes com a Paraconsistência teremos a seguinte tabela:

FALSO	1	NÃO DOENTE = ESPECIFICIDADE = VALOR PREDITIVO NEGATIVO
QUASE FALSO TENDENDO A PARACOMPLETO	2	QUASE NÃO DOENTE TENDENDO A FALSO NEGATIVO
PARACOMPLETO TENDENDO A FALSO	3	FALSO NEGATIVO TENDENDO A NÃO DOENTE
PARACOMPLETO	4	FALSO NEGATIVO = 1 - SENSIBILIDADE

PARACOMPLETO TENDENDO AO VERDADEIRO	5	FALSO NEGATIVO TENDENDO AO DOENTE
QUASE VERDADEIRO TENDENDO A PARACOMPLETO	6	QUASE DOENTE TENDENDO A FALSO NEGATIVO
VERDADEIRO	7	DOENTE – SENSIBILIDADE = VALOR PREDITIVO POSITIVO
QUASE VERDADEIRO TENDENDO AO INCONSISTENTE	8	QUASE DOENTE TENDENDO AO FALSO POSITIVO
INCONSISTENTE TENDENDO AO VERDADEIRO	9	FALSO POSITIVO TENDENDO DOENTE
INCONSISTENTE	10	FALSO POSITIVO=1 - ESPECIFICIDADE
INCONSISTENTE TENDENDO A FALSO	11	FALSO POSITIVO TENDENDO A NÃO DOENTE
QUASE FALSO TENDENDO A INCONSISTENTE	12	QUASE NÃO DOENTE TENDENDO A FALSO POSITIVO

Tabela 4 – Interpretação da Adaptação e Transição de ParaBayes

Para dois valores de sensibilidade e dois valores de especificidade, estudamos o caso de maximização (Or) e o caso de minimização, nestes casos teremos o melhor ponto de corte. Utilizando o quadriculado de Hasse conforme figura 6.

Tabela - Sensibilidade, especificidade e valores preditivos positivo e negativo da queixa de perda auditiva conforme o sexo e a faixa etária dos pacientes atendidos. Canoas, RS, 1998 a 2001.

Variável	Sensibilidade %	Especificidade %	VP+ %	VP- %
Sexo				
Feminino	79,5	62,5	85,8	51,6
Masculino	82,6	69,7	87,2	68,1
Faixa etária				
<7 anos	58,0	76,1	64,4	70,8
7 a 19 anos	83,4	72,1	80,2	76,2
20 a 60 anos	80,6	59,2	88,8	43,2
>60 anos	86,7	0	98,4	0
Total	80,9	69,7	86,5	60,4

VP: Valor preditivo

Tabela 5. Testes realizados em pacientes homens e mulheres com caso de audição [10]

Analisando estes dois casos e considerando único de modo geral:

Feminino

μ_{1A} = Sensibilidade = 79,5 %

μ_{1B} = Especificidade = 62,5 %

Valor Preditivo Positivo = 85,8%

Valor Preditivo Negativo = 51,6%

Masculino

μ_{2A} = Sensibilidade = 82,6 %; μ_{2B} = Especificidade = 69,7 %; Valor Preditivo Positivo = 87,2%; Valor Preditivo Negativo = 68,1%

Analisando para o modo de **Maximização**:

$\mu_{1A} \vee \mu_{2A} = 79,5 \vee 82,6 = 82,6\%$ e $\mu_{2B} \vee \mu_{1B} = 62,5 \vee 69,7 = 69,7\%$

Analisando para o modo **Minimização**:

$\mu_{1A} \wedge \mu_{2A} = 79,5 \wedge 82,6 = 79,5\%$ e $\mu_{2B} \wedge \mu_{1B} = 62,5 \wedge 69,7 = 62,5\%$

Representação no Plano Cartesiano, temos [5]:

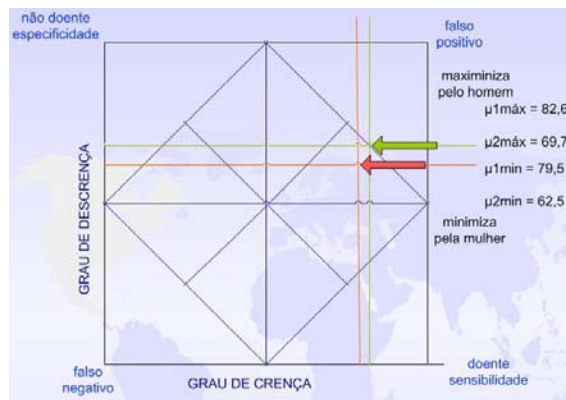


Figura 6. Maximização e Minimização em Graus.

Análise: Falso-Positivo (Max) e Falso-Positivo tendendo a Doente (Min), concluinte.

5. O Algoritmo ParaBayes para Apoio a Decisão Médica

O algoritmo ParaBayes segue o seguinte procedimento:

- Tela Inicial;
- Digite os seguintes valores: verdadeiramente doente, verdadeiramente sadio, falso-positivo, falso-negativo;
- Aplicação das Funções;
- Calcula: Sensibilidade, Especificidade, 1-Especificidade, Valor Preditivo Positivo, Valor Preditivo Negativo, Razão de Verossimilhança Positivo, Razão de Verossimilhança Negativo;
- Calcula Intervalo de Confiança – IC;
- Calcula Erros Possíveis – Erro de Bayes;
- Calcula μ_1 , μ_2 /*valores de crença e descrença*/
- Calcula G_c , G_{ct} /*Grau de certeza e Grau de contradição*/
- Digite Entrada de Valores das Fronteiras – Controle;
- /* Valor superior de controle de certeza */
- /* Valor inferior de controle de certeza */
- /* Valor superior de controle de incerteza */
- /* Valor inferior de controle de incerteza */
- Definido os valores, são analisados os resultados para a região de controle: No plano Cartesiano, no Quadriculado de Hasse;
- Fornece Resultados
- Os resultados são diagnosticados por subárea;
- As providências são sugeridas;
- Pergunta-se: “Deseja modificar os valores de controle para uma nova análise?”
Mesma análise? Retorna a Tela Inicial. Ou recalcula para novos valores;
- Fim.

6. Conclusão

A estrutura básica dos testes diagnósticos é similar aos outros estudos observacionais, e são desenhados para determinar quão bem um teste discrimina a presença ou ausência de uma doença. Possui uma variável preditora, o resultado do teste, e uma variável resultante, a presença ou ausência da doença. A variável preditora do teste embora possa ser categórica ou contínua, habitualmente é dicotomizada, definindo os

critérios (cut-off) que confirmam ou refutam a presença da doença. A variável resultante do teste diagnóstico é a presença ou ausência da doença, confirmada por um exame considerado "Gold Standard".

A decisão médica com base em Curvas ROC (Receiver Operating Curves) é apresentada como aplicação em problemas no âmbito da investigação em medicina. É um gráfico que se propõe a estimar a capacidade de um teste seletivo de discriminar entre pessoas saudáveis e doentes. A análise ROC é uma ferramenta poderosa para medir e especificar problemas no desempenho do diagnóstico em medicina. Esta análise por meio de um método gráfico simples e robusto, permite estudar a variação da sensibilidade e especificidade, para diferentes valores de corte. A área abaixo da curva ROC está associada ao poder discriminante de um teste de diagnóstico. Geometricamente, a curva ROC é um gráfico de pares "x"e"y" (que correspondem, a (1- especificidade) e à sensibilidade, respectivamente) num plano designado por plano ROC unitário. A designação de plano ROC unitário deve-se ao fato das coordenadas deste gráfico representar medidas de probabilidade, e, por conseguinte variarem entre zero e um. Para a visualização da curva ROC, utilizou-se a representação desta no plano ROC unitário e no plano da Paraconsistência. Foi apresentado um estudo de caso ilustrativo, no campo da análise de diagnóstico em medicina, para melhor compreensão da metodologia em estudo. Uma nova metodologia, que se encontra em desenvolvimento, um algoritmo, para comparação de curvas ROC, quando estas se cruzam ou quando os casos se unem para verificar o máximo e o mínimo nesta condição de análise diagnóstica Paraconsistente.

Referências

- [1] Da Silva Filho, João Inácio. Jair Minoro Abe
Introdução à Lógica Paraconsistente Anotada - Editora Emmy – 1.^a Edição, 2000.
- [2] Abe, J. M. “*Fundamentos da Lógica Anotada*” (Foundations of Annotated Logics), (in Portuguese) Ph. D. Thesis, University of São Paulo, São Paulo, 1992.
- [3] Abe, J. M. Paraconsistent Artificial Neural Networks: an Introduction, Lecture Notes In Artificial Intelligence 3214, Springer, Eds. J.G. Carbonell & J. Siekmann, ISSN 0302-9743, ISBN 3-540-23206-0, pp. 942-948, 2004.
- [4] Filho, J. I. Da Silva & J. M. Abe, Fundamentos das Redes Neurais Paraconsistentes – Destacando Aplicações em Neurocomputação, (in Portuguese) Editora Arte & Ciência, ISBN 85-7473-045-9, 247 pp., 2001.
- [5] Da Silva Filho, João Inácio
Métodos de Aplicações da Lógica Paraconsistente Anotada
de anotação com dois valores-LPA2v com construção de
Algoritmo e Implementação de Circuitos Eletrônicos. POLITÉCNICA - USP
São Paulo, 1999. Tese (Doutorado)
- [6] Dempster, A.P. - Generalization of Bayesian inference, *Journal of the Royal Statistical Society*, Series B-30, 205-247, 1968.
- [7] Massad, Eduardo; Renée X. de Menezes; Paulo S.P.Silveira;Neli Regina S. Ortega

Métodos Quantitativos em Medicina, 2004
Editora Manole Ltda - 1.^a Edição

[8] Silva, Aristófanés Correia
Algoritmos para Diagnóstico Assistido de Nódulos
Pulmonares Solitários em Imagens de Tomografia
Computadorizada/ Aristófanés Corrêa Silva; orientador:
Marcelo Gattass; co-orientador: Paulo Cezar Pinto
Carvalho. — Rio de Janeiro : PUC-Rio, Departamento de Informática, 2004.v., 140 f: il.30 cm

[9] Braga, Ana Cristina da Silva.
Tese de doutorado, Curvas ROC:Aspectos Funcionais e Aplicações. Universidade do Minho, dezembro, 2000.

[10] Marini, Ana Lúcia Sant'Anna, Ricardo Halpern e Denise Aerts. Artigo: Sensibilidade, especificidade e valor preditivo da queixa auditiva. Baseado na dissertação de mestrado apresentada ao Programa de Pós-Graduação em Saúde Coletiva da Universidade Luterana do Brasil, em 2003. Recebido em 4/2/2005. Reapresentado em 9/5/2005. Aprovado em 22/7/2005. Revista Saúde Pública, 39(6) :982-4, ano: 2005.