



FABRÍCIO REIS DA SILVA

**DIAGNÓSTICO DE GLAUCOMA BASEADO EM
CLASSIFICADORES DE APRENDIZAGEM DE MÁQUINA
UTILIZANDO DADOS DO *SPECTRAL DOMAIN-OCT* E
PERIMETRIA AUTOMATIZADA ACROMÁTICA**

**CAMPINAS
2013**



**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS MÉDICAS**

FABRÍCIO REIS DA SILVA

**DIAGNÓSTICO DE GLAUCOMA BASEADO EM
CLASSIFICADORES DE APRENDIZAGEM DE MÁQUINA
UTILIZANDO DADOS DO *SPECTRAL DOMAIN-OCT* E
PERIMETRIA AUTOMATIZADA ACROMÁTICA**

Orientador: Prof. Dr. Vital Paulino Costa

**Tese de Mestrado apresentado à Comissão de Pós-Graduação da
Faculdade de Ciências Médicas da Universidade Estadual de Campinas
para obtenção de Título de Mestre em Ciências Médicas, Área De
Concentração em Oftalmologia.**

**ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA
TESE DEFENDIDA PELO ALUNO FABRÍCIO REIS DA SILVA
ORIENTADO PELO PROF. DR. VITAL PAULINO COSTA
Assinatura do Orientador: _____**

CAMPINAS

2013

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Ciências Médicas
Maristella Soares dos Santos - CRB 8/8402

Si38d Silva, Fabrício Reis, 1979-
Diagnóstico de glaucoma baseado em classificadores de aprendizagem de máquina utilizando dados do *Spectral Domain-OCT* e perimetria automatizada acromática / Fabrício Reis da Silva. – Campinas, SP : [s.n.], 2013.

Orientador: Vital Paulino Costa.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Ciências Médicas.

1. Glaucoma. 2. Inteligência artificial. 3. Tomografia de coerência óptica. I. Costa, Vital Paulino. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Machine learning classifiers for glaucoma diagnosis using *Spectral Domain-OCT* and standard automated perimetry

Palavras-chave em inglês:

Glaucoma

Artificial intelligence

Optical coherence tomography

Área de concentração: Oftalmologia

Titulação: Mestre em Ciências Médicas

Banca examinadora:

Vital Paulino Costa [Orientador]

Mônica de Cássia Alves de Paula

José Paulo Cabral de Vasconcelos

Data de defesa: 26-08-2013

Programa de Pós-Graduação: Ciências Médicas

BANCA EXAMINADORA DA DEFESA DE MESTRADO

FABRÍCIO REIS DA SILVA

Orientador (a) PROF(A). DR(A). VITAL PAULINO COSTA

MEMBROS:

1. PROF(A). DR(A). VITAL PAULINO COSTA



2. PROF(A). DR(A). MONICA DE CASSIA ALVES DE PAULA



3. PROF(A). DR(A). JOSÉ PAULO CABRAL DE VASCONCELLOS



Programa de Pós-Graduação em Ciências Médicas da Faculdade de Ciências Médicas
da Universidade Estadual de Campinas

Data: 26 de agosto de 2013

Dedicatória

*Aos meus pais, pelos exemplos de vida que são;
à minha esposa, Suzana, pela compreensão, apoio e paciência;
E ao meu maior tesouro, Felipe.*

Agradecimentos

Ao Prof. Dr. Vital Paulino Costa, meu grande orientador, pela oportunidade e pela atenção dispendida neste trabalho;

Ao Dr. Edson Gomi e Dr. Marcelo Dias, por me auxiliarem na compreensão do fascinante mundo da inteligência artificial;

Aos colegas do Setor de Glaucoma da UNICAMP, em especial Vanessa Vidotti, pelo apoio e incentivo;

Aos funcionários da UNICAMP, em especial, Sílvia e Márcia, que trabalharam para tornar este projeto possível;

À FAPESP, que disponibilizou recursos para esta pesquisa;

Aos meus familiares, irmãos, cunhados e sogros, pelo apoio, incentivo e amizade.

SUMÁRIO

| | |
|---|-------|
| RESUMO | xxxv |
| ABSTRACT | xxxix |
| 1. Introdução | 43 |
| 1.1 Glaucoma | 45 |
| 1.1.2 Epidemiologia | 46 |
| 1.1.3 Relação estrutura e função | 47 |
| 1.2 Campo visual | 50 |
| 1.2.1 “Mean Deviation” | 50 |
| 1.2.2 “Pattern Standard Deviation” | 51 |
| 1.2.3 “Glaucoma Hemifield Test” | 51 |
| 1.2.4 Estratégia SITA | 53 |
| 1.3 OCT | 55 |
| 1.3.1 OCT Cirrus | 58 |
| 1.3.2 Reprodutibilidade do SD-OCT | 62 |
| 1.3.2 Capacidade diagnóstica | 63 |
| 1.4 Conceitos sobre aprendizagem de máquina | 64 |
| 1.5 Classificadores de aprendizagem de máquina | 67 |
| 1.5.1. Árvore de Decisão (Ctree) | 67 |
| 1.5.2. Rede Neural Multicamadas: “Multilayer Perceptrons” | 74 |
| 1.5.3. Rede de Funções de Base Radial | 78 |
| 1.5.4. Classificador “Naive Bayes” | 80 |
| 1.5.5. “Support Vector Machines” | 85 |
| 1.5.5.1 Svm linear | 89 |
| 1.5.5.2 Svm não linear (Gaussiana) | 90 |
| 1.5.6. Classificador “Bagging” | 91 |
| 1.5.7. Classificador “Adaboosting” | 95 |
| 1.5.8. “Random Forest” | 104 |
| 1.5.9. “Ensemble selection” | 107 |
| 2. Justificativa | 109 |
| 3. Objetivos | 113 |

| | |
|--|-----|
| 4. Métodos | 117 |
| 4.1. População | 119 |
| 4.1.1. Critérios de inclusão para todos os participantes | 119 |
| 4.1.2. Critérios de exclusão para todos os participantes | 120 |
| 4.1.3. Critérios de inclusão para indivíduos normais | 120 |
| 4.1.4. Critérios de inclusão para Indivíduos glaucomatosos | 121 |
| 4.2. Métodos | 122 |
| 4.3. Tomografia de Coerência Óptica (OCT) | 123 |
| 4.4. Perimetria Automatizada Acromática (PAA) | 124 |
| 4.5. Classificadores de Aprendizagem de Máquina | 124 |
| 4.6. Validação cruzada (“cross validation”) | 126 |
| 4.7. WEKA | 127 |
| 4.8. Sensibilidade | 128 |
| 4.9. Especificidade | 128 |
| 4.10. Curva ROC | 129 |
| 4.11. Análise Estatística | 132 |
| 5. Resultados | 133 |
| 6. Discussão | 145 |
| 7. Conclusão | 155 |
| 8. Referências bibliográficas | 159 |
| 9. Anexos | 171 |

LISTA DE ABREVIATURAS:

ADA:ada boosting M1

AM: aprendizagem de máquina

aROC: área abaixo da curva roc

AV: acuidade visual

BAG: bootstrap aggregating

CAM: classificador de aprendizagem de máquina

CCD: charge-coupled device

CCI: coeficiente de correlação intraclasse

CFN: camada de fibras nervosas

COV: coeficiente de variação

CTREE: árvore de decisão

CV: campo visual

D: dioptria

dB: decibel

EE: equivalente esférico

EMGT: Early Manifest Glaucoma Trial

ENS: ensemble selection

FDT: perimetria de frequência dupla

GCP: glaucoma change probability

GDX VCC: polarimetria de varredura à laser

GHT: glaucoma hemifield test

GPAA: glaucoma primário de ângulo aberto

HD: alta definição (High Definition)

HRT II: Heidelberg Retina Tomograph II

IA: inteligência artificial

IC: intervalo de confiança

LOCS-III: Lens Opacification Classification System III

MD: mean deviation

CAM OCT+CV: classificador de aprendizagem de máquina utilizando dados combinados do campo visual e tomografia de coerência óptica

MLC: machine learning classifier

MLP: rede neural multicamada (multilayer perceptrons)

N: número de elementos da amostra

NB: naive bayes

OCT: Tomografia de coerência óptica

OHTS: Ocular Hypertension Treatment Study

PAA: perimetria automatizada acromática

PSD: pattern standard deviation

RAN: random forest

RBF: rede de funções de base radial (Radial Basis Function Network)

RNA: rede neural artificial

RNFL: retinal nerve fiber layer

ROC: Receiver operating characteristic

SAP: standard automated perimetry

SD-OCT: tomografia de coerência óptica de domínio espectral (spectral domain OCT)

Se: sensibilidade

SLD: diodo superluminescente de baixa coerência

Sp: especificidade

SVM: máquinas de vetores de suporte (support vector machines)

SVMg: support vector machine gaussian

SVMl: support vector machine linear

SWAP: perimetria azul-amarelo

TAE: teoria de aprendizado estatístico

TD-OCT: tomografia de coerência óptica de domínio temporal (Time Domain-OCT)

UNICAMP: Universidade de Campinas

USP: Universidade de São Paulo

VC: Vapnik-Chervoneskis

WEKA: Waikato Environment for Knowledge Analysis

LISTA DE FIGURAS:

Figura 1: Exemplo de paciente portador de glaucoma apresentando neuropatia óptica glaucomatosa com defeito de campo visual correspondente.

Figura 2: Exemplo de campo visual utilizando estratégia SITA de paciente com defeito arqueado em hemisfério superior em olho esquerdo.

Figura 3: Desenho esquemático do sistema TD-OCT. A fonte de luz é um diodo superluminescente de baixa coerência acoplado a uma fibra óptica. O espelho de referência (canto superior direito) é movimentado para produzir diferentes tempos de atraso (referência) que são comparados com o feixe de luz refletido do olho (canto inferior direito).

Figura 4: Desenho esquemático do sistema SD-OCT. A fonte de luz é um diodo superluminescente de baixa coerência (SLD) (canto superior esquerdo) acoplado a uma fibra óptica e conectado ao espectrômetro + câmera CCD (canto inferior esquerdo). O espelho de referência (canto superior direito) permanece fixo.

Figura 5: Parâmetros avaliados pelo Cirrus OCT “optic disc cube 200x200”.

Figura 6: Fluxograma para construção de um classificador.

Figura 7: Representação esquemática de uma árvore de decisão. Os círculos representam os nós e os retângulos representam as folhas, sendo o círculo superior o nó raiz. Cada nó de decisão contém um teste para o atributo ($<a$ ou $>a$, negativo ou positivo para o teste do nó em questão (exemplos: a_1 , a_2 , a_3 ,...)). Cada folha (retângulo) está associada a uma classe; o percurso na árvore (da raiz à folha) corresponde a uma regra de classificação.

Figura 8: Árvore de decisão gerada a partir dos dados da tabela 2. A árvore possui cinco nós folha (retângulos verdes), que apresentam os valores (número) e probabilidades (número entre parêntese) da classe e classificam de acordo com os atributos.

Figura 9: Representação de uma rede neural MLP. Apresenta uma camada de entrada, uma de saída e apenas uma camada escondida. O sentido de propagação é único (da esquerda para direita), da camada de entrada em direção à camada de saída, passando pela camada escondida.

Figura 10: Representação de uma rede MLP com retro propagação. O sentido de propagação é representado pela linha contínua (da esquerda para direita), da camada de entrada em direção à camada de saída. Por outro lado, o sentido da retro propagação, representado pela linha tracejada, ocorre no sentido contrário (da direita para esquerda), da camada de saída em direção à camada de entrada.

Figura 11: Representação de uma rede de base radial. A RBF apresenta uma camada de entrada (com neurônios X_1 a X_m representados pelos quadrados vazios); uma camada oculta, que apresenta neurônios (representados pelos círculos de W_1 a W_x) que possuem funções de ativação não linear, funções de base radial; e uma camada de saída (representada pelo quadrado com símbolo Σ no seu interior) que possui função de ativação linear que obtém uma saída y .

Figura 12: Amostra desconhecida para ser classificada de acordo com o classificador NB.

Figura 13: Exemplos de três classificadores distintos. O classificador é representado pela linha contínua que separa os dados pertencentes a duas classes (triângulos e círculos).

Figura 14: Representação de margens de separação de hiperplanos. Duas classes (quadrados vermelhos e círculos azuis) são separadas por um hiperplano (linha contínua) e

a maior distância entre as classes (espaço entre as linhas tracejadas) corresponde à margem de separação. A figura (a) apresenta um hiperplano de separação com margem pequena. A figura (b) apresenta um hiperplano de margem máxima, pois apresenta a maior distância entre as linhas tracejadas.

Figura 15: Exemplo de criação de um espaço com maior dimensão separável por um hiperplano. Na figura (a) temos um conjunto de dados (triângulos pretos e círculos brancos) não separáveis linearmente. Na figura (b) é representada uma fronteira não linear onde é possível separar as duas classes. Na figura (c) é representada a criação de um espaço de maior dimensão (x_3) a partir de um espaço original (x_1 e x_2) possibilitando a criação de um hiperplano (retângulo com linhas tracejadas) que pode separar as duas classes corretamente.

Figura 16: Representação esquemática das etapas para criação do classificador “Bagging”. Na primeira etapa, são geradas amostras aleatórias, a partir do conjunto de treinamento. Na segunda etapa, são gerados classificadores derivados de cada amostra. Na terceira etapa, é criado um classificador composto, que inclui todos os classificadores individuais e, então, é realizado uma votação simples (quarta etapa) para eleger a classificação mais popular.

Figura 17: Classificadores gerados para as três amostras da Tabela 4.

Figura 18: Representação do funcionamento do método “Boosting”. As amostras são geradas de forma ponderada, ou seja, diminui-se a probabilidade de incluir amostras classificadas corretamente e aumenta-se a probabilidade de instâncias classificadas incorretamente. Após a geração de K amostras com K classificadores, uma votação ponderada é realizada para classificar a instância desejada.

Figura 19: Primeira amostra com distribuição uniforme, criada a partir da tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado. Todas as instâncias apresentaram a mesma probabilidade de serem incluídas na amostra.

Figura 20: Segunda amostra gerada utilizando a taxa de erro (ponderação), criada a partir da tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado.

Figura 21: Terceira amostra gerada utilizando a taxa de erro (ponderação), criada a partir da tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado.

Figura 22: Representação do gráfico ROC e dos pontos A (0,0), B (100,100), C(0,100) e D (100,0).

Figura 23. Exemplo de comparação de duas curvas ROCs sem intersecção (curvas vermelho e azul). A curva vermelha está mais próxima do ponto (0,100) que a curva azul, por isso, o modelo que é representado por esta curva apresenta melhor desempenho que o modelo representado pela curva azul.

Figura 24. Exemplo de comparação de duas curvas ROCs com intersecção (curvas vermelho e azul). Próximo ao ponto (40%,70%) há a intersecção das duas curvas. Analisando as curvas à esquerda do ponto (40%,70%), a curva vermelha está mais próxima do ponto (0,100%) que a curva azul, por isso, o modelo que é representado por esta curva apresenta melhor desempenho que o modelo representado pela curva azul. Analisando as curvas à direita do ponto (40%,70%) ocorre o inverso, melhor desempenho do modelo representado pela curva azul.

Figura 25: Representação das curvas ROC dos parâmetros que apresentaram maiores aROCs do classificador de aprendizagem de máquina usando dados combinados do OCT e

perimetria (RAN=0,946), do OCT (região inferior=0,813) e da perimetria acromática (PSD=0,915).

LISTA DE TABELAS E QUADROS

Tabela 1: Porcentagem dos indivíduos incluídos nos bancos de dados do Cirrus OCT e Stratus OCT de acordo com a etnia.

Tabela 2: Atributos (parâmetros) para geração de um exemplo de árvore de decisão baseada nos dados dos associados do clube (idade, salário, ensino superior, dependentes, pagamento).

Tabela 3: Dados de uma empresa financeira para geração de um exemplo de um classificador NB.

Tabela 4: Base de dados para criação de exemplos de classificadores “Bagging” e “Adaboosting”. São 14 instâncias que apresentam três atributos (frequência cardíaca, frequência respiratória e perda do apetite) e um atributo preditor (uso de medicação).

Tabela 5: Três amostras, com a mesma quantidade de instâncias (14 instâncias), geradas aleatoriamente a partir da tabela 4 pelo método de reposição.

Tabela 6: Três instâncias a serem classificadas pelo método “Bagging” baseados nos dados da tabela 4 e nas amostras geradas aleatoriamente da tabela 5.

Tabela 7: Instâncias a serem classificadas pelo classificador “Boosting”.

Tabela 8: Aspectos clínicos e demográficos (idade, sexo, raça, acuidade visual, equivalente esférico, pressão intraocular, uso de medicações, “Mean Deviation”, “Pattern Standard Deviation”) dos indivíduos normais e glaucomatosos e a significância estatística da diferença entre os grupos.

Tabela 9: Áreas abaixo da curva ROC (aROC), sensibilidade (%) com especificidade fixa de 80% e 90% para cada parâmetro da perimetria acromática (MD, PSD, GHT) e os

parâmetros do SD-OCT (espessura global, 4 quadrantes e 12 segmentos em horas de relógio) sem utilização dos classificadores de aprendizagem de máquina.

Tabela 10: Avaliação das aROCs dos classificadores de aprendizagem de máquina (BAG, NB, MLP, RBF, RAN, ENS, CTREE, ADA, SVMg e SVMl) usando dados da perimetria e OCT. Foram avaliados: as aROCs, utilizando todos os 20 atributos, a melhor aROC com o menor número de parâmetros, e a sensibilidade (%) com especificidade fixa de 80% e 90%.

Tabela 11: Comparação das áreas abaixo da curva ROC dos classificadores OCT+CV (com número de parâmetros que apresentaram o melhor desempenho) com os melhores parâmetros da perimetria (PSD) e do OCT (região inferior).

Quadro 1: Estudos que avaliaram a capacidade diagnóstica do Cirrus OCT e compararam com o Stratus OCT.

Quadro 2: Vetor de pesos das instâncias da primeira amostra gerada pelo “Boosting” no passo 1.

Quadro 3: Vetor de pesos das instâncias da primeira e segunda amostra gerada pelo “Boosting” no passo 2.

RESUMO

Objetivo: Avaliar a sensibilidade e especificidade dos classificadores de aprendizagem de máquina no diagnóstico de glaucoma usando dados do Spectral Domain OCT (SD-OCT) e perimetria automatizada acromática (PAA).

Desenho do estudo: Estudo transversal observacional.

Método: Sessenta e dois pacientes com glaucoma e 48 indivíduos normais foram incluídos. Todos os pacientes foram submetidos a exame oftalmológico completo, PAA (24-2 SITA; Humphrey Field Analyzer II, Carl Zeiss Meditec, Inc., Dublin, CA) e exame de imagem da camada de fibras nervosas utilizando SD-OCT (Cirrus HD-OCT; Carl Zeiss Meditec Inc., Dublin, California). Curvas ROC (Receiver Operating Characteristic) foram obtidas para todos os parâmetros do SD-OCT e índices globais da perimetria acromática (MD, PSD, GHT). Subsequentemente, os seguintes classificadores de aprendizagem de máquina (CAMs) foram testados usando parâmetros do OCT e PAA: Bagging (BAG), Naive-Bayes (NB), Multilayer Perceptron (MLP), Radial Basis Function (RBF), Random Forest (RAN), Ensemble Selection (ENS), Classification Tree (CTREE), Ada Boost M1(ADA), Support Vector Machine Linear (SVML) e Support Vector Machine Gaussian (SVMG). Áreas abaixo da curva ROC (aROC) obtidas com os parâmetros isolados do campo visual (CV) e OCT foram comparados com os CAMs usando dados combinados do OCT e CV.

Resultados: A média de idade foi de $57,0 \pm 9,2$ anos para os indivíduos normais e $59,9 \pm 9,0$ anos para os pacientes glaucomatosos ($p=0,103$). As médias dos valores do Mean Deviation (MD) foram $-4,1 \pm 2,4$ dB para glaucomatosos e $-1,5 \pm 1,6$ dB para indivíduos normais ($p<0,001$). Os parâmetros do SD-OCT com maiores aROCs foram: quadrante inferior (aROC=0,813) e espessura média (aROC=0,807). O melhor parâmetro do CV foi

PSD (aROC=0,915). Combinando os dados do OCT e do CV, aROCs dos CAMs variaram entre 0,777 (CTREE) e 0,946 (RAN). A maior aROC dos CAMs OCT+CV, obtida com RAN (0,946) foi significativamente maior que o melhor parâmetro do OCT ($p=0,003$), mas não houve diferença estatisticamente significativa com o melhor parâmetro do CV ($p=0,37$).

Conclusão: Os classificadores de aprendizagem de máquina treinados com dados do OCT e CV podem discriminar entre olhos normais e glaucomatosos com sucesso. A combinação das medidas do OCT e CV melhoraram a acurácia diagnóstica comparados aos parâmetros do OCT.

ABSTRACT

Purpose: To evaluate the sensitivity and specificity of machine learning classifiers (MLCs) for glaucoma diagnosis using Spectral Domain OCT (SD-OCT) and standard automated perimetry (SAP).

Design: Observational cross-sectional study.

Population: Sixty two glaucoma patients and 48 healthy individuals were included.

Methods: All patients underwent a complete ophthalmologic examination, SAP and retinal nerve fiber layer (RNFL) imaging with SD-OCT (Cirrus HD-OCT; Carl Zeiss Meditec Inc., Dublin, California). Receiver operating characteristic (ROC) curves were obtained for all SD-OCT parameters and some SAP parameters (MD, PSD, GHT). Subsequently, the following MLCs were tested using parameters from SD-OCT and SAP: Bagging (BAG), Naive-Bayes (NB), Multilayer Perceptron (MLP), Radial Basis Function (RBF), Random Forest (RAN), Ensemble Selection (ENS), Classification Tree (CTREE), Ada Boost M1 (ADA), Support Vector Machine Linear (SVML) and Support Vector Machine Gaussian (SVMG). Areas under the ROC (aROC) obtained with SAP and OCT parameters were compared with those obtained with MLCs using combined OCT+SAP data.

Results: The mean age was 57.0 ± 9.2 years for healthy individuals and 59.9 ± 9.0 years for glaucoma patients ($p=0.103$). Mean Deviation (MD) values were -4.1 ± 2.4 dB for glaucoma patients and -1.5 ± 1.6 dB for healthy individuals ($p<0.001$). SD-OCT parameters with the greater aROCs were inferior quadrant (0.813) and average thickness (0.807). The best SAP parameter was PSD (aROC=0.915). Combining OCT and SAP data, MLCs' aROCs varied from 0.777(CTREE) to 0.946 (RAN). The best combined OCT+SAP aROC, obtained with RAN (0.946), was significantly larger than the best single OCT

parameter ($p=0.003$), but was not significantly different from the aROC obtained with the best single SAP parameter ($p=0.37$).

Conclusion: Machine learning classifiers trained with OCT and SAP data can successfully discriminate between healthy and glaucomatous eyes. The combination of OCT and SAP measurements improved the diagnostic accuracy compared with OCT data alone.

1.INTRODUÇÃO

1.1 GLAUCOMA

Glaucoma é uma neuropatia óptica progressiva que apresenta degeneração das células ganglionares retinianas e seus axônios, ocasionando alterações típicas do nervo óptico e campo visual.¹ O glaucoma primário de ângulo aberto (GPAA) é caracterizado por uma grande quantidade de parâmetros quantificáveis², incluindo a medida da pressão intraocular, a facilidade do escoamento do humor aquoso, e uma variedade de medidas geométricas do disco óptico e da camada de fibras nervosas (CFN), além de várias maneiras para testar e descrever os defeitos de campo visual. Apesar disso, muitos aspectos do glaucoma permanecem ainda em debate. O diagnóstico dessa condição pode ser muito difícil em estágios iniciais. A existência de termos como “glaucoma de pressão normal” para dano do nervo óptico glaucomatoso na ausência de pressão ocular elevada; “hipertensão ocular” para aumento da pressão intraocular e ausência de dano glaucomatoso no nervo e “glaucoma suspeito” para indivíduos com achados clínicos que indicam um grande risco para o desenvolvimento do glaucoma primário de ângulo aberto, servem para ilustrar a dificuldade no diagnóstico do glaucoma.³

O glaucoma é definido por alterações estruturais características do disco óptico acompanhados de defeitos funcionais que correspondem às áreas afetadas da camada de fibras nervosas. Em 2002, um consenso internacional publicou as definições de glaucoma de ângulo aberto e glaucoma de ângulo fechado, que são atualmente aceitas mundialmente.^{4,5} O glaucoma é reconhecido como uma neuropatia óptica, com presença, em pelo menos um olho, de alterações estruturais típicas e defeitos funcionais;⁶ o defeito estrutural é caracterizado por perdas de fibras nervosas adjacentes ao disco, que pode ser detectado tanto por exame clínico quanto por métodos de imagem, como a tomografia de

coerência óptica (OCT). O defeito funcional é avaliado pela medida da sensibilidade à luz em pontos situados dentro dos 30° do campo de visão.⁷



Figura 1: Exemplo de paciente portador de glaucoma apresentando neuropatia óptica glaucomatosa com defeito de campo visual correspondente. (Fonte: Serviço de Glaucoma da Unicamp)

1.1.2. EPIDEMIOLOGIA

Glaucoma é a mais importante causa de cegueira irreversível em todo o mundo, e a segunda causa mais comum de cegueira após catarata.^{8,9} É responsável por 14% da cegueira em todo o mundo.¹⁰ Dados de estudos populacionais indicam que, atualmente, aproximadamente 2,5% dos adultos acima de 40 anos têm glaucoma, afetando 60 milhões de pessoas, das quais 8,4 milhões estão cegas bilateralmente.⁷ Uma recente estimativa sugere que, por volta de 2020, 80 milhões de pessoas terão glaucoma, entre as quais 11 milhões serão cegas bilateralmente.¹¹

Poucos dados estão disponíveis sobre a prevalência do glaucoma na população brasileira. Um recente estudo populacional conduzido na região sul do Brasil encontrou uma prevalência de 3,4% (95% CI: 2,5 a 4,3%) numa amostra de 1636 indivíduos acima dos 40 anos. O glaucoma primário de ângulo aberto foi a forma mais comumente encontrada, com a prevalência estimada de 2,4%, que cresce acentuadamente com a idade, atingindo 4% em maiores de 60 anos.¹² Cerca de 90% dos pacientes glaucomatosos desconheciam seu diagnóstico, um número muito maior do que estimativas sobre desconhecimento da doença reportados em países desenvolvidos, que está por volta de 50%.^{13,14,15} É importante enfatizar que o estudo de prevalência do glaucoma na região sul do Brasil pode não caracterizar completamente a condição dessa doença em todo país, devido à alta heterogeneidade étnica da população brasileira.¹⁶

1.1.3. RELAÇÃO ESTRUTURA-FUNÇÃO

A relação entre alterações estruturais e funcionais no glaucoma tem sido debatida. Vários estudos demonstraram que alterações estruturais tendem a aparecer antes do desenvolvimento de defeitos de campo visual avaliados com perimetria acromática.^{17,18,19} Entretanto, com a progressão da doença, a natureza exata da relação estrutura e função se torna mais complicada. No estágio final da doença, avaliações funcionais parecem ser mais sensíveis para detectar progressão do dano do que alterações estruturais do disco óptico.²⁰

Existem vários estudos transversais que investigaram a relação entre a estrutura do disco óptico, definido por avaliação subjetiva de fotografias do disco óptico, e defeito de campo visual, definidos por alteração na perimetria acromática.²¹ Em geral, esses estudos mostraram boa correlação entre avaliação subjetiva da rima neuroretiniana, defeitos das

camadas de fibras nervosas e atrofia peripapilar com a presença, distribuição e gravidade do defeito do campo visual.²²

Em contraste, apenas alguns estudos longitudinais foram conduzidos para demonstrar a relação temporal entre dano glaucomatoso do disco óptico e defeito de campo visual.^{23,24,25} Apesar de a maioria desses estudos demonstrarem que danos estruturais frequentemente precedem defeitos de campo visual, o número de participantes era limitado.

Sommer et al.¹⁷ estudaram 1344 pacientes com hipertensão ocular seguidos por seis anos com retinografia “red-free” e perimetria (automática e manual). Oitenta e oito sujeitos apresentaram defeito de campo visual durante o seguimento. Deste grupo, 50 a 85% tinham desenvolvido defeitos na camada de fibras nervosas (CFN) antes do aparecimento de defeito de campo visual. Em 60% dos casos, o defeito da CFN precedeu o aparecimento de defeito de campo visual (CV) por 6 anos. O mesmo grupo de pesquisadores comparou 37 hipertensos oculares que converteram para glaucoma baseado no campo visual com 37 indivíduos controles obtidos de um grupo de 813 pacientes hipertensos oculares que não converteram e foram seguidos por período de 5 anos. Nesse estudo, a avaliação da camada de fibras nervosas, através de retinografia, foi mais sensível que a avaliação do disco óptico na detecção subsequente de progressão de campo visual.¹⁹

Dados do “Ocular Hypertension Treatment Study” (OHTS)²⁶ demonstraram que aproximadamente metade dos sujeitos converteram para glaucoma devido a alterações no disco óptico, enquanto a outra metade converteu devido ao aparecimento de defeitos do campo visual. Entretanto, como o diâmetro da escavação foi maior nos grupos que converteram para glaucoma, é possível que boa parte desses indivíduos já apresentassem dano ao disco óptico no momento da inclusão no estudo.

Resultados do “Early Manifest Glaucoma Trial” (EMGT) ²⁷ demonstraram que a maioria dos indivíduos (já com glaucoma) progrediram baseados em mudanças do campo visual. Progressão no CV foi definida pelo “Glaucoma Change Probability” (GCP) e a avaliação do disco óptico foi baseada em fotografias estereoscópicas. Os resultados do EMGT demonstraram que 51% dos indivíduos do grupo controle e 41% dos indivíduos do grupo tratado progrediram baseados apenas no exame de campo visual, enquanto 1% do grupo tratado e nenhum do grupo controle progrediram quando avaliados pela fotografia de disco óptico. Apenas 3% dos indivíduos tratados e 11% do grupo controle progrediram em ambas análises. Em contraste com outros estudos, o resultado do EMGT parece indicar que, em indivíduos com glaucoma inicial, progressão de defeito de campo visual pode ocorrer mais frequentemente que progressão estrutural.

Recentemente, Jonhson et al¹⁸ avaliaram a significância prognóstica das mudanças estruturais do nervo óptico no desenvolvimento de defeito de campo visual em hipertensos oculares utilizando dados longitudinais coletados em 2 centros, que totalizavam 479 olhos de 295 pacientes com seguimento médio de aproximadamente 6 anos. A incidência de conversão do campo visual foi significativamente maior em indivíduos que já apresentavam alterações sugestivas do disco óptico no exame inicial, sugerindo que dano glaucomatoso estrutural precede o desenvolvimento de defeito de campo visual.

1.2. CAMPO VISUAL

A perimetria automatizada acromática (PAA), que utiliza um estímulo branco num fundo branco (branco no branco), tem sido usada por mais de duas décadas na prática clínica para pesquisar o campo visual dos pacientes. Apesar do uso tanto para diagnóstico quanto para determinação de progressão do glaucoma, esse método é pouco sensível para detectar perda de células ganglionares retinianas em estágios iniciais da doença.^{28,29}

A PAA se baseia na determinação do limiar de sensibilidade para cada ponto do campo visual. O limiar de sensibilidade é o estímulo de intensidade mínima visto pelo indivíduo examinado naquele ponto. A estimativa do limiar de sensibilidade é fornecida em unidade chamada decibél (dB), que corresponde à unidade de atenuação do estímulo luminoso. Decibel é uma unidade logarítmica que indica a proporção de intensidade em relação a um nível especificado ou implícito. Uma relação em dB é igual a dez vezes o logaritmo de base 10 da razão entre duas quantidades de energia. Portanto, quanto maior é o número registrado em dB, mais fraco foi o estímulo visto e melhor é a sensibilidade retiniana nesta área.³⁰

O perímetro Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc., Dublin, CA) possui, entre outros programas, a estratégia 24-2 Swedish Interactive Threshold Algorithm (SITA) que fornece vários parâmetros, incluindo gráficos e índices, que analisam dados a respeito do campo visual do paciente (Figura 2), tais como o MD, PSD e GHT.

1.2.1 MD (“Mean Deviation”)

O MD fornece uma ideia da sensibilidade geral do campo visual do paciente. Ele representa a média ponderada dos valores mostrados no gráfico “Total Deviation”. O

gráfico “Total Deviation” apresenta as diferenças, em decibéis, entre os valores obtidos no exame e os valores normais corrigidos para a idade. O MD pode ser negativo, indicando uma sensibilidade geral abaixo da média esperada para a idade, ou positivo, indicando sensibilidade acima da média.³¹

1.2.2. PSD (“Pattern Standard Deviation”)

O PSD é utilizado para dimensionar a profundidade dos defeitos eventualmente presentes. O índice PSD corresponde ao desvio-padrão da média das diferenças entre cada valor da sensibilidade encontrado no exame e o valor esperado para o próprio paciente. Em outras palavras, ele representa o desvio-padrão da média dos valores dispostos no gráfico numérico do “Pattern Deviation”. O gráfico “Pattern Deviation” deriva do gráfico “Total Deviation”, retirando o efeito da depressão difusa (como a causada por opacidades de meios), ou mesmo eliminando a perda homogênea do defeito, tornando mais aparentes os defeitos localizados ou mais profundos.³²

O PSD apresenta-se elevado quando alguns setores do campo visual sofreram uma maior alteração de sensibilidade do que em outros, como ocorre com os escotomas glaucomatosos. Assim, o PSD é um indicador de defeitos localizados no campo visual.³²

1.2.3. GHT (Glaucoma Hemifield Test)

O GHT é um teste que compara as metades superior e inferior do campo visual central. A utilização deste teste se baseia em uma característica da lesão glaucomatosa do nervo óptico, ou seja, no fato do acometimento dos feixes de fibras nervosas geralmente ser

assimétrico, levando a um acometimento de um hemisfério (superior ou inferior) antes que o outro.³¹

A comparação é feita entre cinco zonas simétricas dos hemisférios superior e inferior. Um escore é atribuído a cada zona com base no mapa de probabilidades do “Pattern Deviation”, e então, é feita uma comparação de cada zona superior com a zona inferior correspondente, e a diferença entre os escores das duas zonas é comparada com a existente na população normal. De acordo com esta análise, o GHT fornece cinco resultados possíveis:

- “Outside normal limits”: Esta mensagem é fornecida quando a diferença em pelo menos um par de zonas excede daquela encontrada em 99% da população; ou quando os escores individuais em ambos os pares de qualquer zona excedem os encontrados em 99,5% da população normal.
- “Borderline”: A diferença entre os escores, em pelo menos um par de zonas, excede a encontrada em 97% dos indivíduos normais.
- “General reduction of sensitivity”: Esta mensagem aparece quando nenhuma das condições para “outside normal limits” é satisfeita e há uma diminuição generalizada na sensibilidade observada em menos que 0,5% dos indivíduos normais.
- “Abnormally high sensitivity”: Esta mensagem aparece quando a sensibilidade global é maior que a encontrada em 99,5% da população normal. Quando isto ocorre, não é feita comparação entre as zonas correspondentes. Costuma ocorrer em exames com elevado número de falsos-positivos.
- “Within normal limits”: Esta mensagem aparece quando nenhuma das condições acima é satisfeita.

1.2.4 Estratégia SITA

A estratégia SITA foi desenvolvida para otimizar a resposta do paciente durante o exame, diminuindo o tempo de duração do teste e por conseguinte, tornando-o mais confiável. Está disponível no perímetro Humphrey Field Analyzer II (Humphrey Systems, Dublin, CA) e é baseada no procedimento de previsão (“Forecasting”) empregando estatística bayesiana.³²

Existem vários estudos a respeito da estratégia SITA. Sekhar e colaboradores³³ confirmaram que a estratégia SITA Standard e SITA Fast, produziam resultados similares àqueles obtidos pela estratégia “Full Threshold”, apesar de o tempo médio gasto durante o exame ter sido aproximadamente 50% menos com o SITA Standard e 70% menos com o SITA Fast. A confiabilidade teste e reteste foi excelente com o SITA Standard, porém foi pior com o SITA Fast. Está claro que a estratégia SITA trouxe um impacto positivo no uso dos testes perimétricos automatizados e, atualmente, se tornou o exame padrão (“gold standard”) para avaliação funcional no glaucoma.

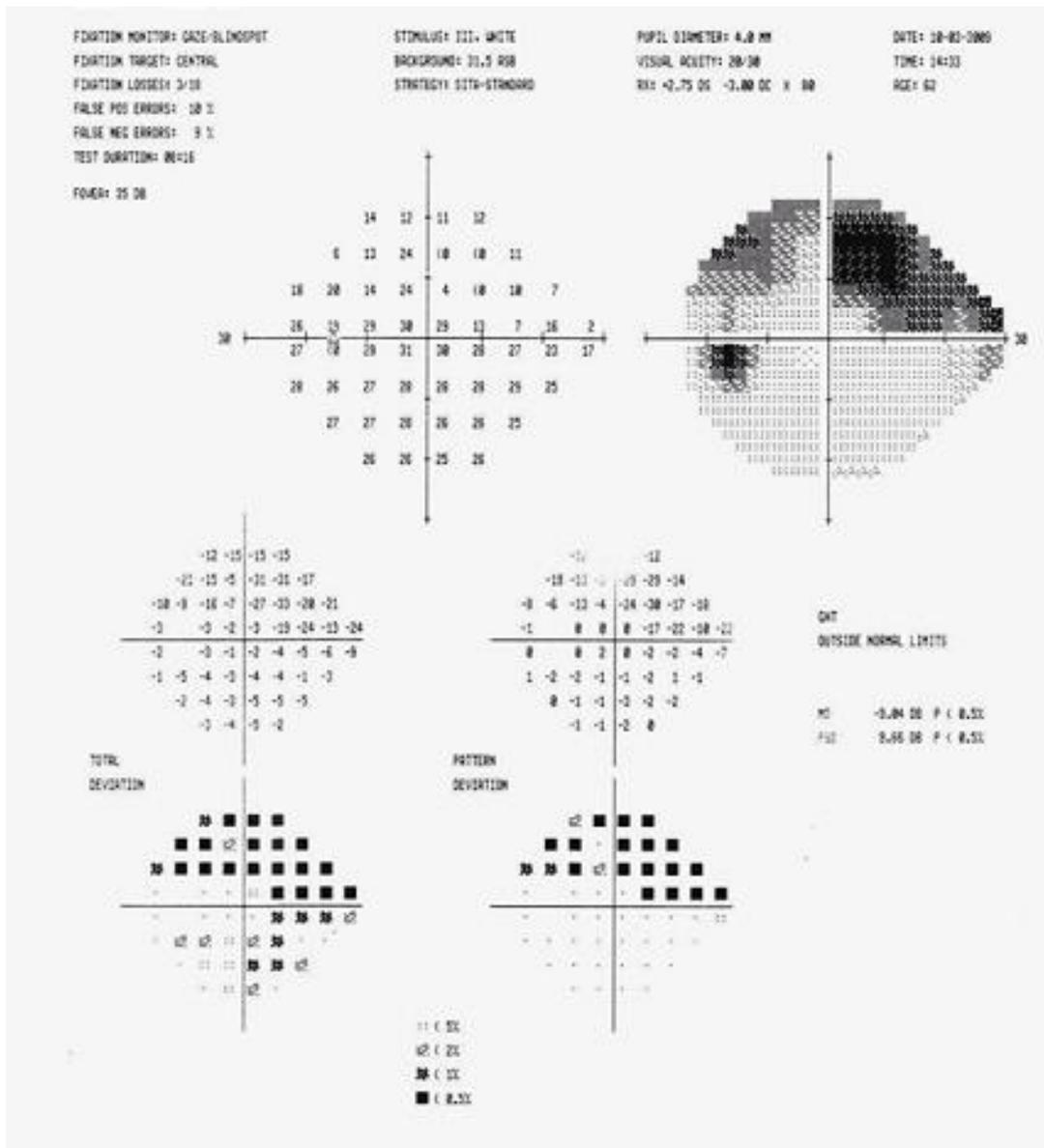


Figura 2: Exemplo de campo visual utilizando estratégia SITA de paciente com defeito arqueado em hemicampo superior em olho esquerdo. (Fonte: Serviço de Glaucoma da Unicamp)

1.3. OCT

As primeiras descrições de imagens unidimensionais (A-SCAN) usando a medição do tempo de atraso do eco de luz e a magnitude da luz refletida ou refratada em uma estrutura (o princípio do OCT) foram publicadas na década de 1980.^{34,35,36} Huang³⁷ e colaboradores demonstraram, “in vitro”, as primeiras imagens retinianas de OCT (imagens bidimensionais ou seccionais, semelhantes à ultrassonografia em modo b) em 1991. Após dois anos, as primeiras imagens de OCT de retinas humanas, “in vivo”, foram capturadas.^{38,39}

A tomografia de coerência óptica tornou-se disponível comercialmente para a oftalmologia em 1996 (Zeiss OCT), mas a sua adoção como ferramenta diagnóstica prosseguiu lenta. Uma segunda geração de instrumentos de OCT (Zeiss OCT 2) começou a ser comercializada em 2000, porém apenas após a introdução da terceira geração do aparelho (Zeiss Stratus OCT), em 2002, essa tecnologia ganhou ampla aceitação na área oftalmológica.⁴⁰

Recentemente, várias companhias desenvolveram novas versões de OCT (4ª geração), entre elas o Cirrus High Definition-HD (Karl Zeiss Med Tec) que usa a tecnologia de domínio espectral. Esses novos OCTs têm como principal característica maior resolução axial e maior velocidade de “Scan” que o OCT convencional de domínio temporal (Stratus OCT).⁴⁰

A obtenção de imagens pela OCT baseia-se na técnica de interferometria de baixa coerência, na qual o interferômetro óptico usa o princípio de comparação entre diferentes feixes de luz, o que possibilita medição de alta resolução de estruturas, correspondente a medição de tempo de atraso da escala de fentossegundos. Um dos feixes é refletido e o

outro é transmitido ao olho, sendo então refletido por estruturas oculares a partir de distâncias axiais variadas. O feixe de luz refletido a partir do olho consiste de múltiplos ecos que informam sobre a distância e espessura das diferentes estruturas intraoculares. O segundo feixe é refletido a partir de um espelho de referência situado a uma distância conhecida. A chave do processo consiste no fato de que o interferômetro pode medir o tempo de atraso de ecos ópticos pela comparação do feixe de luz refletido com o feixe de luz padrão de referência.⁴⁰

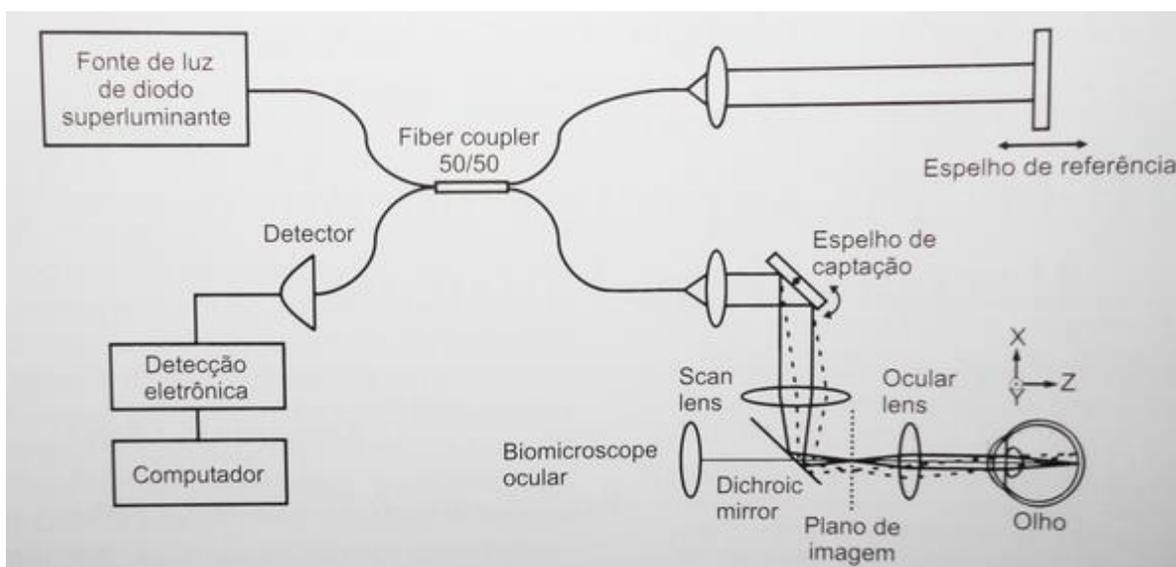


Figura 3: Desenho esquemático do sistema TD-OCT. A fonte de luz é um diodo superluminescente de baixa coerência acoplado a uma fibra óptica. O espelho de referência (canto superior direito) é movimentado para produzir diferentes tempos de atraso (referência) que são comparados com o feixe de luz refletido do olho (canto inferior direito).⁴¹

Diferente do sistema de detecção “time domain” (TD-OCT), em que o cálculo do tempo da luz refletida necessitava de um espelho de referência móvel, na detecção do tipo

“spectral domain” a posição do espelho de referência não necessita ser ajustada. A imagem é adquirida usando espectro de interferência dos atrasos dos ecos da luz, medidos por um espectrômetro e uma câmara CCD (do inglês “charge-coupled device”) de alta velocidade. Como o espectro de interferência é composto por oscilações cujas frequências são proporcionais ao tempo de atraso do eco da luz, a medição da imagem axial pode ser obtida usando cálculo de transformação de Fourier (uma operação matemática que extrai a frequência contida em um sinal). Isso tornou possível a detecção de todos os sinais ópticos (A-Scan) simultaneamente, em contraste com a detecção “time domain”, que os detecta sequencialmente, levando a uma melhora significativa na sensibilidade e na velocidade de captura de imagem.⁴¹

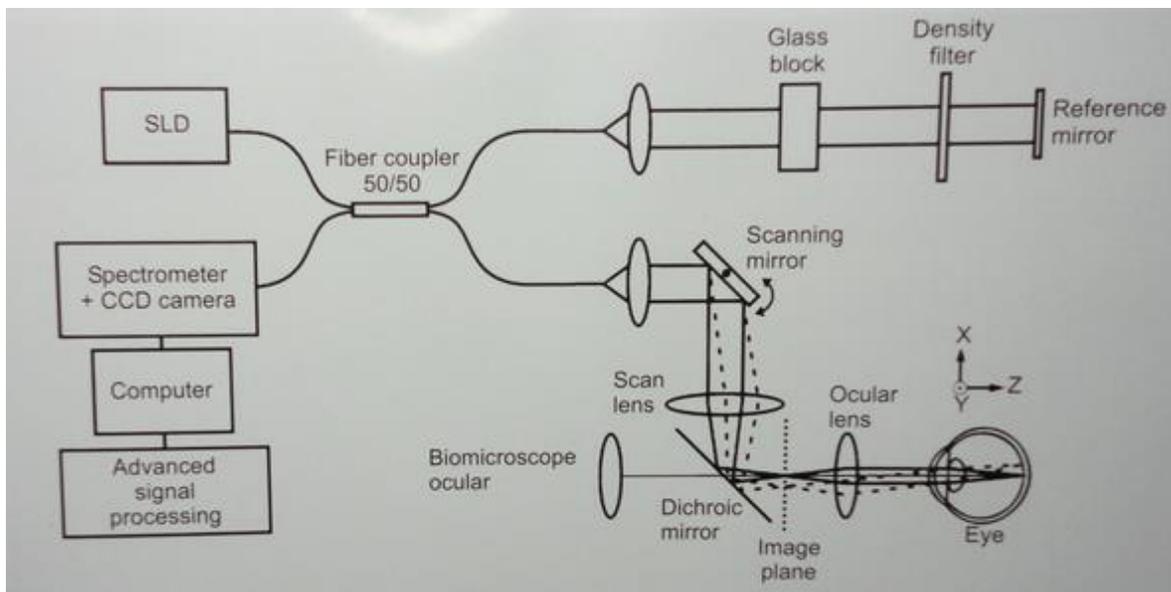


Figura 4: Desenho esquemático do sistema SD-OCT. A fonte de luz é um diodo superluminescente de baixa coerência (SLD) (canto superior esquerdo) acoplado a uma fibra óptica e conectado ao espectrômetro + câmera CCD (canto inferior esquerdo). O espelho de referência (canto superior direito) permanece fixo.⁴¹

Além da maior velocidade de aquisição de imagens, a fonte de luz utilizada no sistema SD-OCT também melhorou a resolução axial quando comparada ao TD-OCT. A resolução axial dos sistemas de OCT é governada pelo comprimento da coerência da fonte de luz, e é inversamente proporcional à largura de banda. Os sistemas de OCT comerciais utilizam fontes de luz de baixa coerência com comprimento de onda próximo ao infravermelho (820-840 nm). O Stratus OCT (TD-OCT) utiliza uma largura de banda de aproximadamente 25 nm e atinge uma resolução axial de 8 a 10 micras em tecidos. O Cirrus OCT (SD-OCT) utiliza uma largura de banda de aproximadamente 50 nm e atinge 5 micras de resolução axial.⁴¹

1.3.1 OCT Cirrus

O aparelho de imagem comercialmente disponível Cirrus HD-OCT (software version 3.0, Carl Zeiss Meditec, Inc.), empregado neste estudo, usa a tecnologia de domínio espectral. Ele apresenta vários programas para análise da CFN da retina e peripapilar, entre eles, o programa “optic disc cube 200x200”, que foi utilizado para coleta de dados do estudo. Este programa consiste em 200 A-scans derivados de 200 B-scans cobrindo uma área peripapilar de 6mm^2 centrados no disco óptico. Após a criação do mapa da espessura das camadas do cubo tridimensional, o software determina automaticamente o centro do disco e extrai um círculo peripapilar (com raio de 1.73mm) para medida da espessura da camada de fibras nervosas.

A estratégia “optic disc cube 200x200” fornece:

1: valores dos parâmetros (espessura da CFN global, simetria das CFN, área da rima neural, área do disco óptico, média da proporção escavação/disco, proporção da

escavação/disco vertical, volume da escavação) e comparação com banco de dados do aparelho.

2: “RNFL Thickness Map”: exibição topográfica da CFN.

3: “RNFL Deviation Map”: comparação topográfica da CFN com banco de dados.

4: “Neuro-retinal Rim thickness”: espessura da rima neuro-retiniana e sua comparação com dados normativos.

5: “RNFL TSNIT graph”: medida da espessura da CFN em 360⁰ peripapilar e comparação com dados normativos.

6: “RNFL Quadrant e Clock Hour”: valores médios da espessura da CFN peripapilar localizadas nos 4 quadrantes e separadas em 12 regiões (horas do relógio).

7: “Horizontal and vertical b-Scans”: imagem de corte seccional nos sentidos horizontal e vertical.

8: “RNFL calculation circle”: imagem extraída do círculo peripapillar que mede a espessura da CFN.

Para cada parâmetro o aparelho fornece uma classificação baseada na comparação com dados normativos internos do aparelho. Um parâmetro é classificado como fora dos limites de normalidade (representado pela cor vermelha) se seus valores estiverem abaixo de 99% do intervalo de confiança (IC) dos indivíduos normais da idade correspondente. O resultado “limítrofe” (representado pela cor amarela) indica que os valores estão entre 95 e 99% do intervalo de confiança e o resultado “dentro dos limites de normalidade” (representado pela cor verde) indica que os valores estão dentro do IC de 95 %.

O programa “optic disc cube 200x200” fornece dados sobre a qualidade da imagem capturada, ou seja, o valor da força do sinal (“Signal Strenght Value”). Este parâmetro

fornece valores de 0 a 10, onde 10 representa o valor máximo da força do sinal e 0 o valor para imagens com sinal muito fraco. Valores acima de 6 indicam boa qualidade do sinal da imagem adquirida.⁴²

O banco de dados do Cirrus difere do Stratus OCT,⁴³ seu antecessor, pois apresenta maior proporção de indivíduos nos grupos estratificados de acordo com a etnia (Tabela 1). Foram incluídos no banco de dados do Cirrus, para análise da CFN, 284 indivíduos (idade entre 19 e 84 anos) de 7 localidades diferentes. A distribuição em relação ao gênero foi semelhante entre os sexos (134 indivíduos do sexo masculino e 150 do sexo feminino). A distribuição em relação a etnia apresenta: 43% caucasianos, 24% asiáticos, 18% afro-americanos, 12% hispânicos, 1% indiano, e 6% de etnia mista. O banco de dados é ajustado pela idade, e não é ajustado em relação ao comprimento axial, refração, área do disco óptico, força do sinal ou outro parâmetro.⁴⁴

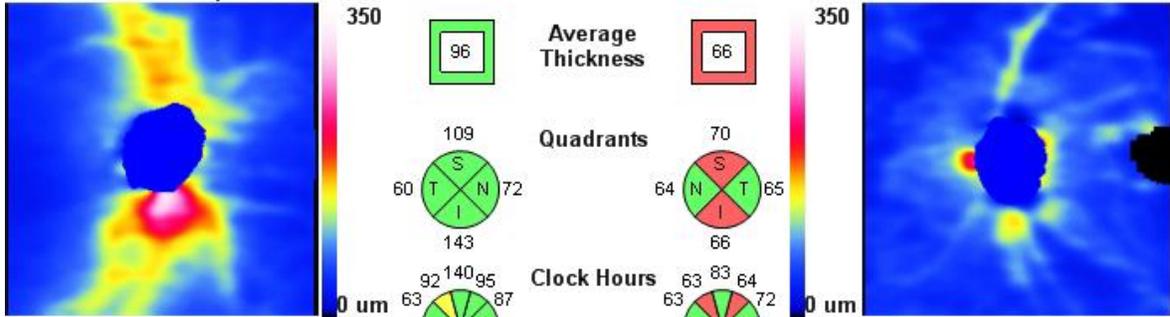
Tabela 1: Porcentagem dos indivíduos incluídos nos bancos de dados do Cirrus OCT e Stratus OCT de acordo com a etnia.

| Etnia | Cirrus (284 indivíduos) | Stratus (328 indivíduos) |
|-----------------|-------------------------|--------------------------|
| Caucasianos | 43% | 63% |
| Asiáticos | 24% | 3% |
| Afro-americanos | 18% | 8% |
| Hispânicos | 12% | 24% |
| Indiano | 1% | 0% |
| Etnia mista | 6% | 0% |

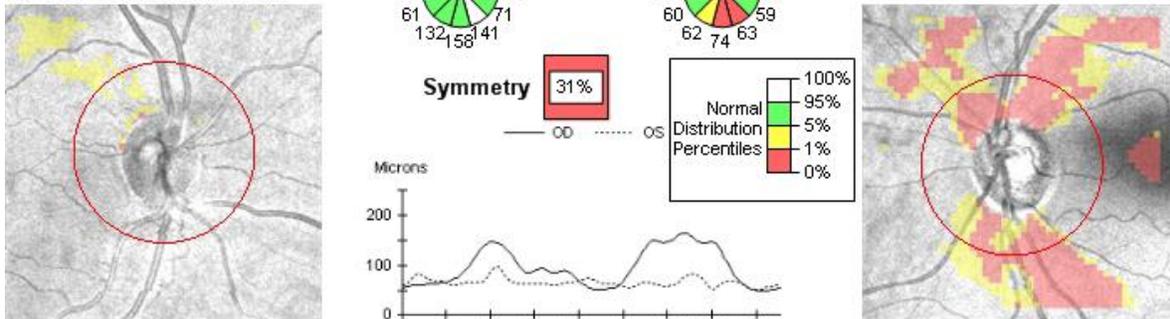
Optic Disc Cube 200x200

OD ● | ● OS

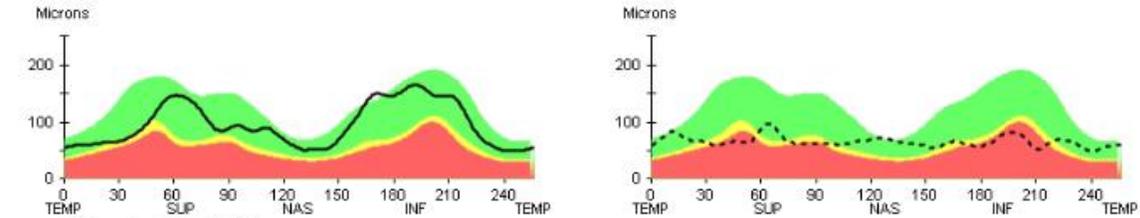
RNFL Thickness Map



RNFL Thickness Deviation



RNFL TSNIT Normative Data



Extracted RNFL Tomogram

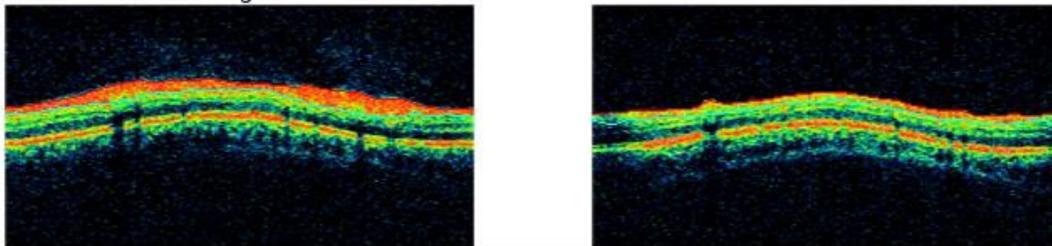


Figura 5: parâmetros avaliados pelo Cirrus OCT “optic disc cube 200x200”. (Fonte: Serviço de Glaucoma da Unicamp)

1.3.2. REPRODUTIBILIDADE DO SD-OCT

A reprodutibilidade é um importante fator a ser avaliado em aparelhos que se propõem a realizar diagnóstico e avaliar progressão. Vários estudos investigaram a reprodutibilidade da medida da espessura das camadas de fibras nervosas do SD-OCT Cirrus.^{45,46,47,48,49} De acordo com o resultado desses estudos, os aparelhos SD-OCTs comercialmente disponíveis apresentaram boa reprodutibilidade e se mostraram comparáveis ou até mesmo melhores que o TD-OCT. A reprodutibilidade pode ser avaliada pelo coeficiente de variação (COV) e coeficiente de correlação intraclasse (CCI). Cremasco et al.⁴⁹ encontraram elevado CCI e baixo COV para avaliação intra e entre seções e para avaliação entre examinadores. Apresentaram as seguintes variações do ICC e COV: ICC=0,82-0,99 e COV=3,68%-10,50% para avaliação entre seções; ICC=0,74-0,99 e COV=2,92-11,84% para avaliação intra seção; e ICC=0,55-0,98 e COV=2,04-14,97% para avaliação entre examinadores.

Medidas setoriais da CFN mostraram maior variabilidade que medidas gerais.⁵⁰ Medidas setoriais são mais facilmente afetadas pela inconsistência da amostra do círculo peripapilar e outros fatores de confusão. Os resultados sugerem que o SD-OCT pode atingir níveis melhores de reprodutibilidade do que a versão anterior de TD-OCT, especialmente em análise setorial. Melhor reprodutibilidade da medida de camadas de fibras nervosas setoriais do SD-OCT comparadas com TD-OCT podem ser explicadas pela melhora da resolução e da tecnologia do registro dos dados do SD-OCT.^{46,47}

1.3.3. CAPACIDADE DIAGNÓSTICA DO SD-OCT

Vários estudos já foram publicados para avaliar a capacidade diagnóstica de glaucoma pelo SD-OCT.^{51,52,53,54,55} A maioria desses estudos compararam a capacidade diagnóstica das medidas das CFN do SD-OCT com aquelas obtidas pelo TD-OCT, usando área abaixo da curva Roc (aROC) para discriminação entre sadios e olhos glaucomatosos. É difícil comparar valores das aROCs obtidas em diferentes estudos, pois valores de aROCs podem variar de acordo com o estágio da doença com que os participantes se apresentam, e com os critérios de inclusão. Entretanto, os aparelhos SD-OCTs testados mostraram boa capacidade diagnóstica para o glaucoma. Adicionalmente, a maioria dos estudos consistentemente mostrou não haver diferença estatisticamente significativa entre SD-OCT e TD-OCT.^{46, 51, 52, 53, 54, 55} (Quadro 1)

Quadro 1: Estudos que avaliaram a capacidade diagnóstica do Cirrus OCT e compararam com o Stratus OCT.

| Autores | População | Resultados |
|--------------------------------------|--|---|
| Leung et al. ⁴⁶ | 116 normais e 116 glaucomatosos | As aROCs foram de 0,962 do Cirrus e 0,956 do Stratus (P>0,05). |
| Park et al. ⁵³ | 74 normais e 100 glaucomatosos | As aROCs foram de 0,953 do Cirrus e 0,934 do Stratus (P=0,15). O cirrus apresentou melhor performance em glaucoma iniciais. |
| Moreno-Montañés et al. ³⁴ | 130 normais e 86 glaucomatosos | As aROCs foram de 0,837 do Cirrus e 0,829 do Stratus (P=0,70). |
| Jeung et al. ⁵⁵ | 55 normais e 55 com glaucoma pré-perimétrico | As aROCs foram de 0,728 do Cirrus e 0,760 do Stratus (P=0,47). |
| Sung et al. ⁵¹ | 60 normais, 48 com glaucoma suspeito e 55 glaucomatosos | Cirrus apresentou maior Se (63,6%) que o Stratus (40,0%) utilizando a classificação normativa dos aparelhos. |
| Chang et al. ⁵² | 50 normais e 54 glaucomatosos | A Se e Sp do Cirrus (65% e 100%, respectivamente) foram semelhantes às Se e Sp do Stratus (61% e 100%, respectivamente) utilizando a classificação normativa dos aparelhos (parâmetro “average thickness” utilizando a classificação normativa dos aparelhos). |
| Huang et al. ⁵⁶ | 62 normais, 47 glaucomatos estágio inicial e 30 glaucoma pré-perimétrico | A aROC do Cirrus (0,951) foi maior que Stratus (0,881) (P=0,006) em indivíduos com glaucoma pré-perimétrico. O desempenho foi semelhante entre os aparelhos em indivíduos com glaucoma inicial. |

aROC= área abaixo da curva ROC; Se= sensibilidade; Sp= especificidade.

1.4. CONCEITOS SOBRE SISTEMAS DE APRENDIZAGEM DE MÁQUINA

A medicina moderna é afrontada com o desafio de adquirir, analisar e aplicar uma vasta quantidade de conhecimento necessária para resolver problemas clínicos complexos. O desenvolvimento da inteligência artificial (IA) médica está relacionado com o desenvolvimento de programas destinados a ajudar os médicos na formulação de hipóteses diagnósticas, nas decisões terapêuticas e na previsão do desenlace da doença. Eles são desenvolvidos para ajudar os trabalhadores na área de saúde em suas atividades diárias, assistindo-os em tarefas que requerem a manipulação de dados e conhecimentos.⁵⁷

A aprendizagem de máquina é um subcampo da inteligência artificial dedicada ao desenvolvimento de algoritmos e técnicas que permitem ao computador adquirir conhecimento de forma automática, aperfeiçoando seu desempenho em uma determinada tarefa. Segundo Ritch et al.,⁵⁸ o processo de aprendizagem incluiria “mudanças adaptáveis no sistema, que permitem que este faça, da próxima vez, a mesma tarefa ou tarefas, tiradas do mesmo grupo, com mais eficiência e eficácia”. Um sistema de aprendizagem tem a função de analisar informações e generalizá-las, para extração de novos conhecimentos. Para isso, usa-se um programa de computador para automatizar o aprendizado.⁵⁹

O aprendizado utiliza o princípio da indução (inferência lógica) com intuito de obter conclusões genéricas a partir de um conjunto de exemplos. Um conceito é aprendido efetuando-se inferência indutiva sobre os exemplos apresentados. O objetivo do algoritmo de indução é construir um classificador que possa determinar a classe a que um exemplo não rotulado pertence. É possível rotular um novo exemplo devido à generalização.⁵⁹

O aprendizado indutivo pode ser dividido em supervisionado e não supervisionado. No aprendizado supervisionado, fornece-se ao algoritmo de aprendizagem, ou indutor, um

conjunto de exemplos de treinamento onde os rótulos das classes associadas são conhecidos. Em geral, cada exemplo é descrito por um vetor de valores de características, ou atributos, e o rótulo da classe associada. O objetivo do algoritmo de indução é construir um classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados. Para rótulos de classe categóricos (não numéricos), esse problema é conhecido como classificação e para valores contínuos (numéricos) como regressão. No aprendizado não supervisionado, o indutor analisa os exemplos fornecidos e tenta determinar se alguns deles podem ser agrupados de alguma maneira formando agrupamentos ou se existe associação. Agrupamentos têm como objetivo relacionar instâncias com características em comum. Associações visam encontrar regras que procuram relações presentes no conjunto de dados. Após a determinação dos agrupamentos, normalmente, é necessário uma análise para determinar o que cada agrupamento significa no contexto do problema que está sendo resolvido.⁶⁰

Atualmente existem diversos “softwares” desenvolvidos para mineração dos dados, entre eles o WEKA (“Waikato Environment for Knowledge Analysis”),⁶¹ que é amplamente utilizado e aceito, e oferece uma grande variedade de ferramentas úteis para o aprendizado de máquinas.⁶²

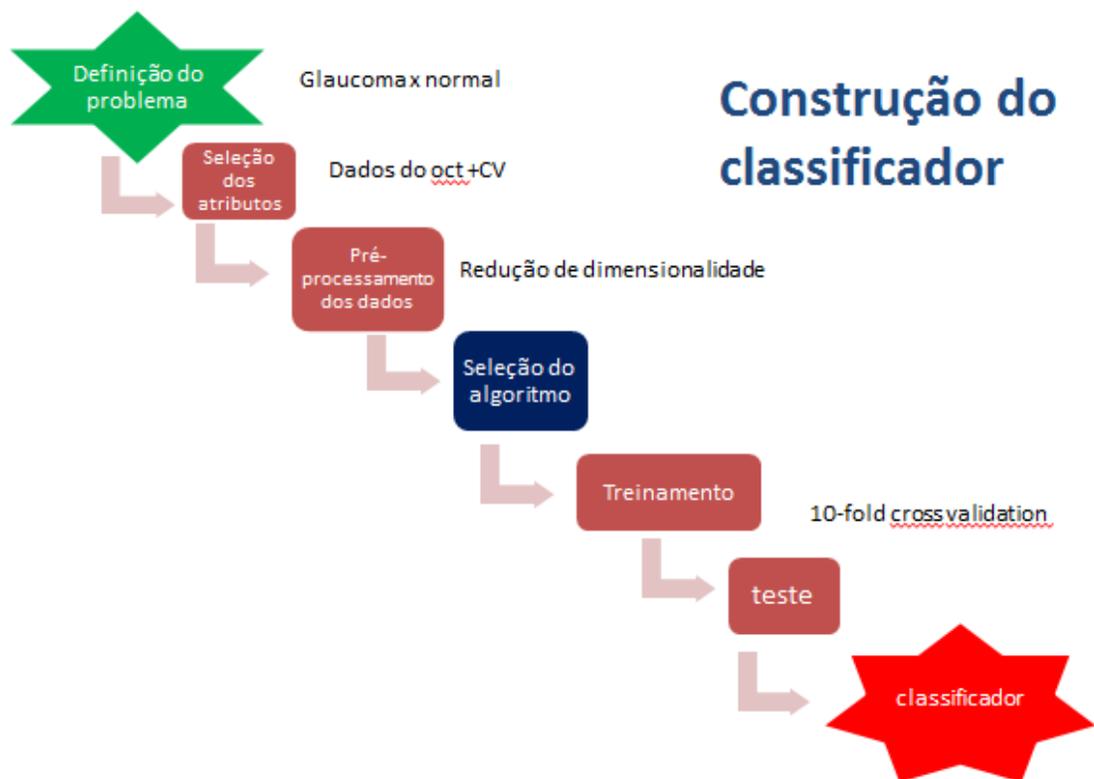


Figura 6: Fluxograma para construção de um classificador. (Fonte: Fabrício Reis da Silva)

A construção de algoritmos de aprendizagem de máquina supervisionados é descrito como na Figura 6. O primeiro passo é a coleta de dados para um determinado problema. Se há à disposição um especialista, ele pode sugerir quais dados (atributos ou parâmetros) são mais informativos. Atributo ou parâmetro se referem a alguma característica, ou aspecto, de um determinado exemplo.⁵⁹ Se não há este profissional, então o método mais simples é aquele chamado de força bruta, que significa medir tudo o que está disponível na esperança que os parâmetros relevantes possam ser isolados. Entretanto, a coleta de dados pelo método de força bruta não é o mais apropriado para aprendizagem de máquina pela técnica de indução.⁶³

O segundo passo é a preparação de dados e pré-processamento. A seleção de atributos não é apenas usada para lidar com ruídos (dados imperfeitos devido ao processo de geração, aquisição ou transformação dos dados ou mesmo devido a classes rotuladas incorretamente), mas para também lidar com uma amostra de dados muito grande. A seleção de atributos serve para otimização do problema, mantendo a qualidade da mineração, ou classificação, enquanto minimiza o tamanho da amostra. A seleção de atributos é o processo que identifica e remove atributos irrelevantes e redundantes. Este passo reduz a dimensionalidade dos dados e torna possível ao algoritmo operar mais rápido e efetivamente. Além disso, a descoberta dos principais atributos contribui para melhorar a compreensibilidade da produção do classificador e o melhor entendimento do conceito de aprendizagem.⁶³

A escolha de qual algoritmo será usado é um passo crítico. Não existe um único algoritmo que apresente o melhor desempenho para todos os problemas.⁵⁹ A avaliação do classificador é geralmente baseada na predição da acurácia (a porcentagem das predições corretas dividida pelo total de número de predições). Existem várias técnicas utilizadas para calcular a acurácia dos classificadores, sendo uma das mais utilizadas a validação cruzada.⁶³ Após a escolha do algoritmo, deve-se fazer o treinamento e testá-lo para verificar a sua capacidade de classificação.

1.5. CLASSIFICADORES DE APRENDIZAGEM DE MÁQUINA

1.5.1.ÁRVORE DE DECISÃO (CTREE)

O algoritmo de árvore de decisão gera uma estrutura de árvore que ajuda na classificação e predição de amostras desconhecidas. O algoritmo classifica as instâncias

baseado nos valores dos atributos, ou seja, com base nos registros do conjunto de treinamento uma árvore é montada e, a partir desta árvore, pode-se classificar a amostra desconhecida.⁶⁴

A árvore é composta de um conjunto de nós que são conectados por ramificações. Basicamente, existem três tipos de nós: um nó raiz, que inicia a árvore; os nós comuns, que dividem um determinado atributo e geram ramificações e os nós folha, que contém as informações de classificação do algoritmo. Já as ramificações possuem todos os valores possíveis do atributo indicado no nó.⁶⁵

A ideia do algoritmo é montar uma árvore onde cada nó indica um teste do atributo. Os atributos escolhidos para os nós da árvore são chamados de atributos divisores ou atributos teste. A escolha dos atributos é feita com base no maior ganho de informação, isto é, na qualidade de classificação do atributo. Deste modo, podemos dizer que o atributo que melhor classificar os dados deve ser escolhido como o nó raiz da árvore. As instâncias percorrem o trajeto desde o nó raiz até os nós folhas e são classificadas através dos valores dos seus atributos.⁶⁵

As árvores de decisão são similares à regra *se-então* (“if-then”), e por isso, o mecanismo de seu funcionamento é facilmente compreensível. O aspecto que mais consome tempo na construção da árvore é a classificação de instâncias que contém atributos numéricos, logo, a melhor maneira de otimizar é através da estratificação dos dados. Isso é realizado através de atribuição (seleção) de limiares dos atributos que são determinados apenas uma vez, convertendo efetivamente os atributos numéricos em intervalos discretos.⁶³

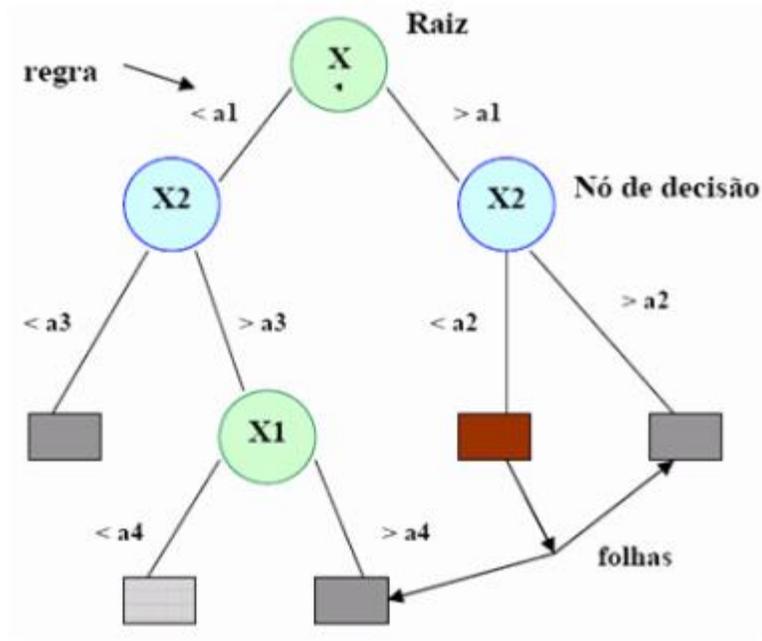


Figura 7: Representação esquemática de uma árvore de decisão.⁶⁶ Os círculos representam os nós e os retângulos representam as folhas sendo o círculo superior o nó raiz. Cada nó de decisão contém um teste para o atributo (<a ou >a, negativo ou positivo para o teste do nó em questão (exemplos: a1, a2, a3,...)). Cada folha (retângulo) está associada à uma classe; o percurso na árvore (da raiz à folha) corresponde a uma regra de classificação.⁶⁶

Segue abaixo exemplo de árvore de decisão:⁶⁵

Vamos considerar o seguinte cenário para a utilização do algoritmo: um sistema de contas a receber de um clube esportivo envia para um banco no início de cada mês um boleto contendo a mensalidade do clube a ser paga pelos associados. O banco então envia pelo correio a fatura para os clientes e espera os recebimentos. No final do mês, o banco insere no sistema do clube quais clientes pagaram o boleto, quais não pagaram e quais pagaram com atraso, dentre outras informações. Com o objetivo de diminuir a quantidade de clientes que pagam o boleto com atraso, decide-se realizar uma mineração de dados na base de associados para identificar o perfil de quem paga com atraso. Um pré-processamento de dados separou as informações dos clientes em alguns atributos que podem ser visualizados na Tabela 2.

Tabela 2: Atributos (parâmetros) para geração de um exemplo de árvore de decisão baseada nos dados dos associados do clube (idade, salário, ensino superior, dependentes, pagamento).

| IDADE | SALARIO | SUPERIOR_COMPLETO | DEPENDENTES | ATRASOU |
|--------|---------|-------------------|-------------|---------|
| <=30 | Alto | Não | Não | Não |
| <=30 | Alto | Não | Sim | Não |
| 31..40 | Alto | Não | Não | Sim |
| >40 | Médio | Não | Não | Sim |
| >40 | Baixo | Sim | Não | Sim |
| >40 | Baixo | Sim | Sim | Não |
| 31..40 | Baixo | Sim | Sim | Sim |
| <=30 | Médio | Não | Não | Não |
| <=30 | Baixo | Sim | Sim | Sim |
| >40 | Médio | Sim | Não | Sim |
| <=30 | Médio | Sim | Sim | Sim |
| 31..40 | Médio | Não | Sim | Sim |
| 31..40 | Alto | Sim | Não | Sim |
| >40 | Médio | Não | Sim | Não |

Ao montar uma árvore de decisão, o primeiro passo é gerar o nó raiz. Para isso, deve-se primeiro calcular a probabilidade para cada um dos valores do atributo de classificação, que no exemplo é o atributo ATRASOU. As probabilidades do nó raiz são:

Probabilidade de ATRASOU=NÃO: $5/14 = 0,36$

Probabilidade de ATRASOU=SIM: $9/14 = 0,64$

Este nó ainda não tem um atributo e por isso é um nó que pode ser dividido.

O próximo passo é escolher atributos para os nós que podem ser divididos. Como só temos um nó, vamos analisar todos os atributos para verificar aquele que melhor classifica os dados.

- Para o atributo IDADE:

Valor ≤ 30

Probabilidade de ATRASOU = NÃO: $3/14 = 0,214$

Probabilidade de ATRASOU = SIM: $2/14 = 0,143$

Valor 31-40

Probabilidade de ATRASOU = NÃO: $0/14 = 0$

Probabilidade de ATRASOU = SIM: $4/14 = 0,286$

Valor > 40

Probabilidade de ATRASOU = NÃO: $2/14 = 0,143$

Probabilidade de ATRASOU = SIM: $3/14 = 0,214$

- Para o atributo SALÁRIO:

Valor alto

Probabilidade de ATRASOU = NÃO: $2/14 = 0,143$

Probabilidade de ATRASOU = SIM: $2/14 = 0,143$

Valor médio

Probabilidade de ATRASOU=NÃO: $2/14 = 0,143$

Probabilidade de ATRASOU=SIM: $4/14 = 0,286$

Valor baixo

Probabilidade de ATRASOU=NÃO: $1/14 = 0,071$

Probabilidade de ATRASOU=SIM: $3/14 = 0,214$

- Para o atributo SUPERIOR_COMPLETO :

Valor Não

Probabilidade de ATRASOU=NÃO: $4/14 = 0,286$

Probabilidade de ATRASOU=SIM: $3/14 = 0,214$

Valor Sim

Probabilidade de ATRASOU=NÃO: $6/14 = 0,428$

Probabilidade de ATRASOU=SIM: $1/14 = 0,071$

- Para o atributo DEPENDENTES :

Valor Não

Probabilidade de ATRASOU=NÃO: $2/14 = 0,143$

Probabilidade de ATRASOU=SIM: $5/14 = 0,357$

Valor Sim

Probabilidade de ATRASOU=NÃO: $3/14 = 0,214$

Probabilidade de ATRASOU=SIM: $4/14 = 0,286$

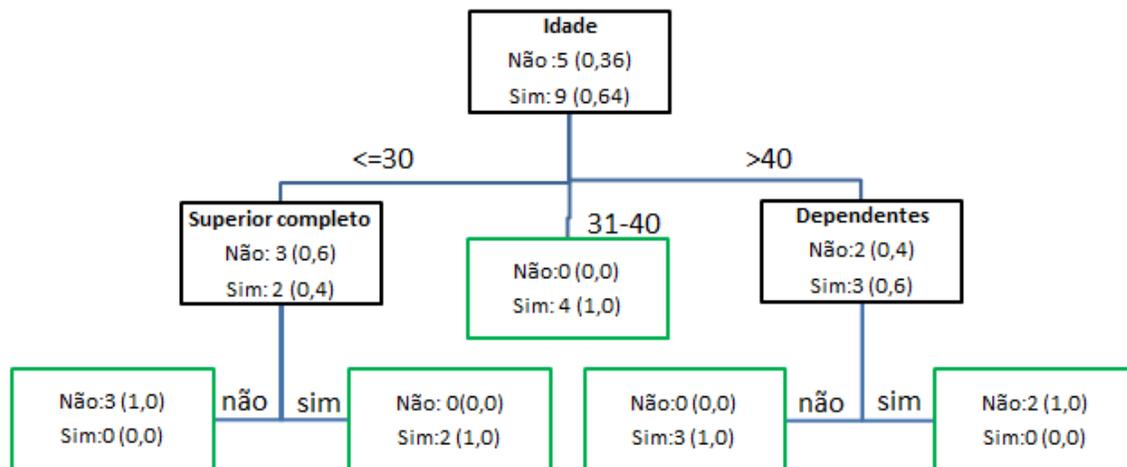


Figura 8: Árvore de decisão gerada a partir dos dados da Tabela 2. A árvore possui cinco nós folha (retângulos verdes), que apresentam os valores (número) e probabilidades (número entre parêntese) da classe, e classificam de acordo com os atributos.

A árvore de decisão montada possui cinco nós folhas (verde) que classificam os valores da classe de acordo com os seus atributos. Notem que o atributo salário não foi utilizado, pois o algoritmo não considerou o atributo como relevante para a classificação.

Desta maneira, fica fácil extrair as regras de classificação do tipo *se-então* desta árvore:

Se idade ≤ 30 e superior completo = não, então a amostra é classificada como atrasa= não.

Se idade ≤ 30 e superior completo =sim, então a amostra é classificada como atrasa= sim.

Se idade =30-40, então a amostra é classificada como atrasa= sim.

Se idade ≥ 40 e dependentes = não, então a amostra é classificada com atrasa= sim.

Se idade ≥ 40 e dependentes = sim, então a amostra é classificada com atrasa= não.

1.5.2. “MULTILAYER PERCEPTRONS”: REDE NEURAL MULTICAMADA (MLP)

As redes neurais artificiais (RNAs) são inspiradas na maneira como se organiza e funciona o cérebro humano. O conhecimento de uma RNA está codificado na estrutura da rede, onde se destacam as unidades (neurônios) e suas conexões (sinapses). Nestas redes artificiais obtidas por simulação em computadores, associa-se a cada conexão um peso sináptico (valor numérico), que caracteriza a força da conexão entre dois neurônios. O aprendizado em uma RNA é realizado por um processo de adaptação de seus pesos sinápticos.⁶⁷

Se os dados a serem classificados podem ser separados linearmente, ou seja, uma linha ou plano pode ser desenhado para separar a instância de entrada entre suas categorias de conexão, então uma rede neural de camada única pode ser treinada e classificar corretamente estes dados. Se as instâncias não podem ser linearmente separadas, o classificador nunca atingirá o ponto ideal que classificará apropriadamente as instâncias. Redes neurais multicamadas (“multilayer perceptron”, MLP) foram criadas para resolver este problema.⁶⁸ A rede neural multicamada (MLP) consiste em um grande número de unidades (neurônios) agrupados num padrão de conexões (Figura 9). Unidos em forma de rede, são usualmente segregados em três classes: camada de entrada (“input layer”), que recebe a informação a ser processada; camada de saída (“output layer”), onde os resultados do processamento são fornecidos; e unidades intermediárias, entre as camadas de entrada e saída, conhecidas como camadas escondidas (“hidden layers”). Com exceção da camada de entrada, todas as outras camadas são constituídas por neurônios e, portanto, apresentam capacidade computacional.⁶⁹

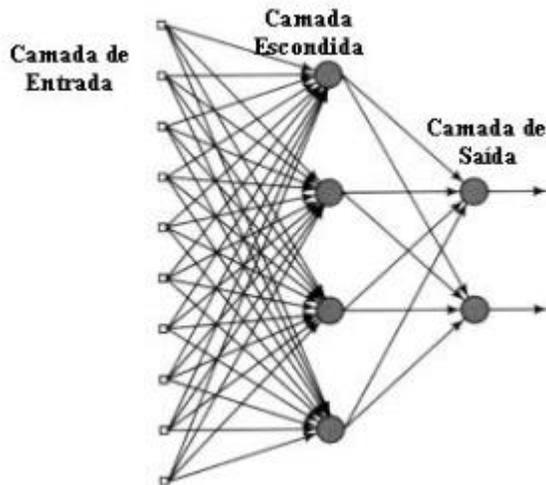


Figura 9: Representação de uma rede neural multicamada. Apresenta uma camada de entrada, uma de saída e apenas uma camada escondida. O sentido de propagação é único (da esquerda para direita), da camada de entrada em direção a camada de saída, passando pela camada escondida.⁷⁰

O sinal caminha pela rede num sentido único, da camada de entrada até a camada de saída, percorrendo as camadas escondidas. Primeiramente, a rede é treinada com os dados pareados para determinar o mapeamento de entrada e saída. Os pesos das conexões entre neurônios são fixos e a rede é usada para determinar classificações de novas amostras.

Durante a classificação, os sinais das camadas de entrada seguem todos na mesma direção e percorrem a rede para determinar os valores de ativação e de todas as camadas de saída. Cada camada de entrada tem um valor de ativação que representa algum parâmetro externo à rede. Então, toda unidade de entrada envia seus valores de ativação para a cada camada escondida no qual está conectado. Cada uma dessas unidades escondidas calcula seu próprio valor de ativação e seu sinal é, então, passado para as unidades de saída. O valor de ativação de cada unidade recebida é calculado de acordo com a função simples de ativação. A função soma todas as contribuições enviadas às unidades. A contribuição de cada unidade é definida como o peso da conexão entre as unidades recebidas e enviadas multiplicados pelo valor da unidade de ativação.⁶³

As RNAs dependem basicamente de três aspectos fundamentais: funções de entrada e ativação das unidades, arquitetura da rede e o peso de cada conexão. Visto que os dois primeiros aspectos são fixos, o comportamento da RNA é definido pelo peso dos valores entre as conexões. Os pesos da rede podem ser treinados inicialmente com valores aleatórios, e então, as instâncias de treinamento são repetidamente expostas à rede. O valor de entrada de uma instância é colocado na unidade de entrada e então a camada de saída da rede é comparada com a unidade de saída desejável para esta instância. Assim, todos os pesos da rede são ajustados na direção que possa trazer os valores de saída da rede mais próximos dos valores desejáveis.⁶³ Existem vários algoritmos que podem treinar a rede, entretanto, o mais conhecido e amplamente utilizado é o algoritmo de retro propagação do inglês “back propagation”(Figura 10).⁷¹ O algoritmo de retro propagação estima os valores dos pesos da rede baseado no aprendizado por correção de erro (em que o erro é retro propagado da camada de saída para as camadas intermediárias da RNA).

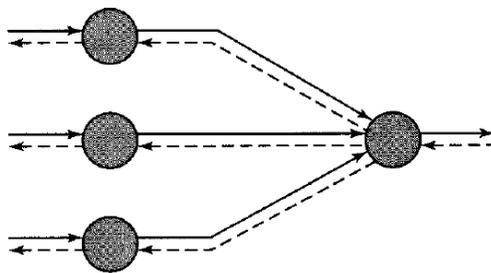


Figura 10: Representação de uma rede MLP com retro propagação. O sentido de propagação é representado pela linha contínua (da esquerda para direita), da camada de entrada em direção à camada de saída. Por outro lado, o sentido da retro propagação, representado pela linha tracejada, ocorre no sentido contrário (da direita para esquerda), da camada de saída em direção à camada de entrada.⁷⁰

Basicamente, o algoritmo de retro propagação consiste de dois passos através das diferentes camadas do MLP: um passo direto e um passo reverso.

No passo direto, um padrão de atividade do processo a ser apreendido (ou vetor de entrada) é aplicada aos nós de entrada do MLP e o seu efeito se propaga através da rede, camada por camada, produzindo na camada de saída a resposta do MLP à excitação aplicada (vetor de saída). Durante o passo direto os pesos sinápticos são todos fixos.⁷⁰

Durante o passo reverso, os pesos sinápticos são todos ajustados de acordo com a regra de aprendizado por correção de erro. Especificamente, a resposta do MLP à excitação é subtraída de um padrão de resposta desejado para aquela excitação aplicada, de forma a produzir um sinal de erro. Esse sinal de erro é então propagado de volta através dos mesmos neurônios utilizados no passo direto, mas no sentido contrário ao fluxo de sinal das conexões sinápticas – daí o nome “back propagation” -. Os pesos sinápticos são, então, ajustados de forma que a resposta obtida do MLP aproxime-se mais da resposta do padrão desejado.⁷⁰

A presença de distribuições não lineares e a alta conectividade tornam difícil a análise teórica das redes MLPs. Em uma rede MLP, o conhecimento aprendido sobre o ambiente é representado pelos valores assumidos pelos pesos sinápticos da rede. A natureza distribuída desse conhecimento ao longo da rede torna difícil sua interpretação.⁷⁰

Na medida em que o conjunto de dados usados para treinar uma RNA MLP seja grande o suficiente para ser representativo do ambiente no qual a rede está inserida, a rede MLP treinada através do algoritmo de retro propagação desenvolverá a capacidade de generalizar. Geralmente, o maior problema consiste na determinação do tamanho das camadas escondidas, porque uma subestimação do número de neurônios pode criar uma capacidade ruim de aproximação e generalização, enquanto nós excessivos podem resultar

em sobre ajuste (“overfitting”, quando o modelo se ajusta em demasiado ao conjunto de dados da amostra) e eventualmente fazer da pesquisa uma má representação da realidade.⁶³

1.5.3. REDE DE FUNÇÕES DE BASE RADIAL (RBF)

Rede de Funções de Base Radial (“Radial Basis Function Network”, RBF) é uma rede de arquitetura neural que tem praticamente as mesmas aplicações da rede neural multicamada (MLP), ou seja, aproximação de mapeamento (funções) não lineares e também classificação de padrões.⁷²

A rede RBF, em sua concepção básica, apresenta apenas uma camada oculta, além de uma camada de saída, diferentemente da rede MLP que pode apresentar uma ou mais camadas de neurônios ocultos. Os neurônios da camada oculta (Figura 11), doravante chamados de funções de base radial, possuem funções de ativação não linear que é bem diferente das funções sigmóides que a rede MLP usa. Por outro lado, os neurônios da camada de saída possuem, em geral, função de ativação linear.⁷²

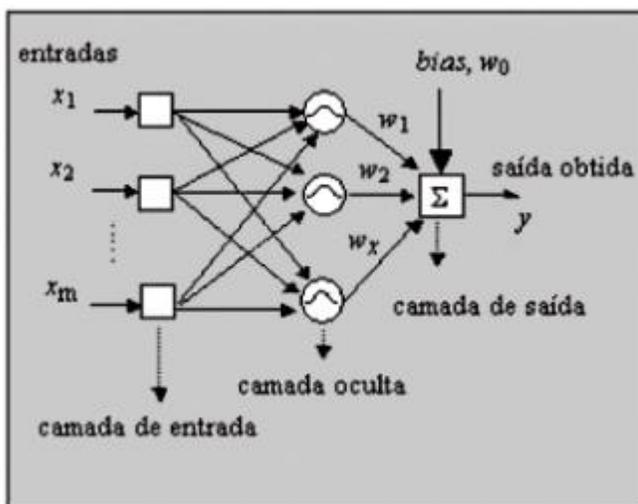


Figura 11: Representação de uma rede de base radial. A RBF apresenta uma camada de entrada (com neurônios X_1 a X_m representados pelos quadrados vazios); uma camada oculta que apresenta neurônios (representados pelos círculos de W_1 a W_x) que possuem funções de ativação não linear, funções de base radial; e uma camada de saída (representada pelo quadrado com símbolo Σ no seu interior) que possui função de ativação linear que obtém uma saída y .⁷³

Para utilizar a rede RBF, é necessário um conjunto de treinamento com N exemplos com dados na forma de pares de vetores (x,d) : $x_{(1)},d_{(1)}$; $x_{(2)},d_{(2)}$; ..., $x_{(n)},d_{(n)}$.

A partir dos dados de treinamento, assume-se que estes vetores estão relacionados de acordo com uma lei matemática e por isso a rede RBF pode gerar uma função matemática, onde espera-se que a camada de saída gerada seja muito próxima da camada de saída real.⁷⁴

O funcionamento da rede RBF envolve basicamente a especificação do número de funções de base, determinação dos seus respectivos parâmetros e determinação dos pesos dos neurônios de saída.⁷⁴

O procedimento de treinamento é dividido em três estágios:

- 1- Determinar os centros das funções de base: A forma mais comum de se determinar os centros das funções de base é por meio de um algoritmo de formação de agrupamentos (“clustering”).
- 2- Determinar os parâmetros das funções de base (raio e largura): utiliza-se de métodos heurísticos (método que leva a máquina a aprender por si mesma) para determinar o raio (“radius”) ou a abertura (“spread”) de cada função de base.
- 3- Determinação dos pesos dos neurônios de saída: existem várias maneiras de se calcular o peso das camadas de saída. Uma das mais utilizadas consiste em usar as regras de aprendizagem de redes neurais de camada única (“single layer perceptron”), onde a cada apresentação de um vetor de entrada, as saídas das funções de base são recalculadas e seus valores são utilizados para atualizar os pesos de saída.

O problema da seleção do número apropriado de funções de base permanece um aspecto crítico das redes RBF. O número de funções de base controla a complexidade e a

habilidade de generalização da rede RBF. A rede RBF com poucas funções-base pode não se adaptar aos dados de treinamento adequadamente por causa da flexibilidade limitada. Por outro lado, aquelas com muitas funções base podem ter pouca capacidade de generalização, pois se tornam muito flexíveis e erroneamente se adaptam aos ruídos existentes nos dados de treinamento. A principal desvantagem da RBF é a complexidade de compreensão da maneira como se obtém os resultados nas camadas de saída.⁶³

1.5.4. CLASSIFICADOR “NAIVE BAYES”

O classificador “Naive Bayes” (NB) é um exemplo de algoritmo de aprendizagem estatístico, pois é completamente baseado em probabilidades. Contrariamente às redes neurais artificiais, ele apresenta um modelo estatístico explícito que provê uma probabilidade que uma instância pertença a uma determinada classe, ao invés de simplesmente classificá-la.⁶³

O classificador é denominado ingênuo (“naive”) por assumir que os atributos são condicionalmente independentes, ou seja, a informação de um evento não é informativa sobre nenhum outro, o que simplifica os cálculos envolvidos.⁷⁵

Ele obtém melhores resultados quando os dados são discretos (categóricos) ao invés de contínuos. Se lidamos com dados contínuos (numéricos), o procedimento usual é estratificá-los durante o processo de pré-processamento dos dados,⁷⁶ apesar de que é possível usar distribuição normal para calcular as probabilidades.⁷⁷

O objetivo principal do classificador NB é calcular a probabilidade que uma amostra desconhecida pertença a uma determinada classe entre as classes possíveis, ou seja, determinar a classe mais provável. Ele necessita de um conjunto de treinamento que já foi

classificado previamente (“cluster”) e quando recebe uma amostra sem classificação, ele retorna como saída a classe mais provável para esta amostra de acordo com cálculos probabilísticos.⁷⁸

Basicamente o algoritmo funciona seguindo quatro estágios:

- 1- Cálculo da probabilidade das classes: o cálculo é feito dividindo-se o número de instâncias de determinada classe pelo número total de instâncias do conjunto de treinamento.
- 2- Cálculo das probabilidades da amostra desconhecida para cada classe possível.
- 3- Cálculo das probabilidades da amostra desconhecida: as probabilidades calculadas para os valores da amostra desconhecida de uma mesma classe são multiplicadas. Em seguida, o valor obtido é multiplicado pela probabilidade da classe calculada no estágio 1.
- 4- Verificação da classe com maior probabilidade: com as probabilidades de cada classe calculadas, verifica-se qual é a classe que possui maior probabilidade para a amostra desconhecida. Com isso, o algoritmo apresenta como saída a classe que possui maior probabilidade de conter a amostra desconhecida.

Exemplo de uso de algoritmo NB:⁷⁸

Neste exemplo, vamos considerar que uma empresa financeira deseja prever se um cliente será inadimplente ou não. Para isso, a financeira deve levar em consideração a sua base histórica de clientes e alguns atributos. Para facilitar o entendimento do cenário e do modelo de dados, vamos utilizar um conjunto de treinamento com quinze instâncias (linhas) com quatro atributos (sexo, estado civil, escolaridade, rendimentos). A Tabela 3 mostra o conjunto de treinamento que será utilizado neste exemplo.

Tabela 3: Dados de uma empresa financeira para geração de um exemplo de um classificador NB.

| cliente | Sexo | Estado civil | Escolaridade | Rendimentos | Inadimplente |
|---------|-----------|--------------|---------------------|-----------------------------|--------------|
| 1 | Feminino | Solteiro | Médio incompleto | Acima de 3 salários mínimos | Não |
| 2 | Feminino | Solteiro | Médio incompleto | Um salário mínimo | Não |
| 3 | Masculino | Solteiro | Médio completo | Acima de 3 salários mínimos | Sim |
| 4 | Feminino | Solteiro | Médio completo | Um salário mínimo | Não |
| 5 | Feminino | Solteiro | Superior incompleto | Dois salários mínimos | Não |
| 6 | Masculino | Casado | Médio completo | Um salário mínimo | Não |
| 7 | Feminino | Casado | Médio completo | Acima de 3 salários mínimos | Não |
| 8 | Masculino | Casado | Médio incompleto | Acima de 3 salários mínimos | Não |
| 9 | Masculino | Casado | Médio incompleto | Um salário mínimo | Não |
| 10 | Masculino | Casado | Superior completo | Dois salários mínimos | Não |
| 11 | Feminino | Solteiro | Médio completo | Acima de 3 salários mínimos | Não |
| 12 | Feminino | Casado | Médio incompleto | Dois salários mínimos | Sim |
| 13 | Feminino | Casado | Superior incompleto | Um salário mínimo | Sim |
| 14 | Masculino | Casado | Superior incompleto | Dois salários mínimos | Sim |
| 15 | Feminino | Solteiro | Superior completo | Dois salários mínimos | Não |

As colunas apresentadas na Tabela 3 são descritas a seguir:

CLIENTE: Esta coluna possui um identificador inteiro sequencial. Para o algoritmo esta coluna é opcional, mas ela ajuda a organizar as linhas de histórico da tabela.

SEXO: Esta coluna identifica o gênero do cliente. Pode assumir apenas os valores **MASCULINO** e **FEMININO**.

ESTADO CIVIL: Esta coluna traz informações sobre o estado civil do cliente. Pode assumir apenas os valores **CASADO** e **SOLTEIRO**.

ESCOLARIDADE: Esta coluna traz informações sobre a escolaridade do cliente. Pode assumir apenas quatro categorias diferentes: **ENSINO MÉDIO INCOMPLETO**, **ENSINO MÉDIO COMPLETO**, **SUPERIOR INCOMPLETO** e **SUPERIOR COMPLETO**.

RENDIMENTOS: Esta coluna traz informações sobre a faixa salarial do cliente. Pode assumir apenas as categorias UM SALÁRIO MÍNIMO, DOIS SALÁRIOS MÍNIMOS e ACIMA DE TRÊS SALÁRIOS MÍNIMOS.

INADIMPLENTE: Esta é a coluna que apresenta a classificação das amostras. Neste exemplo a classificação indica se o cliente é devedor, isto é INADIMPLENTE=SIM, ou se o cliente não é devedor, INADIMPLENTE=NÃO. Para facilitar a visualização, os clientes do conjunto de treinamento que são inadimplentes foram marcados em amarelo e os clientes que não são inadimplentes foram marcados em verde.

Vamos agora executar a classificação Bayesiana para a amostra desconhecida. Com base nos dados históricos apresentados na Tabela 3, deseja-se classificar por meio da classificação Bayesiana a amostra desconhecida apresentada na Figura 12.

| SEXO | ESTADOCIVIL | ESCOLARIDADE | RENDIMENTOS | INADIMPLENTE |
|-----------|-------------|-------------------------|-------------------|--------------|
| MASCULINO | SOLTEIRO | ENSINO MÉDIO INCOMPLETO | UM SALARIO MÍNIMO | ? |

Figura 12: Amostra desconhecida para ser classificada de acordo com o classificador NB utilizando os dados e atributos da tabela 3.

Passo 01: Cálculos das probabilidades das classes.

Existem apenas duas classes, uma que indica que o cliente é devedor (INADIMPLENTE=SIM) e outra que indica que o cliente não é devedor (INADIMPLENTE=NÃO). Calculando as probabilidades das classes temos:

Probabilidade INADIMPLENTE=SIM: $4/15 = 0,2667$

Probabilidade INADIMPLENTE=NÃO: $11/15 = 0,7334$

Passo 02: Cálculo das probabilidades da amostra desconhecida.

Para o primeiro atributo da amostra desconhecida SEXO=MASCULINO, vamos calcular a probabilidade de INADIMPLENTE=SIM:

Probabilidade de SEXO=MASCULINO e INADIMPLENTE=SIM: $2/4 = 0,5$

E para o caso onde o cliente é masculino e não é devedor temos:

Probabilidade de SEXO=MASCULINO e INADIMPLENTE=NÃO: $4/11 = 0,3636$

Para os demais valores dos atributos da amostra desconhecida, temos:

Probabilidade de ESTADO CIVIL=SOLTEIRO e INADIMPLENTE=SIM: $1/4 = 0,25$

Probabilidade de ESTADO CIVIL=SOLTEIRO e INADIMPLENTE=NÃO: $6/11 = 0,5455$

Probabilidade de ESCOLARIDADE=ENSINO MÉDIO INCOMPLETO e INADIMPLENTE=SIM: $1/4 = 0,25$

Probabilidade de ESCOLARIDADE = ENSINO MÉDIO INCOMPLETO e INADIMPLENTE=NÃO: $4/11 = 0,3636$

Probabilidade de RENDIMENTOS= UM SALÁRIO MÍNIMO e INADIMPLENTE = SIM:

$1/4 = 0,25$

Probabilidade de RENDIMENTOS = UM SALÁRIO MÍNIMO e

INADIMPLENTE=NÃO: $4/11 = 0,3636$

Passo 03: Calcular as probabilidades da amostra desconhecida.

Multiplicando as probabilidades da amostra desconhecida para o caso de INADIMPLENTE=SIM pela probabilidade de inadimplência calculada no PASSO 1 temos:

$$0,5 \times 0,25 \times 0,25 \times 0,25 \times 0,2667 = 0,0021$$

Multiplicando as probabilidades da amostra desconhecida para o caso de INADIMPLENTE=NÃO pela probabilidade de não inadimplência calculada no PASSO 1, temos:

$$0,3636 \times 0,5455 \times 0,3636 \times 0,3636 \times 0,7334 = 0,0192$$

Passo 04: Verificação da classe com maior probabilidade.

Como $0,0192 > 0,0021$, o algoritmo classifica a amostra desconhecida como INADIMPLENTE=NÃO, ou seja, este novo cliente tem uma probabilidade maior de não se tornar devedor do que de se tornar inadimplente, de acordo com os dados históricos e a classificação Bayesiana.

1.5.5. SUPORT VECTOR MACHINES

As máquinas de Vetores de Suporte (SVM, “Support Vector Machines”) são embasadas na teoria de aprendizado estatístico, desenvolvido por Vapnik⁷⁹ a partir de

estudos iniciados em 1971.⁸⁰ A teoria de aprendizado estatístico (TAE) estabelece condições matemáticas que auxiliam na escolha de um classificador gerado a partir de um conjunto de dados de treinamento. Essas condições levam em conta o desempenho do classificador no conjunto de treinamento e a sua complexidade, com o objetivo de obter um bom desempenho também para novos dados do mesmo domínio.⁸¹

As SVMs se baseiam na noção de criação de uma “margem” – cada lado de um hiperplano, que é separado em duas classes. Maximizando a margem e, conseqüentemente, criando uma maior distância possível entre os hiperplanos e as instâncias a serem separadas, reduz-se o erro de generalização esperado.⁶³

O objetivo do processo de aprendizagem estatística é, então, obter uma função indicadora que minimize o risco funcional, por meio de relações extraídas deste conjunto. O risco funcional é a probabilidade da saída desejada ser diferente da saída obtida pela máquina após a escolha de uma função indicadora.⁸² Para que isto ocorra, utiliza-se o princípio da minimização do risco estrutural, que é fundamentado no fato de que a taxa de erro nos dados de teste (taxa de erro de generalização) é limitada pela soma da taxa de erro de treinamento e por um termo que depende da dimensão Vapnik-Chervoneskis (dimensão VC). A dimensão VC não tem conexão com a noção geométrica de dimensão, é uma medida da capacidade ou poder de expressão de um conjunto de funções.⁸³

A dimensão VC, pode ser definida como o maior número de pontos que podem ser separados de todas as maneiras possíveis, usando-se funções da classe escolhida. O SVM constrói um conjunto de hiperplanos cujos limites da dimensão VC possam ser computados e usa, então, o princípio de minimização do risco estrutural para identificar o hiperplano ótimo que maximize a margem dos exemplos mais próximos.⁸³

De maneira simplificada, o funcionamento de uma SVM pode ser descrito da seguinte forma: dadas duas classes e um conjunto de pontos que pertencem a essas classes, uma SVM determina o hiperplano que separa os pontos de forma a colocar o maior número de pontos da mesma classe do mesmo lado, enquanto maximiza a distância de cada classe a esse hiperplano. A distância de uma classe a um hiperplano corresponde à menor distância entre ele e os pontos dessa classe e é chamada de margem de separação. O hiperplano gerado pela SVM é determinado por um subconjunto dos pontos das duas classes, chamado vetores de suporte.⁸³

Considere, por exemplo, o conjunto de treinamento da Figura 13.⁸¹ Objetivo do processo de aprendizagem é encontrar um classificador que crie uma margem que separe os dados das classes triângulo e círculo. As funções ou hipóteses são ilustradas na figura por meio das bordas, também denominadas fronteiras de decisão, traçadas entre as classes.

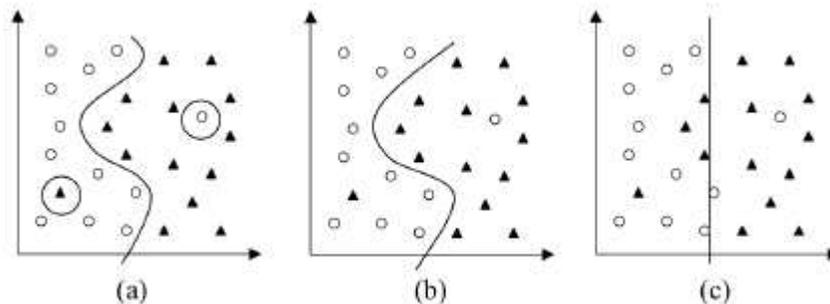


Figura 13: Exemplos de três classificadores distintos. O classificador é representado pela linha contínua que separa os dados pertencentes a duas classes (triângulos e círculos).⁸¹

Na imagem da Figura 13(a), tem-se uma hipótese que classifica corretamente todos os exemplos do conjunto de treinamento, incluindo dois possíveis ruídos (dados classificados incorretamente). Por ser muito específico para o conjunto de treinamento, essa função é muito susceptível a cometer erros quando confrontada com novos dados. Esse

caso representa a ocorrência de sobre ajustamento (“overfitting”) do modelo aos dados de treinamento.

Outro classificador poderia desconsiderar pontos pertencentes a classes opostas que estejam muito próximos entre si. A ilustração da Figura 13(c) representa esta alternativa. A nova hipótese considerada, porém, comete muitos erros, mesmo para casos considerados simples. Nessa situação, ocorre um sub ajustamento (“underfitting”), pois o classificador não é capaz de se ajustar aos exemplos de treinamento.

O melhor classificador é representado na Figura 13(b). Esse preditor tem complexidade intermediária e classifica corretamente grande parte dos dados, sem se fixar demasiadamente em qualquer ponto individual.

Outro exemplo de margem de separação de hiperplanos é representado na Figura 14. Na Figura 14(a) temos uma margem pequena e na Figura 14(b) temos um hiperplano com margem máxima, ou seja, com distância máxima entre os dados de cada classe.

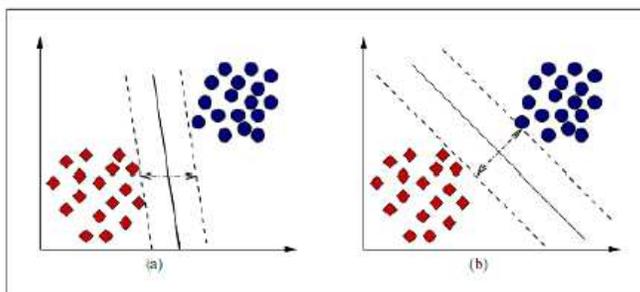


Figura 14: Representação de margens de separação de hiperplanos. Duas classes (quadrados vermelhos e círculos azuis) são separadas por um hiperplano (linha contínua) e a maior distância entre as classes (espaço entre as linhas tracejadas) corresponde à margem de separação. A figura (a) apresenta um hiperplano de separação com margem pequena. A figura (b) apresenta um hiperplano de margem máxima, pois apresenta a maior distância entre as linhas tracejadas. ⁸³

A complexidade do modelo de uma SVM não é afetada pelo número de parâmetros encontrados nos dados de treinamento (o número de vetores de suporte selecionados pelo algoritmo SVM de treinamento é usualmente pequeno). Por esse motivo, SVMs se adaptam bem quando lidam com tarefas de aprendizagem que envolvem um grande número de parâmetros em relação ao número de instâncias de treinamento.⁶³

1.5.5.1 SVM linear

Para padrões linearmente separáveis, a solução do problema de treinamento de SVMs consiste em achar um hiperplano que separe perfeitamente os pontos de cada classe e cuja margem de separação seja máxima (SVM de margem rígida). Esse hiperplano é chamado de hiperplano ótimo.⁸⁴ Uma vez que o plano de separação ótimo é encontrado, a solução é representada como uma combinação linear dos pontos que pertencem à margem (esses pontos são os vetores de suporte). Os outros pontos são ignorados.⁸⁵

Embora a margem máxima permita ao SVM selecionar entre vários hiperplanos possíveis, para muitos conjuntos de dados pode não ser possível encontrar algum hiperplano que separe todos os dados das classes porque os dados contém ruídos (dados classificados incorretamente), “outliers” (dados de treinamento muito diferentes dos dados a serem classificados) ou pela própria natureza do problema, que pode ser não linear. O problema pode ser resolvido usando uma margem suave (SVM de margem suave), que aceita algumas instâncias de treinamento classificadas erroneamente.⁸⁶

1.5.5.2 SVM não linear (Gaussiana)

A maioria dos problemas reais envolvem dados não separáveis, nos quais não existe um hiperplano que possa separar com sucesso as instâncias positivas e negativas do conjunto de treinamento. Uma solução para este problema de separação é o mapeamento dos dados de treinamento de seu espaço original para um espaço de maior dimensão, e definir um hiperplano separável nele. Esse espaço de alta dimensão é chamado de espaço de característica transformado (“transformed feature space”).⁶³ Um exemplo é apresentado na Figura 15:

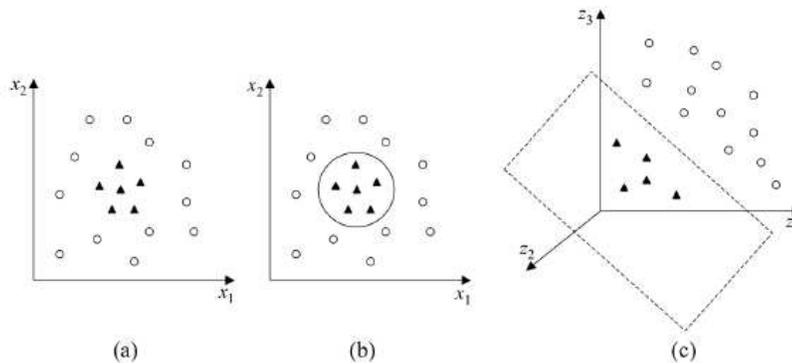


Figura 15: Exemplo de criação de um espaço com maior dimensão separável por um hiperplano. Na figura (a) temos um conjunto de dados (triângulos pretos e círculos brancos) não separáveis linearmente. Na figura (b) é representada uma fronteira não linear onde é possível separar as duas classes. Na figura (c) é representada a criação de um espaço de maior dimensão (x_3) a partir de um espaço original (x_1 e x_2) possibilitando a criação de um hiperplano (retângulo com linhas tracejadas) que pode separar as duas classes corretamente.⁸¹

O uso desse procedimento é motivado pelo teorema de Cover.⁸⁷ Dado um conjunto de dados não linear no espaço de entrada X , esse teorema afirma que X pode ser transformado em um espaço onde os dados apresentem altas probabilidades de serem linearmente separáveis.

Com a escolha apropriada do espaço de características com dimensionalidade suficiente, qualquer conjunto de treinamento consistente pode se tornar separável. Uma separação linear no espaço do parâmetro transformado corresponde a uma separação não linear no espaço original de entrada.⁶³

Uma vez que o hiperplano foi criado, uma função Kernel é usada para mapear novos pontos no espaço de características transformado para classificação.⁸⁸ Alguns dos Kernels mais utilizados na prática são os polinomiais, os gaussianos e os sigmoidais.⁸⁹

1.5.6. CLASSIFICADOR BAGGING

O “Bootstrap Aggregating” ou “Bagging” (BAG) é um algoritmo que se utiliza da combinação de classificadores. A agregação de diversos classificadores tem como objetivo obter uma taxa de acerto melhor do que a obtida pela aplicação de um classificador único.⁹⁰ Para obter um classificador, “Bagging” utiliza métodos que manipulam os conjuntos de treinamento.⁹¹

O algoritmo “Bagging” apresenta 4 etapas até a classificação final (Figura 16):⁹²

- 1- Geração de diferentes amostras a partir da mesma base de dados de treinamento. As amostras devem ter tamanhos iguais, obtidas de forma aleatória e com distribuição uniforme. Isso é obtido por meio de reposição, ou seja, para se obter amostras de tamanhos iguais, algumas instâncias se repetirão, enquanto outras serão omitidas;
- 2- Obtenção de um classificador para cada amostra, representado como árvores de decisão;
- 3- Obtenção de um classificador composto, que inclui todos os classificadores individuais gerados na fase anterior;

- 4- Votação simples, para cada classificação de uma instância, para eleger a classificação mais popular dentre os classificadores individuais.

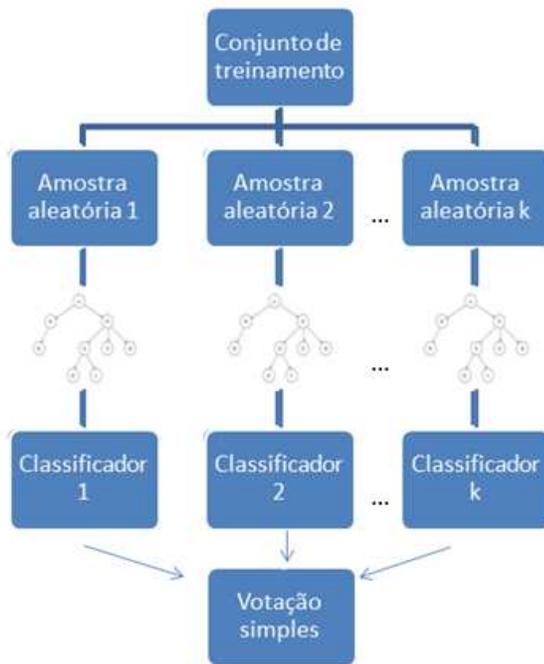


Figura 16: Representação esquemática das etapas para criação do classificador “Bagging”. Na primeira etapa, são geradas amostras aleatórias, a partir do conjunto de treinamento. Na segunda etapa, são gerados classificadores derivados de cada amostra. Na terceira etapa, é criado um classificador composto, que inclui todos os classificadores individuais e, então, é realizada uma votação simples (quarta etapa) para eleger a classificação mais popular. (Fonte: Fabrício Reis da Silva)

O método “Bagging” gera um classificador composto. O resultado final é um conjunto de classificadores que são utilizados de forma integrada para classificar uma nova instância, onde a resposta será aquela escolhida pela maioria dos classificadores.⁹³

O algoritmo “Bagging” é muito útil quando o conjunto de treinamento apresenta dados que possuem comportamento instável, ou seja, pequenas mudanças no conjunto de treinamento acarretam mudanças de classificação.⁹³ Nestes casos, um classificador único

não é capaz de oferecer resposta confiável para todas as situações, mas um conjunto de classificadores (composto) pode ter maior chance de acerto.⁹⁴

Vejamos um exemplo de funcionamento do algoritmo BAG:⁹²

Aplicando o método “Bagging” à base de dados da tabela 4 que possui 14 instâncias (n= 14) utilizando três classificadores (k = 3), teríamos, por exemplo, as três amostras geradas aleatoriamente representadas conforme a tabela 5.

Tabela 4: Base de dados para criação de exemplos de classificadores “Bagging” e “Adaboosting”. São 14 instâncias que apresentam três atributos (frequência cardíaca, frequência respiratória e perda do apetite) e um atributo preditor (uso de medicação).

| instância | Frequência Cardíaca | Frequência Respiratória | Perda do Apetite | Medicação (atributo preditor) |
|-----------|---------------------|-------------------------|------------------|-------------------------------|
| 1 | Normocárdico | Taquipnéia | Não | Sim |
| 2 | Normocárdico | Taquipnéia | Sim | Sim |
| 3 | Taquicárdico | Taquipnéia | Não | Sim |
| 4 | Bradicárdico | Taquipnéia | Não | Não |
| 5 | Bradicárdico | Taquipnéia | Não | Sim |
| 6 | Bradicárdico | Eupnéia | Sim | Não |
| 7 | Taquicárdico | Eupnéia | Sim | Sim |
| 8 | Normocárdico | Taquipnéia | Não | Não |
| 9 | Normocárdico | Eupnéia | Não | Não |
| 10 | Bradicárdico | Eupnéia | Não | Não |
| 11 | Normocárdico | Eupnéia | Sim | Não |
| 12 | Taquicárdico | Taquipnéia | Sim | Não |
| 13 | Taquicárdico | Eupnéia | Não | Sim |
| 14 | Bradicárdico | Taquipnéia | Sim | Não |

Tabela 5: Três amostras, com a mesma quantidade de instâncias (14 instâncias), geradas aleatoriamente a partir da tabela 4 pelo método de reposição.

| Amostra 1 | | | | | Amostra 2 | | | | | Amostra 3 | | | | |
|-----------|----|----|----|---|-----------|----|----|----|---|-----------|----|----|----|---|
| PO | FC | FR | PA | M | PO | FC | FR | PA | M | PO | FC | FR | PA | M |
| 1 | N | T | S | S | 1 | N | T | N | S | 2 | N | T | S | S |
| 2 | N | T | S | S | 3 | T | T | N | S | 2 | N | T | S | S |
| 3 | T | T | N | S | 3 | T | T | N | S | 2 | N | T | S | S |
| 4 | B | T | N | N | 6 | B | E | S | N | 3 | T | T | N | S |
| 5 | B | T | N | S | 6 | B | E | S | N | 3 | T | T | N | S |
| 6 | B | E | S | N | 7 | T | E | S | S | 3 | T | T | N | S |
| 6 | B | E | S | N | 8 | N | T | N | N | 4 | B | T | N | N |
| 7 | T | E | S | S | 9 | N | E | N | N | 5 | B | T | N | S |
| 7 | T | E | S | S | 10 | B | E | N | N | 6 | B | E | S | N |
| 9 | N | E | N | N | 10 | B | E | N | N | 8 | N | T | N | N |
| 10 | B | E | N | N | 11 | N | E | S | N | 12 | T | T | S | N |
| 10 | B | E | N | N | 13 | T | E | N | S | 14 | B | T | S | N |
| 10 | B | E | N | N | 13 | T | E | N | S | 14 | B | T | S | N |
| 14 | B | T | S | N | 13 | T | E | N | S | 14 | B | T | S | N |

PO=paciente, FC=frequência cardíaca, N da coluna FC=normocárdico, B=bradicárdico, T da coluna FC=taquicárdico, FR= frequência respiratória, E= eupnéico, T da coluna FR= taquipnéico, PA= perda de apetite, S= sim , N=não, M=medicação.

Os valores desta tabela foram obtidos por tiragens aleatórias uniformemente distribuídas sobre as instâncias da base original (Tabela 4). Desta forma, para a primeira amostra foram sorteadas as instâncias 1, 2, 3, 4, 5, 6 (2x), 7 (2x), 9, 10 (3x) e 14. Para a segunda amostra, foram sorteadas as instâncias 1, 3 (2x), 6 (2x), 7, 8, 9, 10 (2x), 11, 13(3x). Para a terceira amostra, foram sorteadas as instâncias 2 (3x), 3 (3x), 4, 5, 6, 8, 12, 14 (3x).

Para cada uma das três amostras da Tabela 5, um classificador é gerado de acordo com a classificação mais frequente entre as instâncias. Por exemplo, para a amostra 1 da Tabela 4, três instâncias apresentaram o atributo frequência cardíaca: normocárdico (pacientes 1, 2 e 9). Duas instâncias apresentaram o preditor “uso de medicação: SIM” e apenas uma o preditor “uso de medicação: “NÃO”, conseqüentemente, o classificador gerado irá apresentar a classificação “SIM” para uso de medicação quando a frequência cardíaca for normocárdica. (Figura 17)

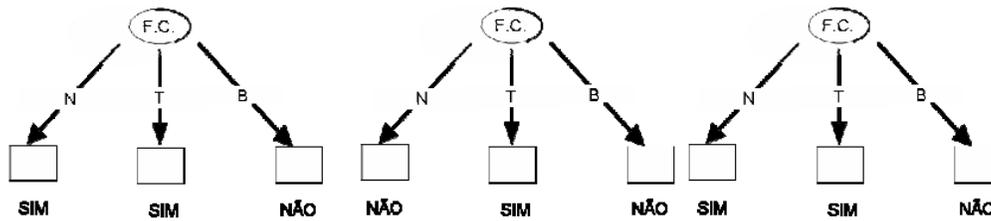


Figura 17: Classificadores gerados para as três amostras da Tabela 5.

Para exemplificar o processo de classificação no método “Bagging”, vamos assumir na Tabela 6 o recebimento de três instâncias a serem classificadas.

Tabela 6: Três instâncias a serem classificadas pelo método “Bagging” baseados nos dados da tabela 4 e nas amostras geradas aleatoriamente da tabela 5.

| instância | Frequência Cardíaca | Frequência Respiratória | Perda do Apetite | Medicação (predição) |
|-----------|---------------------|-------------------------|------------------|----------------------|
| a_1 | Normocárdico | Eupnéia | Não | ? |
| a_2 | Bradicárdico | Eupnéia | Não | ? |
| a_3 | Taquicárdico | Taquipnéia | Sim | ? |

A classificação da primeira destas instâncias (a_1) resultaria na resposta SIM, para o Classificador C1; NÃO, para o Classificador C2; e SIM, para o Classificador C3. Consequentemente a resposta SIM será considerada, pois ganha a votação por 2 votos a 1.

A classificação da instância a_2 resultaria na resposta NÃO para todos os classificadores, ganhando a votação por unanimidade. A mesma situação acontece para a instância a_3 , que resultaria na resposta SIM, também por unanimidade.

1.5.7. CLASSIFICADOR “ADABOOSTING”

O método “Boosting” é outro método de combinação de classificadores, que foi desenvolvido com o intuito de oferecer classificadores mais eficientes. A partir daí, foram

criados vários algoritmos, sendo que o mais conhecido é o “AdaBoost (Adaptative Boosting)” (ADA).⁹⁵

Neste método, como no método “Bagging”, são geradas amostras que dão origem a classificadores que são utilizados de forma integrada. Entretanto, existem duas distinções entre os métodos “Bagging” e “Boosting”: (1) a maneira como são geradas as amostras; e (2) a maneira como são combinados os resultados dos classificadores.⁹²

A maneira como as amostras são geradas no “Boosting” difere do “Bagging” pois não são obtidas através de uma distribuição uniforme e sim através da variação do peso associado a cada um dos exemplos para a geração das próximas amostras.

A segunda diferença está na forma como são combinados os classificadores. No método “Bagging”, a classificação final é obtida por meio de uma votação simples entre as respostas fornecidas por cada classificador, enquanto no método “Boosting” a votação é ponderada segundo um índice de importância entre os classificadores gerados.⁹⁶

De forma mais específica, a geração de amostras no método “Boosting” consiste em (Figura 18):⁹²

- 1) Gerar a primeira amostra assumindo a distribuição uniforme, ou seja, todas as instâncias da base têm a mesma probabilidade ($1/n$) de serem incluídas na primeira amostra gerada;
- 2) Gerar um classificador para esta amostra e aplicar o classificador à base de treinamento original;
- 3) De acordo com o classificador gerado, diminuir a probabilidade de serem incluídas na próxima amostra instâncias que foram corretamente classificadas e aumentar a probabilidade das instâncias que foram incorretamente classificadas;

- 4) Gerar a segunda amostra a partir da base original, levando em conta as novas probabilidades de cada instância;
- 5) Gerar um classificador para a segunda amostra e aplicar o classificador à base de treinamento original; e mais uma vez, diminuir a probabilidade das instâncias bem classificadas e aumentar a probabilidade das instâncias mal classificadas;
- 6) Repetir este processo de geração de amostras, classificadores e alteração de pesos até serem gerados a k-ésima amostra e o k-ésimo classificador.

O resultado desta operação tende a gerar amostras que são mais “especializadas”, ou seja, enquanto uma amostra é eficiente para classificar algumas das instâncias da base, as próximas amostras tendem a privilegiar instâncias mal classificadas anteriormente.

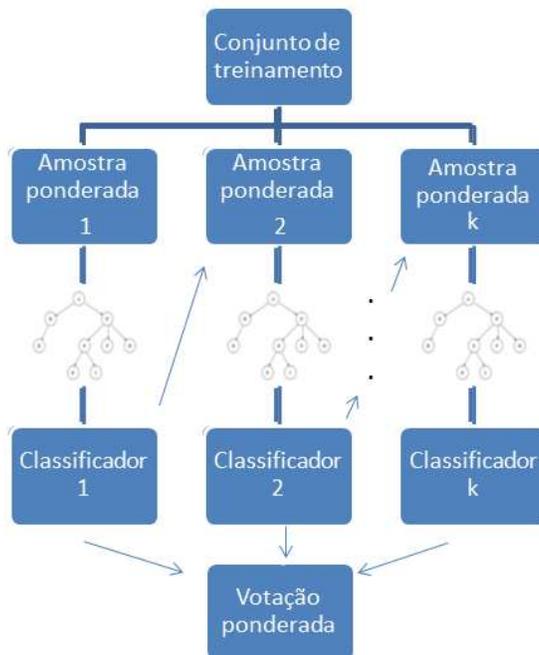


Figura 18: Representação do funcionamento do método “Boosting”. As amostras são geradas de forma ponderada, ou seja, diminui-se a probabilidade de incluir amostras classificadas corretamente e aumenta-se a probabilidade de instâncias classificadas incorretamente. Após a geração de K amostras com K classificadores, uma votação ponderada é realizada para classificar a instância desejada. (Fonte: Fabrício Reis da Silva)

Vejamus um exemplo da aplicação do algoritmo “Boosting”:⁹²

Aplicando o método “Boosting” à base de teste da Tabela 4 (a mesma utilizada para aplicar o método “Bagging” do capítulo anterior) com três classificadores ($k = 3$), o resultado da primeira execução mostra a primeira amostra gerada com uma distribuição uniforme sobre as instâncias da base e o classificador gerado a partir da amostra original.(Figura 19)

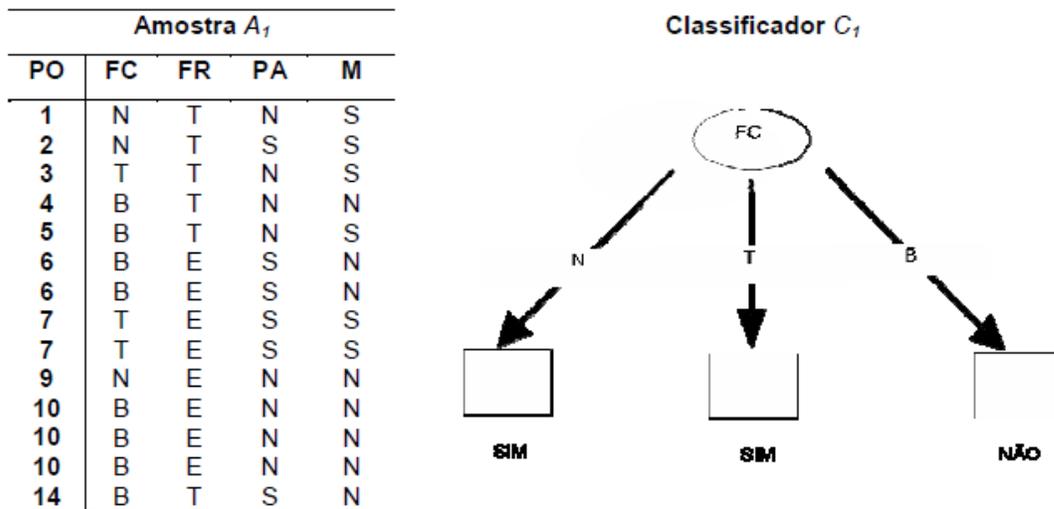


Figura 19: Primeira amostra com distribuição uniforme, criada a partir da tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado. Todas as instâncias apresentaram a mesma probabilidade de serem incluídas na amostra.

Aplicando o classificador C_1 obtido à base de dados original expressa na Tabela 4, verifica-se que as instâncias 5, 8, 9, 11 e 12 foram incorretamente classificadas. Esta tarefa permite o cálculo da taxa de erro e da importância deste classificador: $\epsilon^1 = 0,357$ $\alpha^1 = 0,294$. (passo 1)

A taxa de erro (ϵ^j) de um classificador C_j é feito levando em consideração: o número de instâncias da amostra (n); o peso de cada uma das instâncias (w_j); e a

verificação da efetividade do classificador C_i em classificar corretamente cada uma das instâncias. Numericamente, ela é calculada pela fórmula:

$$\varepsilon^{(j)} = \sum_{j=1..n} w_j \delta(C_i(x_j) \neq y_j)$$

Onde, δ é um operador que retorna o valor 1 caso o parâmetro (x) seja verdadeiro e retorna 0 caso contrário; $C_i(x_j)$ representa a classe prevista pelo classificador C_i para a j -ésima instância da base de treinamento; e y_j representa a classe da j -ésima instância expressa pelo valor do atributo preditor da base de treinamento.

A importância de um classificador ($\alpha^{(i)}$) é calculada com base na sua taxa de erro. Este cálculo é feito pela fórmula:

$$\alpha^{(i)} = (1/2) \ln((1 - \varepsilon_i) / \varepsilon_i)$$

De posse da importância do classificador C_i , decide-se se a amostra e seu classificador serão descartados. Caso a taxa de erro seja superior ao palpite aleatório ($\varepsilon_i > 0,5$) descarta-se a amostra e outra é gerada com os pesos reinicializados para uma distribuição uniforme. Caso a eficiência do classificador seja satisfatória, a atualização dos pesos é feita de acordo com a importância do classificador. Esta atualização de pesos é feita em duas etapas:

Primeiramente, aplica-se um fator multiplicativo aos pesos existentes e depois normaliza-se os pesos obtidos. O peso W_j das instâncias que tiveram suas classes previstas corretamente é decrescido pela divisão pelo fator $e^{\alpha^{(i)}}$, ou seja, ao número de Euler ($e = 2,71828\dots$) elevado à importância do classificador ($\alpha^{(i)}$). De forma análoga, o peso das instâncias que foram previstas incorretamente é acrescido através da multiplicação pelo mesmo fator ($e^{\alpha^{(i)}}$). Feita esta operação, o vetor de pesos w é normalizado, ou seja, todos

seus elementos são divididos por um mesmo fator de normalização de forma que a soma continue igual a 1.

No exemplo dado, como o valor da taxa de erro não supera 0,5, recalculam-se os pesos das instâncias da base de treinamento, obtendo-se os valores apresentados no Quadro 2.

| j | w_j inicial | w_j passo 1 |
|-----|---------------|---------------|
| 1 | 0,071 | 0,056 |
| 2 | 0,071 | 0,056 |
| 3 | 0,071 | 0,056 |
| 4 | 0,071 | 0,056 |
| 5 | 0,071 | 0,100 |
| 6 | 0,071 | 0,056 |
| 7 | 0,071 | 0,056 |
| 8 | 0,071 | 0,100 |
| 9 | 0,071 | 0,100 |
| 10 | 0,071 | 0,056 |
| 11 | 0,071 | 0,100 |
| 12 | 0,071 | 0,100 |
| 13 | 0,071 | 0,056 |
| 14 | 0,071 | 0,056 |

Quadro 2: Vetor de pesos das instâncias da primeira amostra gerada pelo “Boosting” no passo 1.

Após o recálculo dos novos pesos para as instâncias da base de dados (segundo passo), uma nova amostra foi gerada. Esta nova amostra, e também o classificador obtido a partir dela, são apresentados na Figura 20.

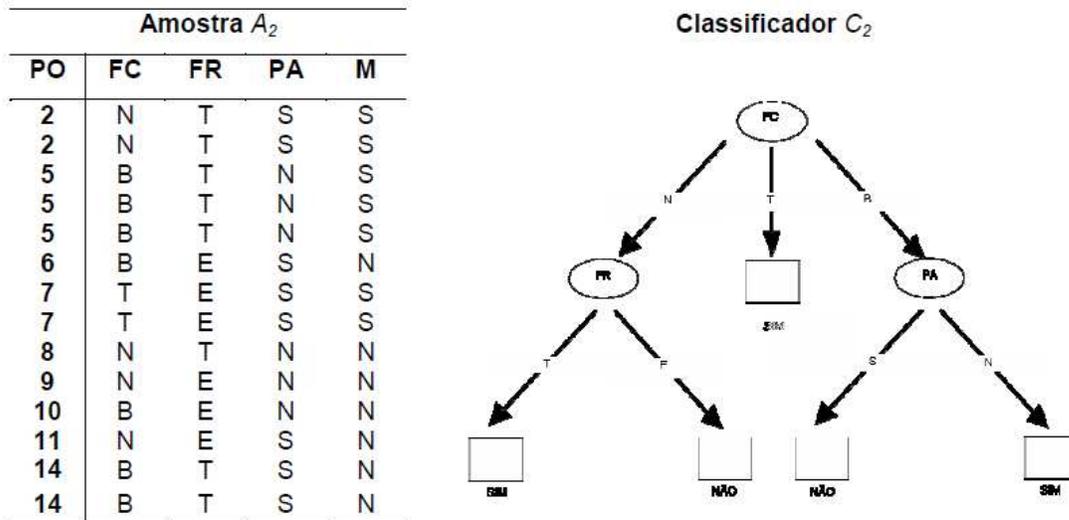


Figura 20: Segunda amostra gerada utilizando a taxa de erro (ponderação), criada a partir da Tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado.

Observa-se que, apesar da amostra gerada ser aleatória, quatro das cinco instâncias mal classificadas pelo Classificador C_1 estão presentes no Classificador C_2 (apenas a instância 12 não foi sorteada). Outro fator marcante é a diferença grande entre os classificadores gerados nos passos 1 e 2. Aplicando-se o novo classificador à base de treinamento original (Tabela 4), verifica-se que as instâncias 4, 8, 10 e 12 foram incorretamente classificadas.

O valor de taxa de erro inferior a 0,5 permite o recálculo dos pesos das instâncias resultando nos valores do Quadro 3. Observa-se neste quadro o aumento da probabilidade das instâncias 8 e 12, que foram erroneamente classificadas nos dois classificadores já gerados. Por outro lado, observa-se que as instâncias 1, 2, 3, 6, 7, 13 e 14 possuem pesos baixos. Estas observações permitem ilustrar o procedimento do método “Boosting” que privilegia a geração de classificadores distintos a cada novo passo.

| j | w_j inicial | w_j passo 1 | w_j passo 2 |
|-----|---------------|---------------|---------------|
| 1 | 0,071 | 0,056 | 0,040 |
| 2 | 0,071 | 0,056 | 0,040 |
| 3 | 0,071 | 0,056 | 0,040 |
| 4 | 0,071 | 0,056 | 0,089 |
| 5 | 0,071 | 0,100 | 0,073 |
| 6 | 0,071 | 0,056 | 0,040 |
| 7 | 0,071 | 0,056 | 0,040 |
| 8 | 0,071 | 0,100 | 0,161 |
| 9 | 0,071 | 0,100 | 0,073 |
| 10 | 0,071 | 0,056 | 0,089 |
| 11 | 0,071 | 0,100 | 0,073 |
| 12 | 0,071 | 0,100 | 0,161 |
| 13 | 0,071 | 0,056 | 0,040 |
| 14 | 0,071 | 0,056 | 0,040 |

Quadro 3: Vetor de pesos das instâncias da primeira e segunda amostra gerada pelo “Boosting” no passo 2.

Na Figura 21 são apresentados a terceira amostra e seu classificador, terminando a geração de classificadores para este exemplo ($k = 3$). Este terceiro classificador é o mais simples possível, pois ele escolhe sempre a classe “NÃO” independente dos valores dos atributos. Apesar de estranho, este classificador monótono não é tão surpreendente assim quando se observa a terceira amostra, onde todas as instâncias, exceto a primeira, possuem o valor “NÃO” no atributo preditor. O dado mais interessante é que, apesar disto, a taxa de erro do classificador C3 não é tão ruim quando aplicado à base de treinamento original.

Os valores da taxa de erro e da importância de C3 são: $\epsilon^3 = 0,274$ $\alpha^3 = 0,487$

| Amostra A_3 | | | | |
|---------------|----|----|----|---|
| PO | FC | FR | PA | M |
| 1 | N | T | N | S |
| 4 | B | T | N | N |
| 4 | B | T | N | N |
| 8 | N | T | N | N |
| 8 | N | T | N | N |
| 8 | N | T | N | N |
| 9 | N | E | N | N |
| 9 | N | E | N | N |
| 9 | N | E | N | N |
| 10 | B | E | N | N |
| 11 | N | E | S | N |
| 12 | T | T | S | N |
| 12 | T | T | S | N |
| 12 | T | T | S | N |

Classificador C_3

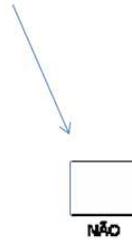


Figura 21: Terceira amostra gerada utilizando a taxa de erro (ponderação), criada a partir da Tabela 4, resultante da execução do classificador “Boosting” e o classificador gerado.

Para exemplificar o processo de classificação no método “Boosting”, vamos assumir na Tabela 7 o recebimento de três novas instâncias a serem classificadas.

Tabela 7: Instâncias a serem classificadas pelo classificador “Boosting”.

| instância | Frequência Cardíaca | Frequência Respiratória | Perda do Apetite | Medicação (predição) |
|-----------|---------------------|-------------------------|------------------|----------------------|
| a_1 | Normocárdico | Eupnéia | Não | ? |
| a_2 | Bradicárdico | Eupnéia | Não | ? |
| a_3 | Taquicárdico | Taquipnéia | Sim | ? |

A classificação da primeira destas instâncias (a_1) resultaria na resposta:

- SIM, para o classificador C_1 , que tem importância 0,294;
- NÃO, para o classificador C_2 , que tem importância 0,398;
- NÃO, para o classificador C_3 , que tem importância 0,487.

Consequentemente, a resposta NÃO será considerada, pois ganha a votação ponderada com o valor 0,885 contra o valor 0,294 da resposta SIM.

A classificação da instância a_2 resulta em:

- NÃO, para o classificador C1, que tem importância 0,294;
- SIM, para o classificador C2, que tem importância 0,398;
- NÃO, para o classificador C3, que tem importância 0,487.

Consequentemente a resposta NÃO será considerada, pois ganha a votação ponderada com o valor 0,781 contra o valor 0,398 da resposta SIM.

A classificação da instância a_3 resulta em:

- SIM, para o classificador C1, que tem importância 0,294;
- SIM, para o classificador C2, que tem importância 0,398;
- NÃO, para o classificador C3, que tem importância 0,487.

Consequentemente a resposta SIM será considerada, pois ganha a votação ponderada com o valor 0,692 contra o valor 0,487 da resposta NÃO.

1.5.8. RANDOM FOREST

Aprendizado em conjunto (“ensemble learning”) inclui métodos que geram vários classificadores e agregam seus resultados⁹⁷. Dois desses métodos muito conhecidos e já descritos são “Boosting”⁹⁵ e “Bagging”⁹¹. Em 2001, Breiman⁹⁸ propôs o método Floresta Randômica ou “Random Forest” (RAN), que acrescenta uma camada de aleatoriedade ao método “Bagging”. Em adição à construção de cada árvore usando uma amostra aleatória dos dados, RAN muda como as árvores classificadoras são construídas. Nas árvores padrão, cada nó é dividido usando a melhor separação entre todas as variáveis. No RAN, cada nó é dividido usando o melhor entre um subgrupo de preditores escolhidos randomicamente

naquele nó. Cada árvore de decisão é construída com profundidade máxima, isto é, sem poda. Isso confere uma melhora na performance quando comparados a outros classificadores.⁹⁷

O algoritmo “Random Forest” é de fácil manuseio, pois o usuário manipula somente dois parâmetros: o número de variáveis (características) do subgrupo randomizado em cada nó e o número de árvores da floresta. O algoritmo RAN funciona da seguinte maneira:⁹⁷

- 1) Desenha $N_{\text{árvores}}$ amostras do conjunto original de dados.
- 2) Para cada amostra, cresce uma árvore de classificação sem poda com a seguinte modificação: em cada nó, ao invés de escolher a melhor separação entre todos os preditores, uma amostra aleatória dos preditores é feita e a melhor separação entre essas variáveis é escolhida.
- 3) A predição de novos dados é feita por meio da combinação de todas as árvores, através do voto majoritário.

O pacote RAN produz opcionalmente uma informação adicional: a medida da importância de uma variável preditora. No geral este é um conceito de difícil definição, porque a importância de uma variável pode ser secundária a sua interação com outras variáveis. O algoritmo RAN estima a importância de uma variável verificando o quanto o erro preditor aumenta quando os dados dessa variável são modificados, enquanto os outros dados são mantidos.⁹⁷

Vejamos o modo como o classificador RAN é gerado:⁹⁹

Seja a coleção de treino onde $D = \{(x_i, c_i) \mid x_i = (y_{i1}, y_{i2}, \dots, y_{im})\}$ é um vetor de características e $c_i \in \{+1, -1\}$, $1 \leq i \leq n$, T o número total de árvores, $M \leq m$ o número de características (atributos) do conjunto de m características.

Antes da construção de cada árvore t , $1 \leq t \leq T$, gera-se uma amostra D_{at} do conjunto total de dados D , obtido de modo randômico, com reposição dos elementos da amostra original. O conjunto da amostra D_{at} é utilizado como conjunto de treino para a construção de t .

Uma vez obtida a amostra D_{at} para a árvore de decisão t , inicia-se a construção da árvore. Cada nó da árvore representa um subconjunto S da amostra D_{at} . Inicialmente, o nó raiz representa o próprio conjunto da amostra D_{at} . Em cada nó opta-se por utilizar a característica (parâmetro) y_j , entre M características, que divida o conjunto S correspondente em dois outros subconjuntos $S1$ e $S2$, que sejam os mais puros possíveis em relação às classes -1 e 1 . Os subconjuntos $S1$ e $S2$ correspondem respectivamente, as subárvores esquerda e direita do nó correspondente a S . Um critério de pureza pré-estabelecido (por exemplo, ganho de informação), é utilizado para decidir não somente qual característica y_j , dentre M características, que melhor divide S , como também, para determinar qual valor v_{y_j} de y_j gera a melhor divisão, em termos de pureza.

A árvore t é construída iniciando-se pelo nó raiz, que representa o conjunto $S=D_{at}$. Uma vez que a floresta com t árvores é construída, a classificação pode ser feita com base nos votos de todas as árvores. Cada árvore emite o voto -1 ou 1 . A classificação dada por uma floresta corresponde ao voto da maioria das árvores.

1.5.9. ENSEMBLE SELECTION

Em 2004, Caruana e colaboradores¹⁰⁰ criaram o algoritmo Seleção de Conjuntos (“Ensemble Selection”, ENS), que é um classificador que utiliza um conjunto de classificadores, a fim de melhorar a sua performance. O algoritmo treina um grande número de classificadores, usando múltiplos algoritmos e parâmetros (características), com a ideia de que pelo menos alguns dos classificadores terão boa performance. O melhor conjunto de classificadores é então, selecionado. Semelhante a outros métodos compostos de aprendizado como “Bagging”⁹¹ e “Boosting”, o algoritmo ENS tem mostrado melhor performance que muitos classificadores isolados. Eles atingem maior acurácia devido, principalmente, à correção de seus erros não correlacionados.¹⁰¹

Nenhum algoritmo é superior aos outros em todas as tarefas, pois o desempenho depende da característica dos dados que estão sendo utilizados. Na prática, diferentes algoritmos são capazes de obter vantagens em diferentes características em relação a um determinado banco de dados. Intuitivamente, o objetivo do ENS é, automaticamente, detectar e combinar as forças desses algoritmos únicos para criar uma soma que é melhor que as partes.¹⁰²

A estrutura do algoritmo ENS inclui:¹⁰¹

- 1- Seleção de uma variedade de algoritmos de aprendizagem;
- 2 - Escolha de um grande conjunto de parâmetros para cada algoritmo;
- 3 - Divisão dos dados entre treinamento, afinamento (“tunning”) e teste;
- 4 - Construção da “biblioteca”;
- 5 - Seleção do objetivo métrico apropriado para o problema (ex. acurácia);

6- Seleção dos classificadores com os melhores desempenhos para o objetivo métrico selecionado;

7- Predição através da média dos preditores constituintes da biblioteca.

O algoritmo ENS se baseia na construção de uma “biblioteca”, que é um amplo e diversificado conjunto de classificadores (modelos) que utilizam diferentes parâmetros. Após a construção desta biblioteca, inicia-se a combinação dos modelos. Começamos com o melhor modelo de toda biblioteca, aquele que faça o melhor trabalho de acordo com o conjunto de validação. Então, adicionamos modelos um de cada vez. Para descobrir qual modelo é adicionado a cada vez, o algoritmo escolhe o modelo que provê a maior melhora na performance do conjunto de classificadores (modelos).¹⁰¹

Algoritmo ENS é muito útil quando existe um grande banco de dados. A maioria dos algoritmos de indução é computacionalmente complexa e sofre com problemas de memória, quando aplicados a uma grande base de dados. A solução para esse problema é a divisão horizontal do banco de dados em partes menores, treinamento do modelo preditivo em cada parte menor e combinação dos modelos preditivos.¹⁰³

“Ensemble Selection” provê um benefício adicional: não requer nenhum técnico humano na seleção de um algoritmo de aprendizagem apropriado ou na escolha do conjunto de parâmetros para obter uma boa performance. Isso é realizado de forma automática pelo algoritmo, que seleciona os classificadores com melhores desempenhos para fazerem parte do conjunto de classificadores.¹⁰¹

2. JUSTIFICATIVA

2. JUSTIFICATIVA

O glaucoma é uma doença crônica, assintomática, de lenta progressão e multifatorial. O raciocínio clínico baseia-se na integração de informações do paciente a partir da anamnese, exame físico e exames complementares. Sendo assim, dados epidemiológicos (idade, raça, história familiar), avaliação anatômica (gonioscopia, do disco óptico, da camada de fibras nervosas (ex.OCT®)) e funcional (ex. perimetria acromática) nos auxiliam no diagnóstico de glaucoma.

Na literatura, há estudos^{104,105,106} que relataram que a associação entre tecnologias aumenta a capacidade diagnóstica de glaucoma. Entretanto, não há um método objetivo de integração de dados epidemiológicos, anatômicos e funcionais que facilitem o diagnóstico de glaucoma. Assim, a aplicação de sistemas de aprendizagem de máquina pode contribuir, de forma inovadora, para um importante e específico domínio do conhecimento visando resolver um problema relevante para a oftalmologia, por propor técnicas alternativas de auxílio ao diagnóstico do glaucoma.

É possível estimar sua potencial relevância social como significativa, visto que o desenvolvimento de um “software” que implemente os métodos que serão avaliados poderá disponibilizar uma nova ferramenta de suporte ao diagnóstico do glaucoma para a comunidade médica, beneficiando milhares de pacientes ao redor do mundo.

3. OBJETIVO

3. OBJETIVO

O objetivo do presente estudo é avaliar a sensibilidade e a especificidade de Sistemas de Aprendizagem de Classificadores Híbridos utilizando dados do SD-OCT e Perimetria Automatizada Acromática para o diagnóstico de glaucoma.

4. MÉTODOS

4.1. POPULAÇÃO

Para realização deste estudo, foram recrutados indivíduos glaucomatosos do Serviço de Glaucoma da Universidade de Campinas (UNICAMP), Campinas-SP, Brasil. Funcionários da universidade e acompanhantes dos pacientes foram convidados para participarem do estudo, formando o grupo controle (normal). Obteve-se por escrito o consentimento de todos os participantes, após esclarecimentos dos mesmos quanto aos aspectos éticos e clínicos do estudo.

O presente estudo foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da Universidade de Campinas (parecer 406/2005, Anexo 1). O protocolo seguiu os princípios da Declaração de Helsinque, assim como todas as regulamentações vigentes.

4.1.1 CRITÉRIOS DE INCLUSÃO PARA TODOS OS PARTICIPANTES

Os seguintes critérios foram utilizados para a inclusão de todos os participantes:

- Ter idade superior a 40 anos.
- Acuidade visual corrigida superior a 20/40 em escala de Snellen.
- Refração esférica entre ± 5.0 D.
- Refração cilíndrica entre ± 3.0 D.
- Ângulo aberto à gonioscopia: visualização do esporão escleral em 360 graus e ausência de goniossinéquias.
- Campos visuais confiáveis: índices de respostas falso-positivas $< 33\%$, índice de respostas falso-negativas $< 33\%$ e índice de perdas de fixação $< 20\%$.

- Exame de OCT confiáveis: centrados, sem artefatos de movimento e com sinal superior ou igual a 7.
- Quando os dois olhos eram elegíveis para inclusão, selecionou-se apenas um aleatoriamente.

4.1.2 CRITÉRIOS DE EXCLUSÃO PARA TODOS OS PARTICIPANTES

Os seguintes critérios de exclusão foram utilizados a todos os participantes:

- Retinopatia diabética
- Doença retiniana
- Uveíte
- Neuropatia óptica não glaucomatosa
- Pseudofacia
- Afacia
- Catarata definida pelos critérios “*Lens Opacification Classification System III*” (LOCS-III) com opacidade máxima nuclear (NC3, NO3), cortical (C3) e subcapsular (P3)(Anexo 2).¹⁰⁷

4.1.3 CRITÉRIOS DE INCLUSÃO PARA INDIVÍDUOS NORMAIS

Os seguintes critérios de inclusão foram aplicados aos indivíduos normais:

- Ausência de história familiar de glaucoma.
- Pressão intraocular ≤ 21 mmHg, sem história de elevação da mesma.
- Dois campos visuais (perimetria automática acromática) consecutivos e confiáveis sem as alterações descritas abaixo.
- Disco óptico de aspecto normal.

4.1.4 CRITÉRIOS DE INCLUSÃO PARA INDIVÍDUOS GLAUCOMATOSOS

Os seguintes critérios de inclusão foram aplicados aos indivíduos glaucomatosos:

- Duas ou mais medidas da pressão intraocular acima de 21mmHg.
- Alteração de campo visual sugestiva de glaucoma

Para caracterizar alteração de campo visual, o paciente devia apresentar¹⁰⁸:

Defeito de campo visual glaucomatoso confirmado em dois exames confiáveis e consecutivos com presença de pelo menos dois dos critérios seguintes:

- Agrupamento de 3 pontos com probabilidade < 5% no gráfico “Pattern Deviation” com pelo menos 1 ponto com probabilidade < 1%.
- “Glaucoma Hemifield Test” (GHT) fora dos limites de normalidade de 99% específica para a idade.
- “Pattern Standard Deviation” (PSD) fora dos limites de normalidade de 95% específica para a idade.

A gravidade do glaucoma foi baseada no valor do “Mean Deviation” (MD) e definida de acordo com os seguintes critérios:¹⁰⁹

- Glaucoma inicial: $MD \geq -6dB$
- Glaucoma moderado: MD entre -6dB e -12dB (inclusive)
- Glaucoma avançado: $MD < -12dB$

Apenas pacientes com glaucoma inicial e moderado foram incluídos no estudo.

- Disco óptico de aspecto glaucomatoso, definido pela presença de pelo menos dois dos seguintes critérios:
 - Escavação do disco maior que 0,6.

- Defeito focal de rima neuroretiniana.
- Hemorragia peripapilar.
- Fosseta adquirida.

4.2. MÉTODOS

Este estudo observacional, transversal, incluiu 110 olhos de 110 participantes (sendo 62 pacientes com glaucoma e 48 indivíduos saudáveis como controle) no período entre agosto 2008 e novembro de 2010. Todos participantes foram submetidos ao exame oftalmológico completo que incluiu:

- 4.2.1. Anamnese completa: história da doença atual, histórico das doenças pregressas, histórico das alterações oculares pregressas, uso de medicações sistêmicas e oculares, histórico familiar para a presença de glaucoma e outras doenças.
- 4.2.2. Medida da acuidade visual com a melhor correção da refração em escala Snellen. Após a inclusão, a acuidade visual era também transformada em escala LogMar de acordo com a tabela do Anexo 3.¹¹⁰
- 4.2.3. Biomicroscopia em lâmpada de fenda modelo Haaig-Streit 900 (Haaig-Streit, Koeniz, Suíça).
- 4.2.4. Medida da pressão intraocular com tonômetro de aplanção de Goldman modelo R900 (Haaig-Streit, Koeniz, Suíça).
- 4.2.5. Gonioscopia com lente de quatro espelhos modelo Volk G4 com e sem indentação.
- 4.2.6. Exame do disco óptico sob midríase com lente de 78-D Volk à lâmpada de fenda.

- 4.2.7. Perimetria Automatizada Acromática utilizando a estratégia “Swedish Interactive Threshold Algorithm” (SITA), programa 24-2, do aparelho Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc., Dublin, CA).
- 4.2.8. Exame da camada de fibras nervosas peripapilar com aparelho Cirrus HD-OCT (Cirrus HD-OCT; Carl Zeiss Meditec Inc., Dublin, California) utilizando o programa “*optic disc cube 200x200*”.

4.3. TOMOGRAFIA DE COERÊNCIA ÓPTICA (OCT)

Os participantes foram submetidos ao exame da camada de fibras nervosas peripapilar utilizando o aparelho de imagem comercialmente disponível Cirrus HD-OCT (software version 3.0, Carl Zeiss Meditec, Inc.), que usa a tecnologia de domínio espectral. Foi usado o programa “*optic disc cube 200x200*”, que consiste em 200 A-scans derivados de 200 B-scans cobrindo uma área peripapilar de 6mm² centrados no disco óptico. Após a criação do mapa da espessura das camadas nervosas do cubo tridimensional, o “software” determina automaticamente o centro do disco e extrai um círculo peripapilar (com raio de 1.73mm) para medida da espessura da camada de fibras nervosas.

As imagens foram obtidas por dois examinadores experientes e treinados, que eram mascarados quanto ao diagnóstico do paciente.

A estratégia “*optic disc cube 200x200*” fornece vários parâmetros e análises, já descritos na seção 1.3. No presente estudo, foram utilizados dados de 17 parâmetros, a saber: espessura global, divisão da espessura em 4 quadrantes (superior, inferior, nasal e temporal), e divisão da espessura em 12 fragmentos, como horas de relógio. Todos os parâmetros foram utilizados na análise do estudo e foram alinhados de acordo com

orientação para olho direito. Assim, a medida peripapilar das 9 horas do relógio representa o lado temporal para ambos os olhos. Apenas exames bem centrados, sem artefatos de movimentação do olho e com sinal de qualidade, fornecida pelo aparelho, igual ou superior a 7 foram incluídos.

4.4. PERIMETRIA AUTOMATIZADA ACROMÁTICA (PAA)

Todos os exames de perimetria foram realizados com aparelho Humphrey Field Analyzer II (Carl Zeiss Meditec, Inc., Dublin, CA) e estratégia “Swedish Interactive Threshold Algorithm” (SITA), programa 24-2. Foram realizados no mínimo dois exames confiáveis e reprodutíveis, com intervalo máximo de três meses entre os exames e a inclusão no estudo. Os parâmetros utilizados para análise dos classificadores foram: “Mean Deviation” (MD), “Pattern Standard Deviation” (PSD) e “Glaucoma Hemifield Test” (GHT). Como o GHT só fornece dados categóricos, atribuímos valores para os resultados: dentro dos limites de normalidade: 1, limítrofe: 2, e fora dos limites de normalidade: 3.

4.5. CLASSIFICADORES DE APRENDIZAGEM DE MÁQUINA (CAMs)

Baseado nos dados obtidos com os exames de OCT e perimetria, classificadores de aprendizagem de máquina foram desenvolvidos utilizando os seguintes algoritmos:

- 1- “Bagging” (BAG)
- 2- “Naive-Bayes” (NB)
- 3- “Multilayer Perceptron” (MLP)
- 4- “Radial Basis Function” (RBF)
- 5- “Random Forest” (RAN)

- 6- “Ensemble Selection” (ENS)
- 7- “Classification Tree” (CTREE)
- 8- “Ada Boost M1” (ADA)
- 9- “Support Vector Machine Linear” (SVMl)
- 10- “Support Vector Machine não linear (Gaussian)” (SVMg)

Os classificadores foram desenvolvidos pelo método de aprendizado supervisionado, cujos exemplos de treinamentos foram rotulados com a presença ou ausência de glaucoma. Cada exemplo foi descrito por um vetor de valores de características (atributos ou parâmetros). Inicialmente, os CAMs foram treinados utilizando todos os 17 parâmetros do OCT e os 3 parâmetros da PAA, totalizando 20. Subsequentemente, realizou-se o processo de redução de dimensionalidade, que consistia em uma seleção retrógrada com eliminação dos parâmetros que apresentavam menores aROCS. Logo que um parâmetro era excluído, fazia-se a análise da aROC com os parâmetros restantes até que fossem eliminados todos, restando apenas o parâmetro que apresentava a maior aROC isoladamente. Este processo foi realizado para cada 1 dos 10 classificadores de aprendizagem de máquina e tinha como objetivo determinar a maior aROC com o menor número de parâmetros para cada classificador.

Os classificadores foram desenvolvidos usando “software” WEKA versão 3.7.0 (“Waikato Environment for Knowledge Analysis,” Universidade de Waikato, Nova Zelândia). Os dados foram coletados no Serviço de Glaucoma da Universidade de Campinas (UNICAMP) e depois transferidos para o Laboratório de Conhecimento em Engenharia (KNOMA) do Programa de Computação e Sistemas Digitais da Universidade de São Paulo (USP) para análise.

4.6. VALIDAÇÃO CRUZADA (“CROSS VALIDATION”)

Para a o desenvolvimento dos CAMs, é idealmente necessário treinar com uma população de indivíduos e testá-los em uma população diferente. Porém, para realizar este processo, é necessário uma população de indivíduos glaucomatosos e controles normais muito grande, o que limita o estudo. Para suprir esta deficiência e diminuir os vieses de seleção, utiliza-se a técnica de “10-fold cross validation” (validação cruzada).

A técnica de validação cruzada, ou “cross-validation”,¹¹¹ é uma das formas mais utilizadas para estimar a taxa de acertos de classificadores. Uma opção simplista de verificar a taxa de acerto de um método para uma base de dados de treinamento com n instâncias é gerar um classificador utilizando a totalidade da base do conjunto de treinamento e utilizar a mesma totalidade da base para testar o classificador. Essa técnica simplista, com certeza, irá gerar um valor de taxa de acerto falsamente elevada pois o mesmo conjunto de instâncias foi utilizado para gerar o classificador e para testá-lo.⁹² A validação cruzada aparece como uma opção mais sofisticada que busca evitar os problemas de uma técnica tão simplista.

Na técnica de validação cruzada, divide-se a base de dados com n instâncias em f amostras (“folds”) aleatoriamente escolhidas, cada uma com tamanho igual a n/f instâncias, ou um tamanho aproximado caso n não seja divisível por f .¹¹²

De forma genérica, a validação cruzada repete f vezes um processo de treinamento e teste. Em cada uma dessas vezes, utiliza-se uma das f amostras aleatórias para serem validadas pelo treinamento feito com as outras $f - 1$ amostras. Cada uma dessas validações

retornará um resultado de taxa de acerto. A taxa de acerto final será dada pela média aritmética simples dos valores de taxa de acerto encontrados para cada uma das amostras.

Desta forma, uma base de dados com 110 instâncias (população do presente estudo), a qual se aplicará a técnica de validação cruzada com $f=10$ dará origem a 10 amostras com 11 instâncias. O processo de validação cruzada consistirá de 10 etapas. Para cada uma delas, 9 das amostras serão utilizadas como conjunto de treinamento, ou seja, uma base de treinamento com 99 instâncias. Essa base de treinamento será utilizada para gerar, através de um determinado método, um classificador que será utilizado para classificar as 11 instâncias da amostra que não foram consideradas na base de treinamento. Por fim, a média de cada um dos 10 testes conduzidos será considerado como a taxa de acerto obtida para a base de dados.⁹²

4.7. WEKA

Weka (“Waikato environment for Knowledge Analyses”, Universidade de Waikato, Nova Zelândia) é uma suíte de mineração de dados muito popular no meio acadêmico e que oferece uma grande variedade de ferramentas para propósitos de aprendizado de máquinas.^{113,114,115,116,117} Ele é desenvolvido utilizando linguagem Java, que tem como principal característica a portabilidade, permitindo utilizá-lo em diferentes formas operacionais.¹¹⁸ Ele provê algoritmos para três tipos de problemas: classificação ou regressão, agrupamento e seleção de atributos.

Existe uma grande variedade de algoritmos. Nenhum deles é superior em todos os problemas de mineração de dados.¹¹⁹ O algoritmo necessita encontrar a estrutura do problema para obter informações úteis ou um modelo mais apurado. O principal objetivo do

desenvolvimento do “software” Weka foi permitir a máxima flexibilidade ao testar o método de aprendizagem de máquinas ou novos algoritmos. Weka provê diversos métodos para construir e analisar algoritmos, e são disponíveis em uma interface comum. Isso torna fácil a comparação de diferentes estratégias para solução do problema baseado no mesmo método de avaliação e identifica aquele mais apropriado para o problema em questão.¹²⁰

Os algoritmos de aprendizagem podem ser avaliados usando validação cruzada (“cross validation”). O Weka provê medidas de performance numérica (exemplo acurácia, erro quadrático médio), além de médias gráficas para visualização da performance do classificador (curva ROC e curva “precision recall”).¹²⁰

4.8. SENSIBILIDADE

A sensibilidade (Se) de um exame diagnóstico mede a proporção de pacientes com determinada doença que conseguem ser identificados positivamente a partir do resultado de um teste. A medida da sensibilidade é expressa através de um número no intervalo entre 0 e 1, pode ser calculada dividindo-se a quantidade de resultados verdadeiros positivos pela quantidade total de pacientes com a doença, ou seja, o somatório de verdadeiros positivos e falsos negativos.

$$Se = \frac{VP}{VP+FN}$$

4.9. ESPECIFICIDADE

A especificidade (Sp) de um exame diagnóstico mede a proporção de pacientes normais (sem a doença) que conseguem ser identificados como tal por meio de um teste. Assim como a sensibilidade, a medida da especificidade é igualmente expressa através de

um número no intervalo entre 0 e 1, sendo calculada através da razão entre a quantidade de resultados verdadeiros negativos e a quantidade total de pacientes normais, ou seja, o somatório de verdadeiros negativos e falsos positivos.

$$Sp = \frac{VN}{VN+FP}$$

A sensibilidade e especificidade de um teste diagnóstico dependem não apenas da qualidade do exame em si, mas dos critérios de normalidade e anormalidade adotados e da população estudada.

4.10. CURVA ROC

A análise ROC – do inglês “receiver operating characteristic” – é um método gráfico para avaliação, organização e seleção de sistemas de diagnósticos e/ou predição. Recentemente, a análise ROC foi introduzida também em aprendizagem de máquina (AM) e mineração de dados como uma ferramenta útil e poderosa para avaliação de modelos de classificação,^{121,122} e também tem sido utilizada para construção e refinamento de modelos.¹²³

O gráfico ROC é baseado na probabilidade de detecção, ou taxa de verdadeiros positivos (sensibilidade), e na probabilidade de falsos alarmes, ou taxa de falsos positivos (1-especificidade) para diferentes valores de limiares de um teste diagnóstico. Para se construir o gráfico ROC, plota-se no eixo das ordenadas (eixo X) valores de 1-especificidade (ou falsos positivos) e no eixo das abcissas (eixo Y) a sensibilidade para diferentes limiares. A curva ROC demonstra a relação entre a sensibilidade e especificidade, onde aumentos na sensibilidade são acompanhados de diminuição na

especificidade. Alguns pontos no espaço ROC merecem destaque (Figura 22). O ponto A (0,0) representa a estratégia de nunca classificar um exemplo como positivo. Modelos que correspondem a esse ponto não apresentam nenhum falso positivo, mas também não conseguem classificar nenhum verdadeiro positivo. A estratégia inversa, de sempre classificar um novo exemplo como positivo, é representada pelo ponto B (100%, 100%). O ponto C (0,100%) representa o modelo perfeito, ou seja, todos os exemplos positivos e negativos são corretamente classificados. O ponto D (100%,0) representa o modelo que sempre faz previsões erradas.¹²⁴

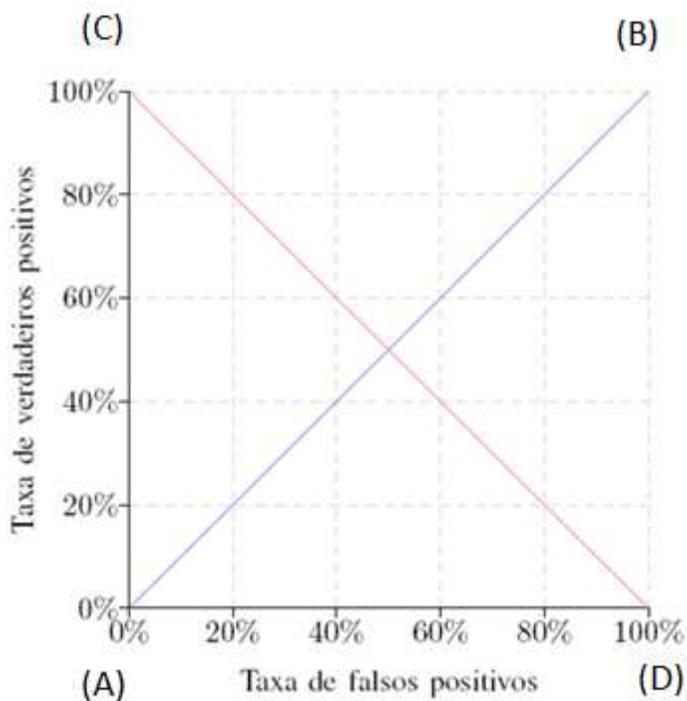


Figura 22: Representação do gráfico ROC e dos pontos A (0,0), B (100,100), C(0,100) e D (100,0)¹²⁴

O desempenho de um classificador é representado por uma curva no gráfico ROC – a curva ROC -. Para comparar dois classificadores basta comparar suas curvas, caso não haja nenhuma intersecção, a curva que mais se aproxima do ponto C é a de melhor desempenho (Figura 23). Caso haja intersecções, cada um dos sistemas tem uma faixa operacional na qual é melhor que o outro (Figura 24). Outra maneira para avaliar e comparar sistema ou classificadores é comparar a área abaixo da curva ROC – aROC. Uma vez que a área da curva ROC é uma fração da área de um quadrado de lado 1, o seu valor está sempre entre 0 e 1. Classificadores com aROCs mais próximos de 1 apresentam um melhor desempenho.¹²¹ A aROC vem gradativamente ganhando espaço como medida de avaliação de modelos em aprendizado de máquina e mineração de dados. Por isso a aROC foi o método escolhido neste trabalho para comparar o desempenho entre os classificadores.

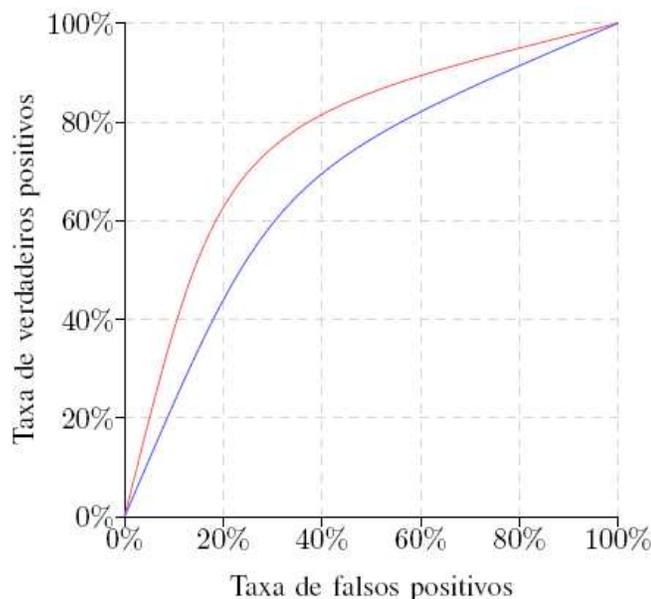


Figura 23. Exemplo de comparação de duas curvas ROCs sem intersecção (curvas vermelho e azul). A curva vermelha está mais próxima do ponto (0,100) que a curva azul, por isso, o modelo que é representado por esta curva apresenta melhor desempenho que o modelo representado pela curva azul.¹²⁴

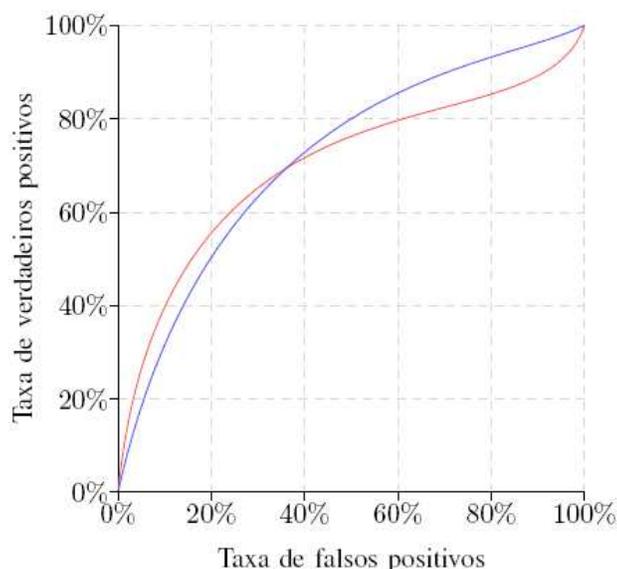


Figura 24. Exemplo de comparação de duas curvas ROCs com intersecção (curvas vermelho e azul). Próximo ao ponto (40%,70%) há a intersecção das duas curvas. Analisando as curvas à esquerda do ponto (40%,70%), a curva vermelha está mais próxima do ponto (0,100%) que a curva azul, por isso, o modelo que é representado por esta curva apresenta melhor desempenho que o modelo representado pela curva azul. Analisando as curvas à direita do ponto (40%,70%) ocorre o inverso, melhor desempenho do modelo representado pela curva azul.¹²⁴

4.10. ANÁLISE ESTATÍSTICA

Todas as análises foram realizadas utilizando “software” MedCalc versão 11.0.1 (MedCalc Software, Mariakerke, Bélgica). As variáveis contínuas foram comparadas usando o teste T de “Student”, enquanto as variáveis categóricas foram analisadas usando o Teste de Qui-quadrado ou Teste Exato de Fisher.

Curvas “Receiver Operating Characteristic” (ROC) foram obtidas para todos os parâmetros de SD-OCT (espessura média, 4 quadrantes, e 12 horas de relógio), parâmetros de campo visual (MD, PSD e GHT), classificadores de aprendizagem de máquina com todos os 20 parâmetros e com dados otimizados (retirados os parâmetros com menores aROCs). Áreas sob a curva ROC (aROCs) foram calculadas e foram comparadas usando o Teste Z. Valores de $P < 0.05$ foram considerados estatisticamente significantes.

5. RESULTADOS

5. RESULTADOS

Cento e dez olhos de 110 indivíduos foram incluídos no estudo (anexos 4 e 5). Destes 110 indivíduos, 62 pacientes tinham diagnóstico de glaucoma e 48 eram indivíduos normais (controles).

A Tabela 8 demonstra as características clínicas dos indivíduos incluídos no estudo. A idade média foi de $57,0 \pm 9,2$ anos para os indivíduos normais e de $59,9 \pm 9,0$ para os pacientes glaucomatosos ($p=0,103$). Dos indivíduos normais, 23 (47,9%) eram do sexo masculino e 25 (52,1%) do sexo feminino; entre os indivíduos glaucomatosos, a proporção de indivíduos do sexo masculino e feminino foi de 50% (31 sexo masculino e 31 sexo feminino) ($p=0,830$).

A distribuição da raça nos dois grupos não apresentou diferença estatisticamente significativa ($p=0,461$). Foram classificados como da raça branca (leucoderma) 37 (77,1 %) do grupo controle e 46 (74,2%) do grupo glaucoma e da raça negra (melanoderma) 11 (22,9%) do grupo controle e 16 (25,8%) do grupo glaucoma.

A acuidade visual (AV), medida pela escala LogMAR, e a correção óptica, medida pelo equivalente esférico (EE), não apresentaram diferença estatisticamente significativa ($p>0,05$) entre os grupos normal e glaucoma.

A pressão intraocular média não apresentou diferença estatisticamente significativa entre os grupos controle e glaucoma ($14,8 \pm 2,8$ mmHg e $13,8 \pm 2,5$ mmHg, respectivamente) ($p=0,062$), entretanto os pacientes glaucomatosos usavam maior número de medicações hipotensoras ($2,2 \pm 1,2$), enquanto os indivíduos controle não usavam medicação ($p<0,001$).

Com relação aos índices do exame de perimetria, o “Mean Deviation” (MD) foi significativamente menor ($p < 0,001$) no grupo glaucoma que no grupo controle ($-4,1 \pm 2,4$ dB e $-1,5 \pm 1,6$ dB, respectivamente). O “Pattern Standard Deviation” (PSD) foi significativamente maior ($p < 0,001$) no grupo glaucoma que no grupo controle ($4,3 \pm 2,4$ dB e $1,8 \pm 0,8$ dB, respectivamente). Entre os pacientes glaucomatosos 51 (82,3%) apresentavam glaucoma em fase inicial e 11 (17,7%) glaucoma moderado. (Tabela 8)

Tabela 2: Aspectos clínicos e demográficos (idade, sexo, raça, acuidade visual, equivalente esférico, pressão intraocular, uso de medicações, “Mean Deviation”, “Pattern Standard Deviation”) dos indivíduos normais e glaucomatosos e a significância estatística da diferença entre os grupos.

| | Normal (n=48) | Glaucoma (n=62) | Valor p |
|------------------------------|---------------------|---------------------|---------|
| Idade (anos; média±DP) | 57,0 ± 9,2 | 59,9 ± 9,0 | 0,103 |
| Variação | 45-82 | 43-78 | - |
| Sexo (Masc[%]:Fem[%]) | 23[47,9] : 25[52,1] | 31[50] : 31[50] | 0,830 |
| Raça (branca[%]:negra[%]) | 37[77,1] : 11[22,9] | 46[74,2] : 16[25,8] | 0,461 |
| AV LogMAR(média ±DP) | 0,0 ± 0,0 | 0,1 ± 0,1 | 0,800 |
| EE (D; média ± DP) | 1,53 ± 2,15 | 1,2 ± 1,9 | 0,467 |
| PIO (mmHg; média ±DP) | 14,8 ± 2,8 | 13,8 ± 2,5 | 0,062 |
| Medicações (média ± DP) | 0 | 2,2 ± 1,2 | <0,001 |
| MD (dB; média ± DP) | -1,5 ± 1,6 | -4,1 ± 2,4 | <0,001 |
| PSD (dB; média ± DP) | 1,8 ± 0,8 | 4,3 ± 2,4 | <0,001 |

AV= acuidade visual; EE = equivalente esférico; PIO= pressão intraocular; D = dioptrias; dB = decibels; DP = desvio padrão; MD = “Mean Deviation”; PSD = “Pattern Standard Deviation”; Masc= masculino; Fem= feminino.

A Tabela 9 mostra o desempenho (medido através da área abaixo da curva ROC) dos aparelhos de perimetria acromática e OCT usados isoladamente e sem utilizar os classificadores de aprendizagem de máquina. Os melhores parâmetros da perimetria foram: PSD (aROC=0,915), GHT (aROC=0,866) e MD (aROC=0,828). Os parâmetros do OCT que apresentaram as maiores aROCs foram : quadrante inferior (aROC=0,813), espessura global média (aROC=0,807), posição das 7 horas (aROC=0,765) e posição das 6 horas (aROC=0,754).

Para uma especificidade fixa de 80%, as maiores sensibilidades observadas com os parâmetros da perimetria foram: 91,9 % para GHT, 88,7% para PSD e 62,9% para MD. Entre os parâmetros do OCT, os que apresentaram melhores sensibilidades com especificidade fixa em 80% foram: posição das 7 horas (64,5%) espessura global média (62,1%), quadrante inferior (61,3%), e quadrante superior (57,6%).

Tabela 3: Áreas abaixo da curva ROC (aROC), sensibilidade (%) com especificidades fixas de 80% e 90% para cada parâmetro da perimetria acromática (MD, PSD, GHT) e os parâmetros do SD-OCT (espessura global, 4 quadrantes e 12 segmentos em horas de relógio) sem utilização dos classificadores de aprendizagem de máquina.

| | aROC (95% IC) | Especificidade 80% | Especificidade 90% |
|------------------------|---------------------|--------------------|--------------------|
| PAA | | | |
| MD | 0,828 (0,745-0,894) | 62,9 | 56,4 |
| PSD | 0,915 (0,846-0,959) | 88,7 | 75,1 |
| GHT | 0,866 (0,787-0,923) | 91,9 | 70,9 |
| OCT | | | |
| Espessura global | 0,807 (0,721-0,876) | 62,1 | 54,0 |
| Quadrante | | | |
| Temporal | 0,675 (0,579-0,761) | 50,0 | 33,1 |
| Superior | 0,737 (0,645-0,816) | 57,6 | 46,7 |
| Nasal | 0,685 (0,590-0,771) | 49,2 | 26,6 |
| Inferior | 0,813 (0,727-0,881) | 61,3 | 53,2 |
| Hora de relógio | | | |
| 1 | 0,703 (0,608-0,786) | 45,2 | 29,8 |
| 2 | 0,723 (0,630-0,804) | 51,6 | 34,7 |
| 3 | 0,574 (0,476-0,668) | 29,8 | 22,6 |
| 4 | 0,605 (0,507-0,696) | 29,0 | 11,3 |
| 5 | 0,671 (0,575-0,757) | 43,5 | 27,4 |
| 6 | 0,754 (0,663-0,831) | 47,6 | 33,9 |
| 7 | 0,765 (0,674-0,840) | 64,5 | 38,7 |
| 8 | 0,631 (0,534-0,721) | 42,7 | 27,4 |
| 9 | 0,625 (0,527-0,715) | 44,3 | 36,3 |
| 10 | 0,699 (0,604-0,783) | 50,8 | 42,0 |
| 11 | 0,740 (0,648-0,819) | 40,3 | 33,9 |
| 12 | 0,672 (0,576-0,759) | 38,7 | 21,0 |

PAA= perimetria acromática automatizada; MD= Mean Deviation; PSD= Pattern Standard Deviation; GHT=Glaucoma Hemifield Test; IC= intervalo de confiança.

A análise do desempenho dos classificadores de aprendizagem de máquina é mostrada na Tabela 10. Observa-se que as aROCs variam de acordo com o tipo de classificador e com o número de parâmetros utilizados para análise: o primeiro valor corresponde ao obtido com todos os 20 parâmetros (17 do OCT e 3 da Perimetria); o segundo valor representa a maior aROC com o menor número de atributos. Vale ressaltar que, para cada classificador, observou-se melhor desempenho com número diferente de atributos. Combinando todos os atributos, as aROCs variaram de 0,777 (CTREE) a 0,933 (RAN). Os classificadores que apresentaram maiores aROCs combinando todos os atributos analisados foram: RAN (0,933), ADA (0,932) e SVM (0,929). Avaliando o desempenho dos classificadores após otimização (exclusão dos atributos com menores AROCs), as aROCs variaram entre 0,879 (CTREE, 2 atributos) e 0,946 (RAN, 4 atributos). Apresentaram melhor acurácia diagnóstica os seguintes classificadores: RAN (0,946, 4 atributos), SVM (0,938, 12 atributos), MLP (0,934, 5 atributos) e RBF (0,934, 4 atributos). No Anexo 6 apresentamos a tabela com as aROCs dos CAMs com todos os parâmetros e com a exclusão gradativa dos parâmetros que apresentavam piores desempenhos (redução de dimensionalidade).

Tabela 4: Avaliação das aROCs dos classificadores de aprendizagem de máquina (BAG, NB, MLP, RBF, RAN, ENS, CTREE, ADA, SVMg e SVMl) usando dados da perimetria e OCT. Foram avaliados: as aROCs, utilizando todos os 20 atributos, a melhor aROC com o menor número de parâmetros, e a sensibilidade (%) com especificidades fixas de 80% e 90%.

| CAM | aROC (20 atributos) | aROC (# atributos) | Especificidade 80% | Especificidade 90% |
|-------|------------------------|-----------------------|--------------------|--------------------|
| BAG | 0,893 | 0,913(8) | 91,93 | 74,19 |
| NB | 0,912 | 0,928 (5) | 93,54 | 83,87 |
| MLP | 0,845 | 0,934 (5) | 90,32 | 77,41 |
| RBF | 0,857 | 0,934 (4) | 93,54 | 79,03 |
| RAN | 0,933 | 0,946 (4) | 95,16 | 82,25 |
| ENS | 0,910 | 0,910 (20) | 90,32 | 79,03 |
| CTREE | 0,777 | 0,879 (2) | 82,25 | 56,45 |
| ADA | 0,932 | 0,933 (15) | 93,54 | 83,87 |
| SVMg | 0,913 | 0,930 (8) | 93,54 | 83,87 |
| SVMl | 0,929 | 0,938 (12) | 93,54 | 83,87 |

CAM- classificador de aprendizagem de máquina; aROC – área abaixo da curva ROC ; BAG – “Bagging”; NB – “Naive-Bayes”; MLP – “Multilayer Perceptron”; RBF – “Radial Basis Function”; RAN – “Random Forest”; ENS – “Ensemble Selection”; CTREE – “Classification Tree”; ADA – “Ada Boost M1”; SVMl – “Support Vector Machine Linear”; SVMG “Support Vector Machine Gaussian”.

A Tabela 11 revela a comparação das áreas abaixo da curva ROC de todos os classificadores de aprendizagem de máquina utilizando dados combinados do OCT e PAA (OCT+CV), otimizados, com os melhores parâmetros da perimetria acromática (PSD) e do OCT (região inferior). Todos os CAMs OCT+CV apresentaram aROCs significativamente maiores que o do melhor parâmetro do OCT ($P < 0.05$), com exceção do classificador CTREE ($p = 0,2$). Quando comparados com o melhor parâmetro da perimetria (PSD), os CAMs OCT+CV não apresentaram aROCs significativamente diferentes ($p > 0,05$).

Tabela 5: Comparação das áreas abaixo da curva ROC dos classificadores OCT+CV (com número de parâmetros que apresentaram o melhor desempenho) com os melhores parâmetros da perimetria (PSD) e do OCT (região inferior).

| CAM OCT+CV | Melhor aROC | IC (95%) | p (PSD)* | p (I)* |
|---------------|-------------|-------------|----------|--------|
| NB | 0,928 | 0,879-0,977 | 0,7244 | 0,0147 |
| MLP | 0,934 | 0,887-0,981 | 0,5991 | 0,0094 |
| SVMI | 0,938 | 0,893-0,983 | 0,5192 | 0,0068 |
| SVMg | 0,930 | 0,882-0,978 | 0,6825 | 0,0127 |
| RBF | 0,934 | 0,887-0,981 | 0,5991 | 0,0094 |
| ADA | 0,933 | 0,886-0,980 | 0,6197 | 0,0101 |
| BAG | 0,913 | 0,859-0,967 | 0,9588 | 0,0394 |
| ENS | 0,910 | 0,855-0,965 | 0,8979 | 0,0470 |
| CTREE | 0,879 | 0,815-0,943 | 0,3956 | 0,2003 |
| RAN | 0,946 | 0,904-0,988 | 0,3713 | 0,0034 |

CAM- classificador de aprendizagem de máquina; aROC – área abaixo da curva ROC ; BAG – “Bagging”; NB – “Naive-Bayes”; MLP – “Multilayer Perceptron”; RBF – “Radial Basis Function”; RAN – “Random Forest”; ENS – “Ensemble Selection”; CTREE – “Classification Tree”; ADA – “Ada Boost M1”; SVML – “Support Vector Machine Linear”; SVMG - “Support Vector Machine Gaussian”, IC:intervalo de confiança; p<0.05 considerado estatisticamente significativo.

*= Teste Z.

A Figura 25 mostra a representação gráfica das curvas ROC dos melhores parâmetros da perimetria acromática (PSD), do OCT (região inferior) e do melhor classificador com atributos otimizados (RAN). Nota-se uma diferença importante entre o classificador RAN treinado com 4 parâmetros (aROC=0,946) e o melhor parâmetro do OCT (aROC=0,813 Inferior) ($p < 0.05$), porém as curvas geradas pelo classificador RAN e o melhor parâmetro da PAA (PSD, aROC=0,915) se situam bem próximas ($p = 0,37$).

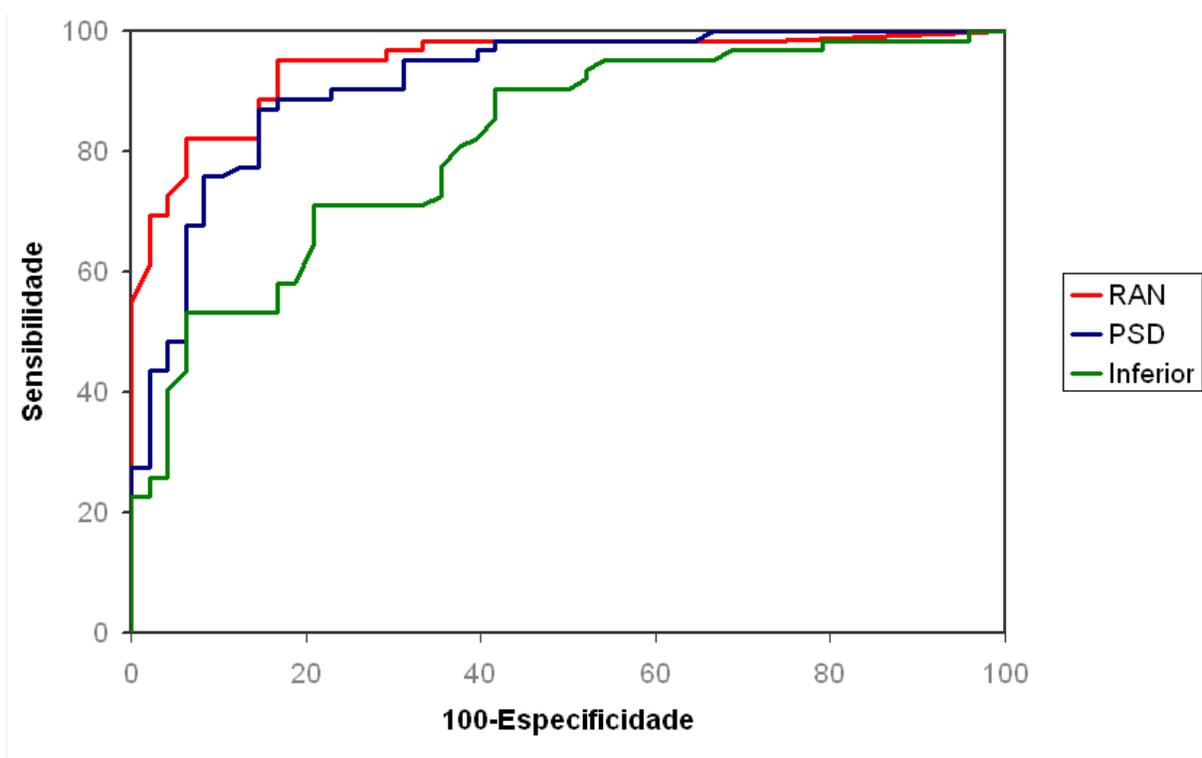


Figura 25: Representação das curvas ROC dos parâmetros que apresentaram maiores aROCs do classificador de aprendizagem de máquina usando dados combinados do OCT e perimetria (RAN=0,946), do OCT (região inferior=0,813) e da perimetria acromática (PSD=0,915).

6. DISCUSSÃO

6. DISCUSSÃO

O presente estudo investigou a sensibilidade e especificidade dos classificadores de aprendizagem de máquina empregando dados obtidos do “Spectral Domain–OCT”. Trata-se de estudo pioneiro, que empregou pela primeira vez dados funcionais e estruturais fornecidos pela nova versão do OCT combinados com dados da perimetria acromática automatizada (PAA).

O diagnóstico de glaucoma requer análise estrutural e funcional e a combinação de testes funcionais e estruturais podem melhorar a detecção de glaucoma. Lauande-Pimentel et al.¹⁰⁶ mostraram melhora da acurácia diagnóstica quando dados estruturais (utilizando dados da polarimetria de varredura a laser, GDX) e funcionais (dados da PAA) foram usados em combinação. Os autores incluíram 91 indivíduos normais e 94 indivíduos com diagnóstico de glaucoma baseado na PAA e aspecto do nervo óptico. Apenas um olho por paciente, escolhido aleatoriamente, foi incluído. O parâmetro com maior aROC do GDX foi o “The Number” (aROC= 0,87) e da PAA foi o MD (aROC= 0,92); por outro lado, ao associar-se dados dos dois aparelhos (GDX+PAA), usando função discriminante linear de Fisher, obtiveram aROC de 0,98.

Shah et al.¹²⁵ também demonstraram haver melhora da acurácia diagnóstica quando se utilizam dados estruturais e funcionais associados. Eles incluíram no estudo 123 participantes, que eram classificados como glaucoma ou normal baseados em estereofotografia ou PAA. Fizeram a análise do desempenho dos aparelhos isoladamente (Stratus-OCT, GDX VCC, “Heidelberg Retina Tomography II” (HRT II), “Frequency Doubling Technology” (FDT) e Perimetria Azul e Amarelo (SWAP)) e após a combinação

entre eles (avaliação funcional + estrutural). Os autores demonstraram que a combinação de parâmetros estruturais (OCT, GDX ou HRT) e funcionais (FDT ou SWAP) levou a um aumento significativo da sensibilidade, que foi maior do que a encontrada na análise isolada dos aparelhos ($p < 0,05$). Utilizando como critério de inclusão a análise do campo visual, ao associar dados do FDT + GDX-VCC, a sensibilidade encontrada pelos autores aumentou de 41,9% para 62,8%, FDT + Stratus-OCT de 58,1% para 72,1%, e FDT + HRT-II de 58,1% para 74,4%. Houve diminuição, porém não estatisticamente significativa ($p > 0,05$), da especificidade após a associação do FDT com GDX-VCC, Stratus-OCT e HRT-II, exceto para a comparação da análise entre FDT e FDT + HRTII ($p = 0,007$). Ao associar dados do SWAP com GDX-VCC, Stratus-OCT ou HRT-II os autores encontraram aumento da sensibilidade e diminuição da especificidade (alterações estas estatisticamente significantes ($p < 0,05$)), em comparação com dados estruturais isolados, exceto para comparação com SWAP isoladamente ($p > 0,05$). A sensibilidade com a combinação do SWAP + GDX-VCC aumentou de 41,9% para 76,7%, SWAP + Stratus-OCT de 58,1% para 76,7%, e SWAP + HRT-II de 58,1% para 76,7%. A especificidade com a combinação do SWAP + GDX-VCC diminuiu de 98,3% para 84,3%, SWAP + Stratus-OCT de 98,3% para 86,2%, SWAP + HRT-II de 84,5% para 72,4%. Quando comparados com SWAP isoladamente, houve aumento da sensibilidade em relação aos dados do SWAP+ HRT II ($p < 0,05$), porém essa diferença não foi encontrada quando utilizaram dados do SWAP + OCT ($p = 0,13$) e SWAP + GDX VCC ($p = 0,13$). A especificidade da SWAP em combinação com dados estruturais (SWAP+OCT, SWAP+HRT II e SWAP+GDX VCC) diminuiu significativamente em relação aos dados estruturais isoladamente ($p < 0,05$), mas não houve decréscimo significativo quando comparados com dados isolados da SWAP ($p > 0,05$). Os autores

encontraram resultados semelhantes quando utilizaram como critério de inclusão a análise do disco óptico pela estereofotografia.

Os CAMs já foram empregados com várias tecnologias na avaliação estrutural e funcional do glaucoma, incluindo TD-OCT,^{126,127,128,129} HRT,^{130,131,132} GDX^{133,134} e PAA.^{135,136,134,137} Recentemente, Vidotti et al.¹³⁸ avaliaram o desempenho de CAMs empregando parâmetros do SD-OCT isoladamente. No entanto, os autores não encontraram diferença estatisticamente significativa ($p > 0,05$) entre a acurácia do melhor parâmetro do OCT (Inferior, aROC = 0,813) e do melhor classificador utilizando os parâmetros do OCT (BAG, aROC=0,818).

Vários estudos usaram os CAMs para combinar dados funcionais e de imagem na intenção de melhorar a acurácia diagnóstica. Bowd et al.¹²⁸ treinaram dois tipos de CAMs (“Relevance Vector Machine”, RVM e “Subspace Mixture of Gaussians”, SSMoG) usando dados do Stratus-OCT e perimetria acromática. Eles incluíram 225 participantes (156 glaucomatosos e 69 normais), que foram classificados de acordo com a aparência do disco óptico ou alteração na perimetria acromática. A idade média dos indivíduos normais ($62 \pm 9,7$ anos) era significativamente menor que a idade média dos indivíduos glaucomatosos ($66,4 \pm 10,2$ anos) ($p=0,005$), e 90% apresentavam glaucoma inicial ($MD \geq -6$ dB). Os autores observaram que o desempenho do CAM RVM usando dados do TD-OCT e PAA isoladamente (aROC RVM OCT=0,806 e aROC RVM PAA =0,815) foi semelhante aos desempenhos dos aparelhos (aROC OCT inferior=0,777 e PAA MD=0,775) ($p > 0,05$). Entretanto, o desempenho melhorou significativamente com a combinação de dados do TD-OCT e PAA (RVM OCT+PAA=0,845) ($p < 0,05$). Avaliando o CAM SSMog, o

desempenho utilizando dados isolados do OCT (aROC SSMog OCT=0,817) não apresentou melhora significativa em relação ao TD-OCT (aROC OCT inferior=0,777) ($p>0,05$), entretanto, o desempenho do CAM usando dados da PAA (aROC SSMog PAA=0,841) foi melhor que o do aparelho de PAA (aROC PAA MD=0,775) ($p<0,05$). A combinação de dados do TD-OCT e PAA usando o CAM SSmog (aROC SSMog OCT+PAA =0,869) melhorou significativamente o desempenho em relação ao CAM SSmog OCT ($p<0,05$), porém a melhora não foi estatisticamente significativa em relação ao SSmog PAA ($p>0,05$).

Mardin et al.¹³⁹ também avaliaram a combinação de dados estruturais e funcionais usando CAMs. Eles incluíram 88 indivíduos glaucomatosos e 88 normais com diagnóstico baseado no aspecto do disco óptico e no exame de perimetria acromática. Utilizaram dados do HRT e campo visual (aparelho Octopus) e testaram três diferentes CAMs (“Liminar Discriminant Analysis”, “Bagging” e “Double-Bagging”). Evidenciaram melhor performance dos CAMs com dados combinados do HRT e PAA do que com o HRT isoladamente ($p<0,05$), porém não observaram diferença em relação aos dados de PAA isoladamente. As aROCs, utilizando o classificador que apresentou melhor desempenho (“Double-bagging”), foram: PAA aROC=0,981; HRT aROC=0,855 e PAA + HRT aROC=0,977. Nosso estudo também não mostrou superioridade na acurácia obtida com CAMs usando dados do OCT e PAA em relação aos CAMs usando apenas dados da PAA ($p=0,37$). Este resultado provavelmente decorre de viés de seleção, uma vez que o campo visual foi utilizado como critério de inclusão em todos esses estudos: indivíduos normais deveriam ter CVs normais e indivíduos glaucomatosos deveriam apresentar defeito de CV caracterizado por alterações no gráfico “Pattern Deviation”, PSD ou GHT. Assim, o

desempenho dos CAMs empregando dados apenas da PAA foi, provavelmente, superestimado.

Uma maneira de reduzir a interferência deste viés de seleção seria testar os CAMs com exames funcionais e estruturais que não foram usados para a definição de glaucoma. Racette et al.¹⁴⁰ empregaram parâmetros da PAA e estereofotografia como critérios de seleção, mas utilizaram outros dispositivos (HRT e SWAP) para testar os CAMs. Os autores incluíram 212 participantes (144 glaucomatosos e 68 controles). Utilizou-se o CAM “Relevance Vector Machine” (RVM), que testou parâmetros do HRT II e SWAP isolados e em combinação. O RVM foi treinado e testado com todos os parâmetros e após a redução de dimensionalidade, com a exclusão de parâmetros que não contribuíam para melhorar o desempenho dos classificadores (otimização). Evidenciaram que o RVM apresentou melhor desempenho com parâmetros otimizados e quando foram utilizados atributos do HRT e SWAP em combinação. A aROC do RVM HRT+SWAP (0,925) foi significativamente maior que as aROCs do RVM HRT (0,878) ($p=0,002$) e RVM SWAP (0,763) ($p=0,000$), demonstrando melhora do desempenho quando se utilizam dados estruturais e funcionais em combinação para o diagnóstico de glaucoma.

Diferente dos estudos apresentados previamente, não se detectou em nosso estudo diferença estatisticamente significativa em relação à idade entre as populações dos grupos controle e glaucomatoso. A idade média foi de $57,0 \pm 9,2$ anos para os indivíduos normais e de $59,9 \pm 9,0$ anos para os pacientes glaucomatosos ($p=0,103$). A diferença estatisticamente significativa de idade entre os grupos, observada nos estudos citados, poderia causar viés no pareamento dos grupos e com isso interferir no poder diagnóstico dos CAMs, pois tanto o

OCT quanto a PAA utilizam a idade do paciente para fazerem análise comparativa com o banco de dados interno dos aparelhos.

Outro fator importante a ser analisado é o estágio da doença em que os indivíduos glaucomatosos se encontravam. A inclusão de indivíduos em estágio avançado da doença pode superestimar a acurácia dos métodos diagnósticos. Por isso, procuramos incluir indivíduos em estágio inicial e moderado da doença e excluimos indivíduos glaucomatosos em estágio avançado. Dos pacientes glaucomatosos, 51 (82,3%) apresentavam glaucoma em fase inicial (MD >-6 dB) e apenas 11 (17,7%) glaucoma moderado (MD entre -6 dB e -12 dB).

No presente estudo, foram testados e comparados 10 diferentes tipos de CAMs. Até o presente momento, não temos conhecimento de algum estudo em que tenham sido analisados tantos CAMs. A análise de vários classificadores tem por finalidade compará-los e identificar aqueles que melhor desempenham determinada tarefa. O nosso interesse em questão era o diagnóstico de glaucoma utilizando SD-OCT e PAA e todos classificadores apresentaram desempenhos superiores ($p < 0.05$) em comparação com o melhor parâmetro do OCT (inferior), com exceção do classificador CTREE ($p = 0,2$). O classificador que teve a melhor acurácia foi o “Random Forest” (RAN). Não encontramos estudos que avaliaram o desempenho desse classificador para o diagnóstico de glaucoma, por isso não foi possível a comparação com dados encontrados na literatura. O RAN é um classificador que gera vários classificadores e agrega seus resultados.⁹⁷ A agregação de diversos classificadores tem como objetivo obter uma taxa de acerto melhor do que a obtida pela aplicação de um classificador único. Esta pode ser uma explicação para a superioridade deste classificador

em relação aos outros. Outra possibilidade que pode ser aventada é ocorrência de ter ocorrido um sobre ajustamento (“overfitting”) ao conjunto de treinamento.

O presente estudo apresenta algumas limitações. Primeiramente, a performance diagnóstica pode ter sido superestimada, pois o desenho do estudo (caso-controle) cria duas populações distintas de indivíduos, normais e doentes, e excluem indivíduos suspeitos.¹⁴¹ Entretanto, este estudo deve ser visto como um estudo preliminar para comparar diferentes classificadores e variáveis (PAA e OCT). Um estudo longitudinal, acompanhando indivíduos suspeitos, serviria para determinar quais indivíduos de fato desenvolveriam glaucoma. Assim, poderia-se avaliar a capacidade dos CAMs em prever os indivíduos suspeitos que progrediriam para glaucoma.

Outra limitação do estudo foi o uso da mesma população para treinar e testar os CAMs. Idealmente, é melhor testar e treinar o classificador em populações distintas, porém torna-se necessária uma amostra muito grande para esta tarefa. Assim como em outros estudos,^{105,128,139,140,142} a validação cruzada foi empregada para contornar essa limitação. Este procedimento possibilita treinar e testar com a mesma população evitando ainda o sobre ajustamento do classificador (“overfitting”). Num estudo subsequente, testaremos os CAMs identificados nesse projeto em uma população independente.

A despeito dessas limitações, acreditamos que nosso estudo demonstrou que é possível empregar classificadores de aprendizagem de máquina e utilizar dados estruturais e funcionais combinados para aumentar a acurácia diagnóstica de glaucoma.

7. CONCLUSÃO

7. CONCLUSÃO

Baseado nos dados apresentados, concluímos que os classificadores de aprendizagem de máquina treinados com dados do Spectral Domain OCT e Perimetria Acromática Automatizada podem, com sucesso, discriminar entre olhos normais e glaucomatosos. O melhor CAM testado foi o RAN, que apresentou $aROC=0,946$ e que revelou sensibilidade de 95,16% quando a especificidade foi fixada em 80%. Os classificadores apresentaram sensibilidade e especificidade significativamente maiores quando utilizaram dados estruturais (OCT) e funcionais (PAA) associados que dados do OCT isoladamente.

8. REFERÊNCIAS BIBLIOGRÁFICAS*

* Baseadas na norma do “international Committee of Medical Journal Editors (Vancouver), 2009.

8. REFERÊNCIAS BIBLIOGRÁFICAS

-
- ¹ Weinreb RN, Khaw PT. Primary open-angle glaucoma. *The Lancet*. 2004;363:1711-20.
- ² Sommer A. Glaucoma: facts and fancies. Doyne lecture. *Eye*. 1996;10:295-301.
- ³ Glaucoma Panel, Preferred Practice Pattern: Primary Open Angle Glaucoma Suspect. San Francisco (CA): American Academy of Ophthalmology (AAO). 2005;pp3.
- ⁴ Quigley HA, Broman A. The number of persons with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90:151-56.
- ⁵ Foster PJ, Buhrmann R, Quigley HA, Johnson GJ. The definition and classification of glaucoma in prevalence surveys. *Br J Ophthalmol*. 2002;86:238-42.
- ⁶ Burgoyne CF, Downs JC, Bellezza AJ, Suh JK, Hart RT. The optic nerve head as a biomechanical structure: a new paradigm for understanding the role of IOP-related stress and strain in the pathophysiology of glaucomatous optic nerve head damage. *Prog Retin Eye Res*. 2005;24:39-74.
- ⁷ Quigley HA. Glaucoma. *The Lancet*. 2011;377(16):1367-77.
- ⁸ Quigley HA. Number of people with glaucoma worldwide. *Br J Ophthalmol*. 1996;80:389-93.
- ⁹ Congdon N, O'Colmain B, Klaver CC, Klein R, Munoz B, Friedman DS, et al. Causes and prevalence of visual impairment among adults in the United States. *Arch Ophthalmol*. 2004;122:477-85.
- ¹⁰ Thyelfors DS, Negrel AD, Pararajasegaram R, Dadzie KY. Global data on blindness. *Bull World Health Organ*. 1995;73:115-21.
- ¹¹ Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006;90(3):262-7.
- ¹² Sakata K, Sakata LM, Sakata VM, Santini C, Hoppker LM, Bernardes R, et al. Prevalence of glaucoma in a south Brazilian population: Projeto Glaucoma. *Invest Ophthalmol Vis Sci*. 2007;47(11):4974-9.
- ¹³ Tielsch JM, Sommer A, Katz J, Royall RM, Quigley HA, Javitt J. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *JAMA*. 1991;266(3):369-74.
- ¹⁴ Sommer A, Tielsch JM, Katz J, Quigley HA, Gottsch JD, Javitt JC. Racial differences in cause-specific prevalence of blindness in east Baltimore. *N Eng J Med*. 1991;325(20):1412-7.
- ¹⁵ Coffey M, Reidy A, Wormald R, Xian WX, Wright L, Courtney P. Prevalence of glaucoma in west of Ireland. *Br J Ophthalmol*. 1993;77:17-21.
- ¹⁶ Leite MT, Sakata LM, Medeiros FA. Managing glaucoma in developing countries. *Ar Bras Oftalmol*. 2011;74(2):83-4.
- ¹⁷ Sommer A, Katz J, Quigley HA, et al. Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. *Arch ophthalmol*. 1991;109:77-83.
- ¹⁸ Johnson CA, Sampl DP, Zangwill LM, et al. Structure and functional evaluation (SAFE): II. Comparison of optic disk and visual field characteristics. *Am J Ophthalmol*. 2003;135:148-54.
- ¹⁹ Quigley HA, Katz J, Derick RJ, et al. An evaluation of optic disk and nerve fiber layer examinations in monitoring progression of early glaucoma damage. *Ophthalmology*. 1992;99:19-28.

-
- ²⁰ Cioffi GA, Liebmann JM, Johnson CA, et al. Structural-functional relationships of the optic nerve in glaucoma. *J Glaucoma*. 2000;9:3-4.
- ²¹ Johnson CA, Cioffi GA, Liebmann J, et al. The relationship between structural and functional alterations in glaucoma: a review. *Semin Ophthalmol*. 2000;15:221-3.
- ²² Girkin CA. Relationship between structure of optic nerve/nerve fiber layer and functional measurements in glaucoma. *Curr Opin Ophthalmol*. 2004;15:96-101.
- ²³ Okubo K. Correlation between glaucomatous optic disc and visual field defects. IV. Mode of cupping formation. *Kobe J Med Sci*. 1986;32:197-202.
- ²⁴ Miglior S, Brigatti L, Lonati C, et al. Correlation between the progression of optic disk and visual field changes in glaucoma. *Curr Eye Res*. 1996;15:145-9.
- ²⁵ Zeyen TG, Caprioli J. Progression of disk and field damage in early glaucoma. *Arch Ophthalmol*. 1993;111:62-5.
- ²⁶ Gordon MO, Beiser JA, Brandt JD, et al. The ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open angle glaucoma. *Arch ophthalmol*. 2002;120:714-20.discussion 829-30.
- ²⁷ Hiji A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression: results from Early Manifest Glaucoma Trial. *Arch Ophthalmol*. 2002;120:1268-79.
- ²⁸ Quigley HA, Dunkelberger GR, Green WR. Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma. *Am J Ophthalmol*. 1989;107:453-64.
- ²⁹ Quigley HA, Kartz J, Derick RJ, Gilbert D, Sommer A. An evaluation of optic disk and nerve fiber layer examinations in monitoring progression of early glaucoma damage. *Ophthalmology*. 1992;99:19-28.
- ³⁰ Junior RS, Medeiros FA. O exame de perimetria automatizada. In: Junior RS, Medeiros FA. *Perimetria Computadorizada: interpretação e discussão de casos*. Rio de Janeiro: Cultura Médica; 2001.3-38.
- ³¹ Junior RS, Medeiros FA. Princípios básicos da interpretação do campo visual. In: Junior RS, Medeiros FA. *Perimetria Computadorizada: Interpretação e Discussão de Casos*. Rio de Janeiro: Cultura Médica; 2001.39-56.
- ³² Schimiti RB, Costa VP. Interpretação do campo visual computadorizado. In: Schimiti RB, Costa VP. *Perimetria computadorizada: Um Guia Básico de Interpretação*. 3ed. Rio de Janeiro: Cultura Médica; 2009.53-72.
- ³³ Sekhar GC, Naduvilath TJ, et al. Sensitivity of Swedish Interactive Threshold Algorithm compared with Standard Full Threshold Algorithm in Humphrey visual field test. *Ophthalmology*. 2000;107:1303-8.
- ³⁴ Fercher AF, Roth E. Ophthalmic Laser Interferometer. *Proc. SPIE*. 1986;651-8.
- ³⁵ Fercher AF, Mengwdoth K, Werner W. Eye length measurement by interferometer with partially coherent light. *Optics letters*. 1988;13:186-8.
- ³⁶ Fujimoto JG, De Silvestri S, Ippen EP, et al. Femtosecond optical ranging in biological systems. *Opt Lett*. 1986;11:150-2.
- ³⁷ Huang D, Swanson EA, Lin CP, et al. Optical Coherence tomography. *Science*. 1991;254 (5035):1178-81.

-
- ³⁸ Fercher AF, Hitzenberger CK, Drexler W, Kamp G, Sattmann H. In-vivo Optical Coherence Tomography. *Am J Ophthalmol.* 1993;116 (1) :113-5
- ³⁹ Swanson EA, Izatt JÁ , Hee MR et al. In-vivo retinal imaging by optical coherence tomography. *Optics Letters.* 1993;18(21)1864-6
- ⁴⁰ Malerbi FK, Andrade REA, Farah ME. OCT no diagnóstico por imagem. In: Farah ME. *Tomografia de Coerência Óptica.* 2 ed. Rio de Janeiro: Guanabara Koogan; 2010.1-8.
- ⁴¹ Castro LC, Duker JS, Filho PAAM. OCT- Spectral/ Fourier Domain (SD-OCT) e de resolução Ultra (UHR OCT): aspectos técnicos e perspectivas futuras. In: Farah ME. *Tomografia de Coerência Óptica.* 2 ed- Rio de Janeiro: Guanabara Koogan; 2010.411-24.
- ⁴² Lumbroso B, Rispoli M. *Guide to interpreting Spectral Domain Optical Coherence Tomography. SD-OCT for The Study and Management of Glaucoma.* Publisher INC Inovation News Communication. Dublin, CA. 2009;6:55-8.
- ⁴³ Budenz DL, Anderson DR, Varma R, et al. Determinants of normal retinal nerve fiber layer thickness measured by Stratus OCT. *Ophthalmology* 2007;114:1046-52.
- ⁴⁴ 510 (k) Premarket notification Cirrus HD-OCT with retinal fiber layer and macular normative database. FDA, Department of Human and Health service. Register: K08329; may, 2009.
- ⁴⁵ Kim JS, Ishikawa H, Sung KR, et al. Retinal nerve fiber layer thickness measurements reproducibility improved with spectral domain optical coherence tomography. *Br j Ophthalmol.* 2009;93:2057-63.
- ⁴⁶ Leung CK, Cheung CY, Weinreb RN, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. *Ophthalmology.* 2009;116:1257-63
- ⁴⁷ Schuman JS. Spectral domain coherence tomography for glaucoma (an AOS thesis). *Trans Am ophthalmol Soc .*2008;106:426-58.
- ⁴⁸ Vizzeri G, Weinreb RN, Gonzalez-Garcia AO, et al. Agreement between spectral-domain and time-domain OCT for measuring RNFL thickness. *Br J Ophthalmol.* 2009;93:775-81.
- ⁴⁹ Cremasco F, Massa G, Vidotti VG, Lupinacci APC, Costa VP. Intrasession, intersession and interexaminer variabilities of retinal nerve fiber layer measurements with spectral-domain OCT. *Eur J Ophthalmol.* 2011;21(3):264-70.
- ⁵⁰ Paunescu LA, Schuman JS, Price LL, et al. Reproducibility of nerve fiber thickness, macular thickness, and optical nerve head measurements using Stratus OCT. *Invest Ophthalmol Vis Sci.* 2004;45:1716-24.
- ⁵¹ Sung KR, KimDY, Park SB, et al. Comparison of retinal fiber layer thickness measured by Cirrus HD and Stratus optical coherence tomography. *Ophthalmology.* 2009;116:1264-70.
- ⁵² Chang RT, Knight OJ, Feuer WJ, et al. Sensitivity and specificity of time-domain versus spectral-domain optical coherence tomography in diagnosing early to moderate glaucoma. *Ophthalmology.* 2009;116:2294-9.
- ⁵³ Park SB, Sung KR, Kang SY, et al. Comparison of glaucoma diagnostic capabilities of Cirrus HD and Stratus optical coherence tomography. *Arch Ophthalmol.* 2009;127:1603-9.

-
- ⁵⁴ Moreno-Montañés J, Olmo N, Alvarez A, et al. Cirrus high-definition optical coherence tomography compared with Stratus optical coherence tomography in glaucoma diagnosis. *Invest Ophthalmol Vis Sci*. 2010;51:335-43.
- ⁵⁵ Jeung JW, Park KH. Comparison of Cirrus OCT and Stratus OCT on the ability to detect localized retinal nerve fiber layer defects in preperimetric glaucoma. *Invest Ophthalmol Vis Sci*. 2010;51:938-45.
- ⁵⁶ Huang L, Fan N, Shen X, He J. Comparison of the diagnostic ability of retinal nerve fiber layer thickness measured using Time Domain and Spectral Domain Optical Coherence Tomography in primary open angle glaucoma. *Yan Ke Xue Bao*. 2011; 26(3):132-7.
- ⁵⁷ Ramesh AN, Kambhampati C, Monson JRT, DreW PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004;86:334-8.
- ⁵⁸ Rich E, Knight K. *Artificial Intelligence*. 2nd ed. MacGraw-Hill (USA). 1991
- ⁵⁹ Monard MC, Baranauskas JA. Conceitos sobre aprendizagem de máquina. In: Rezende SO. *Sistemas inteligentes: Fundamentos e aplicações*. São Carlos: Manole; 2003;89-114.
- ⁶⁰ Damasceno, M. Introdução a mineração de dados utilizando o WEKA. In: <http://connepi.ifal.edu.br/ocs/index.php/connepi/CONNepi2010/paper/view/258/207>. Visitado 03/01/13.
- ⁶¹ Witten, I.H. and Frank, E. *Data mining: practical Machine learning Tools and Techniques*. 2 ed. San Francisco: Morgan Kaufmann; 2005.
- ⁶² Gewehr JE, Szugat M, Zimmer R. Bio-Weka extending the Weka framework for bioinformatics. *Bioinformatics*. 2007;(23):651-3.
- ⁶³ Kotsiantis SB. *Supervised Machine Learning: A review of Classification Techniques*. *Informatica*. 2007;31:249-68.
- ⁶⁴ Murthy KS. Automatic construction of decision trees from data: A multi-Disciplinary Survey. *Data Mining and Knowledge Discovery* 2. 1998;345-89.
- ⁶⁵ Pichialini M. *Data Mining na Prática: Árvores de Decisão*. 2006. In: http://imasters.com.br/artigo/5130/sql_server/data_mining_na_pratica_arvores_de_decisao/, visitado em 08/11/12.
- ⁶⁶ GAMA, J. *Árvores de Decisão*. 2002. In :<http://www.liacc.up.pt/~jgama/Bdc/arv.pdf>, visitado em 10/09/2006. GANASCIA, J. G. – *Inteligência Artificial*.
- ⁶⁷ Kovacs ZL. *Redes Neurais Artificiais: Fundamentos e Aplicações*. 2ed. São Paulo: Collegium Cognitio; 1997;174p.il.
- ⁶⁸ Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In : Rumelhart D.E.,McClelland J.L. et al. *Parallel Distributed processing: Explorations in the microstructure of Cognition*. Cambridge, MA: MIT Press; 1986.
- ⁶⁹ Zhang S, Zhang C, Yang Q. Data Preparation for data Mining. *Applied Artificial Intelligence*. 2002;17:375-81.

-
- ⁷⁰ Castro FC, Castro MCF. Redes Neurais Artificiais. [Tese de mestrado]. Porto Alegre (RS): PUCRS-FENG-DEE; 2001.
- ⁷¹ Neocleous C, Schizas C. Artificial Neural Network Learning: A comparative review. LNAI 2308. 2002;300-13
- ⁷² Barreto, GA. Rede de funções de base radial. In: <http://www.deti.ufc.br/~guilherme/TIP705/rbf.pdf>. Visitado em 03/01/2013.
- ⁷³ Coelho LS, Mariani VC. Sistema híbrido neuro-evolutivo aplicado ao controle de um processo multivariável. Sba Controle & Automação. 2006;17 (1):32-48.
- ⁷⁴ Moody JE, Darken C. Fast learning in networks of locally-tuned processing units. Neural Computation. 1989;1(1):281-94.
- ⁷⁵ Oguri P. Aprendizado de Máquina para o Problema de Sentiment Classification. [Dissertação-mestrado]. Rio de Janeiro: PUC-RIO; 2006.
- ⁷⁶ Yang Y, Webb G. On why discretization works for naïve-bayes classifiers, lecture notes in computer Science. 2003;2903:440-52.
- ⁷⁷ Bouckaert R. Naive bayes classifiers that perform well with continuous variables, lecture notes in computer Science, 2004;3339:1089-94.
- ⁷⁸ Pichiliani M. Data Mining na Prática: Classificação Bayesiana. In: masters.com.br/.../data_mining_na_pratica_classificacao_bayesiana/ visitado em 09/11/12.
- ⁷⁹ Vapnik VN. Statistical learning theory. John Wiley and Sons. New York: Spinger-verlag; 1995.
- ⁸⁰ Vapnik VN, Chervonenkis AY. On the uniform convergence of relative frequencies of events to their probabilities. Theory of probability and its Applications. 1971;16(2):283-305.
- ⁸¹ Lorena AC, Carvalho ACPLF. Uma introdução às Support Vector Machines. Revista de informática teórica e aplicada-RITA. 2007;14(2):43-67.
- ⁸² Vapnik VN. An overview of statistical learning theory. IEEE Trans. Neural Netw. 1999;10:988-99.
- ⁸³ Almeida FFM. Relatório técnico: Support vector machine. Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia. 2007.
- ⁸⁴ Smola AJ, Schölkopf. Learning with Kernels. Cambridge, MA: MIT Press; 2002.
- ⁸⁵ Burges CJC. A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining. 1998;2(2):1-43,199
- ⁸⁶ Veropoulos K, Campbell C, Cristianini N. Controlling the sensitivity of support vector machines. In: Proceedings of the international joint conference on artificial intelligence (IJCAI99); 1999.
- ⁸⁷ Haykin, S. Neural networks- A comprehensive Foundation. 2ed. New Jersey: Prentice-Hall; 1999.
- ⁸⁸ Herbrich, R. Learning kernel classifiers: theory and algorithms. MIT Press; 2001.
- ⁸⁹ Genton M. Classes of Kernels for machine learning: A statistics perspective. Journal of machine learning research. 2001;2:299-312.

-
- ⁹⁰ Witten IH, Frank E. Data mining: practical machine learning tools and techniques with Java implementations. San Francisco: Morgan Kaufmann; 2000.
- ⁹¹ Breiman, L. Bagging predictors. *Machine Learning*. Springer. 1996;24(2),123-40.
- ⁹² Lopes L. Aprendizagem de maquina baseada na combinação de classificadores em bases de dados da área de saúde [Dissertação]. Curitiba (PR): PUC-PR; 2007.
- ⁹³ Quinlan, J.R. Bagging, Boosting and C4.5. In: *Proceedings of the Thirteenth National Conference on Artificial Intelligence, AAAI 96*. University of Sydney (Austrália); 1996.
- ⁹⁴ Rezende, S.O. *Sistemas inteligentes, Fundamentos e Aplicações*. São Paulo: Manole; 2003.
- ⁹⁵ Freund Y, Schapire RE. Experiments with a new boosting algorithm. In: *Proceedings of the 13th international conference on machine learning*. San Francisco: Morgan Kaufman; 1996;148-56.
- ⁹⁶ Barlett P, Freund Y, Lee WS, Schapire RE. Boosting the margin: a new explanation for the effectiveness of voting methods. In: *Proceedings of the 14th international conference on machine learning*. San Francisco: Morgan Kaufman; 1997;320-30.
- ⁹⁷ Liaw A, Wiener M. Classification and Regression by Random Forest. *R News*. 2002;2/3:18-22. [online] URL <http://CRAN.R-project.org/doc/Rnews>.
- ⁹⁸ Breiman L. Random forest. *Machine learning*. 2001;45(1):5-32.
- ⁹⁹ Canuto SDC. Um estudo comparativo entre abordagens supervisionadas para a resolução de referencias a autores [Dissertação-Mestrado]. Goiania (GO): Universidade Federal de Goias; 2011.
- ¹⁰⁰ Caruana R, Niculescu-Mizil A, Crew G, Ksikes A. Ensemble selection from libraries of models. In: *Proceedings of the twenty-first international conference on Machine learning (ICML)*; 2004 Jul. 04-08; Banff, Alberta (Canada) p.18;2004.
- ¹⁰¹ Munson A, Cardie C, Caruana R. Optimizing to arbitrary NLP Metrics using Ensemble selection. *Proceedings of Human language technology conference and conference on empirical methods in natural language processing (HLT/EMNLP)*. 2005;530-46.
- ¹⁰² <http://www.cs.cornell.edu/~caruana/caruana.icml04.revised.rev2.ps>
- ¹⁰³ Tsoumakas G, Partalas I, Vlahavas I. A taxonomy and short review of ensemble selection. In *ECAI 08, Workshop on Supervised and Unsupervised Ensemble Methods and Their Applications, SUEMA*; 2008.
- ¹⁰⁴ Greaney MJ, Hoffman DC, Garway-Heath DF, Nakla M, Coleman AL, Caprioli J. Comparison of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. *Invest Ophthalmol Vis Sci*. 2002;43(1):140-5.
- ¹⁰⁵ Bowd C, Chan K, Zangwill LM, Goldbaum MH, Lee TW, Sejnowski TJ, et al. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Invest Ophthalmol Vis Sci*. 2002;43(11):3444-54.
- ¹⁰⁶ Lauande-Pimentel R, Carvalho RA, Oliveira HC, Gonçalves DC, Silva LM, Costa VP. Discrimination between normal and glaucomatous eyes with visual field and scanning laser polarimetry measurements. *Br J Ophthalmol*. 2001;85(5):586-91.

-
- ¹⁰⁷ Chylack LT Jr, Leske MC, McCarthy D, et al. The Lens opacities Classification System III (locs III). *Arch Ophthalmol.* 1993; 111:831-6.
- ¹⁰⁸ Anderson DR, Patella VM. *Automated static perimetry.* St Louis: Mosby Year Book; 1999.
- ¹⁰⁹ Hodapp E, Parrish RK II, Anderson DR. *Clinical decisions in glaucoma.* St. Louis, The C.V, Mosby Co.; 1993;52-61.
- ¹¹⁰ Messias A, Jorge R, Cruz AAV. Tabelas para medir acuidade visual com escala logarítmica: porque usar e como construir. *Arq Bras Oftalmol.* 2010;(1):96-100.
- ¹¹¹ Diamantidis NA, Karlis D, Giankoumakis EA. Unsupervised stratification of cross-validation for accuracy estimation. *Artificial Intelligence.* 2000;116:1-16.
- ¹¹² Tan M, Steinbach V, Kumar AW. Classification: basic concepts, decision tree and model evaluation. In: Tan M, Steinbach V, Kumar AW. *Introduction to data Mining.* Addison- Wesley. 2006;145-206.
- ¹¹³ Witten IH, Frank E. *Data Mining: Practical Machine Learning Tools and Techniques.* 2ed. San Francisco: Morgan Kaufmann; 2005.
- ¹¹⁴ Bazzan AL, Engel PM, Silva SC. Automated annotation of Keywords for Protein related to mycoplasmataceae using machine learning techniques. *Bioinformatics.*2002; 18(suppl 2):35-43.
- ¹¹⁵ Kretschmann E, Fleischmann W, Apweiler R. Automated rule generation for protein annotation with C4.5 data mining algorithm applied on SWISS-PROT. *Bioinformatics.* 2001;17:920-6.
- ¹¹⁶ Li J, Liu H, Ng SK, Wong L. Discovery of significant rules for classifying cancer diagnosis data. *Bioinformatics.* 2003;19(suppl2):93-102.
- ¹¹⁷ Tobler JB, Molla MN, Nuwaysir EF, Shavlik JW. Evaluating machine learning approaches for aiding probe selection for gene-expression arrays. *Bioinformatics.* 2002;18(suppl 1):164S-71S.
- ¹¹⁸ Gewehr JE, Szugat M, Zimmer R. BioWeka-extending the Weka framework for bioinformatics. *Bioinformatics.* 2007;23:651-3.
- ¹¹⁹ Dietterich TG. Machine Learning Research: Four Current Directions. *AI Magazine.* 1997;18(4):97-136.
- ¹²⁰ Frank E, Hall M, Trigg L, Holmes G, Witten IH. Data Mining in Bioinformatics using Weka. *Bioinformatics.* 2004;20:2479-81.
- ¹²¹ Bradley AP. The use of the área under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition.*1997;30(7):1145-59.
- ¹²² Spackman KA. Signal detection theory: Valuable tools for evaluating inductive learning. In *Proceedings of the 6th Int Workshop on Machine Learning (ICML 1989).* San Francisco: Morgan Kaufmann.1989;160-3.
- ¹²³ Flach P, Wu S. Repairing concavities in ROC curves. In *Proceedings of Nineteenth International Joint Conference on Artificial Intelligence(IJCA 2005).* Edinburgh (UK). 2005;702-7.
- ¹²⁴ Prati RC, Batista GEAPA, Monard MC. Curvas ROC para avaliação de classificadores. *IEEE América Latina.* 2008;6(2).
- ¹²⁵ Shah NN, Bowd C, Medeiros FA, et al. Combining structural and functional testing for detection of glaucoma. *Ophthalmology.* 2006;113:1593-602.

-
- ¹²⁶ Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci.* 2005;46:4147-52.
- ¹²⁷ Bizios D, Heijl A, Hougaard JL, et al. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fiber layer thickness parameters measured by Stratus OCT. *Acta Ophthalmol.* 2010;88:44-52.
- ¹²⁸ Bowd C, Hao J, Tavares IM, et al. Bayesian machine learning classifiers for combining structural and functional measurements to classify healthy and glaucomatous eyes. *Invest Ophthalmol Vis Sci.* 2008;49:945-53.
- ¹²⁹ Huang ML, Chen HY. Development and comparison of automated classifiers for glaucoma diagnosis using Stratus optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2005;46:4121-9.
- ¹³⁰ Townsend KA, Wollstein G, Danks D, et al. Heidelberg retina tomograph 3 machine learning classifiers for glaucoma detection. *Br J Ophthalmol.* 2008;92:814-8.
- ¹³¹ Mardin CY, Peters A, Horn F, et al. Improving glaucoma diagnosis by the combination of perimetry and HRT measurements. *J Glaucoma.* 2006;15:299-305.
- ¹³² Zangwill LM, Chan K, Bowd C, et al. Heidelberg retina tomograph measurements of the optic disc and parapapillary retina for detecting glaucoma analyzed by machine learning classifiers. *Invest Ophthalmol Vis Sci.* 2004;45:3144-51.
- ¹³³ Bowd C, Medeiros FA, Zhang Z, et al. Relevance vector machine and support vector machine classifier analysis of scanning laser polarimetry retinal nerve fiber layer measurements. *Invest Ophthalmol Vis Sci.* 2005;46:1322-9.
- ¹³⁴ Adler W, Peters A, Lausen B. Comparison of classifiers applied to confocal scanning laser ophthalmoscopy data. *Methods Inf Med.* 2008;47:38-46.
- ¹³⁵ Knight OJ, Chang RT, Feuer WJ, et al. Comparison of retinal nerve fiber layer measurements using time domain and spectral domain optical coherent tomography. *Ophthalmology.* 2009;116:1271-7.
- ¹³⁶ Townsend KA, Wollstein G, Danks D, et al. Heidelberg retina tomograph 3 machine learning classifiers for glaucoma detection. *Br J Ophthalmol.* 2008;92:814-8.
- ¹³⁷ Boden C, Chan K, Sample PA, et al. Assessing visual field clustering schemes using machine learning classifiers in standard perimetry. *Invest Ophthalmol Vis Sci.* 2007;48:5582-90.
- ¹³⁸ Vidotti VG, Costa VP, Silva FR, Resende GM, Cremasco F, Dias M, Gomi ES. *Eur J Ophthalmol.* 2012 Jun 15:0. doi: 10.5301/ejo.5000183. [Epub ahead of print]
- ¹³⁹ Mardin CY, Peters A, Horn F, Junemann AG, Lausen B. Improving glaucoma diagnosis by the combination of perimetry and HRT measurements. *J Glaucoma.* 2006;15:299-305.
- ¹⁴⁰ Racette L, Chiou CY, et al. Combining functional and structural tests improves the diagnostic accuracy of relevance vector machine classifiers. 2010;19:167-75.

¹⁴¹ Whiting P, Rutjes AW, Reitsma JB, et al. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med.* 2004;140:189-202



9. ANEXOS

ANEXO 1

FACULDADE DE CIÊNCIAS MÉDICAS COMITÊ DE ÉTICA EM PESQUISA

☒ Caixa Postal 6111, 13083-970 Campinas, SP

☎ (0_19) 3788-8936

FAX (0_19) 3788-8925

🌐 www.fcm.unicamp.br/pesquisa/etica/index.html

✉ cep@fcm.unicamp.br

CEP, 23/08/05.
(Grupo III)

PARECER PROJETO: N° 406/2005

CAAE: 0204.0.146.000-05

I-IDENTIFICAÇÃO:

PROJETO: “DIAGNÓSTICO PRECOCE E DE PROGRESSÃO DO GLAUCOMA BASEADOS EM SISTEMA DE APRENDIZAGEM DE CLASSIFICADORES HÍBRIDOS”

PESQUISADOR RESPONSÁVEL: Vital Paulino Costa

INSTITUIÇÃO: Ambulatório de Glaucoma/HC/UNICAMP

APRESENTAÇÃO AO CEP: 01/08/2005

APRESENTAR RELATÓRIO EM: 23/08/06

II - OBJETIVOS

1. Desenvolver sistema de aprendizagem de classificadores híbridos para integrar dados anatômicos, epidemiológicos e funcionais para aumentar a sensibilidade e especificidade do diagnóstico de glaucoma. 2. Desenvolver sistema de aprendizagem de classificadores híbridos que integrará dados anatômicos e funcionais para avaliação longitudinal de pacientes, facilitando o diagnóstico precoce de glaucoma. 3. Desenvolver sistema de aprendizagem de classificadores híbridos que integrará dados anatômicos e funcionais para avaliação longitudinal de pacientes, facilitando o diagnóstico de progressão do glaucoma.

III - SUMÁRIO

Serão recrutados pacientes do ambulatório de oftalmologia da Unicamp com glaucoma, hipertensos oculares e normais para um exame oftalmológico completo. Os resultados dos exames formarão um banco de dados que será utilizado para formação de um sistema computacional híbrido.

IV - COMENTÁRIOS DOS RELATORES

O projeto é relevante já que pretende diagnosticar o glaucoma precocemente, tratá-lo e assim, evitar maiores danos à visão. O Termo de Consentimento Livre e Esclarecido explica completa e pormenorizadamente, a natureza da pesquisa, seus objetivos, métodos, benefícios, potenciais riscos e incômodos eventuais, além de garantir a liberdade do sujeito em participar ou não da pesquisa e assegurar o sigilo dos sujeitos envolvidos.

V - PARECER DO CEP

O Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da UNICAMP, após acatar os pareceres dos membros-relatores previamente designados para o presente caso e atendendo todos os dispositivos das Resoluções 196/96 e complementares, bem como ter aprovado o Termo do Consentimento Livre e Esclarecido, assim como todos os anexos incluídos na Pesquisa, resolve aprovar sem restrições o Protocolo de Pesquisa supracitado.

O conteúdo e as conclusões aqui apresentados são de responsabilidade exclusiva do CEP/FCM/UNICAMP e não representam a opinião da Universidade Estadual de Campinas nem a comprometem.

VI - INFORMAÇÕES COMPLEMENTARES

O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 196/96 – Item IV.1.f) e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (Item IV.2.d).

Pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou (Res. CNS Item III.1.z), exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade do regime oferecido a um dos grupos de pesquisa (Item V.3.).

O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS Item V.4.). É papel do pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.

Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projeto do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial (Res. 251/97, Item III.2.e)

Relatórios parciais e final devem ser apresentados ao CEP, de acordo com os prazos estabelecidos na Resolução CNS-MS 196/96.

VII - DATA DA REUNIÃO

Homologado na VIII Reunião Ordinária do CEP/FCM, em 23 de agosto de 2005.


Prof. Dra. Carmen Sílvia Bertuzzo
PRESIDENTE DO COMITÊ DE ÉTICA EM PESQUISA
FCM / UNICAMP

ANEXO 2

SISTEMA DE CLASSIFICAÇÃO DE OPACIDADE LENTICULAR III

O sistema de classificação de opacidade lenticular II (LOCS II) foi introduzida em 1989¹⁴¹, validado por outros investigadores em 1989 e 1991¹⁴¹¹⁴¹ e foi utilizado em vários estudos epidemiológicos relacionados à história natural da catarata senil. Em 1993¹⁴¹ foi revisado e modificado sendo então renomeado como sistema de classificação de opacidade lenticular III (Locs III) e desde então amplamente utilizado pelo meio acadêmico para graduação da opacidade cristalíniana.

Graduação pelo sistema LOC III

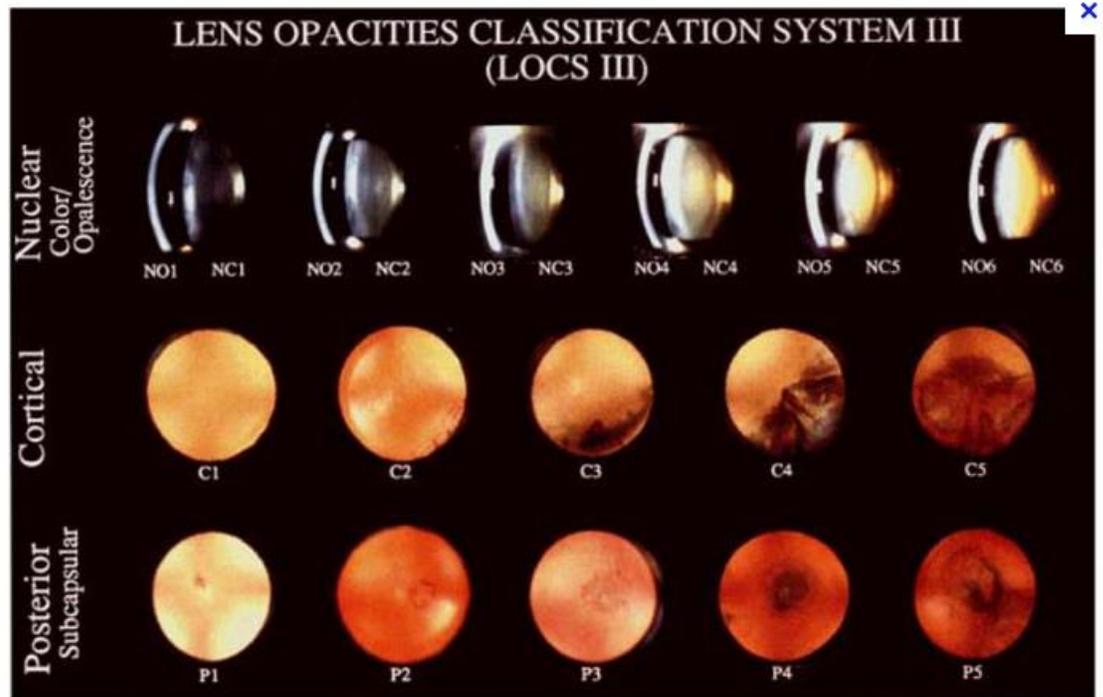
O sistema de classificação de opacidade lenticular é realizado através da comparação da imagem da catarata a ser analisada com imagens padronizadas fornecida pelo LOCS III.(figura). As seguintes observações devem ser consideradas:

Regras gerais

- I- Todas as figuras padrão são limítrofes dos intervalos da escala.
- II- Não existe padrão ou graduação zero no LOCS III.
- III- Fatores a serem avaliados :opalescência nuclear (NO), catarata cortical (C), catarata subcapsular posterior (P) e cor nuclear (NC).
- IV- O graduador decide em qual intervalo a imagem avaliada se encontra; a severidade da opacidade deve ser maior que o padrão inferior e menor ou igual que a próxima escala acima. Em cada intervalo, entre os padrões de referência adjacentes ,é subjetivamente dividido em dez partes iguais, e cada

unidade apresenta intervalo de 0,1. Para cada tipo de catarata ou para opacidade nuclear, maiores valores graduados indicam maior severidade. A escala varia de 0,1 (transparente) até 5,9 (muito opaco, em casos de C e P) ou 6,9 (muito opaco NO e NC).

- V- A graduação decimal usando intervalos de unidade de 0,1 é então utilizado para graduar o grau de opacidade. A escala decimal deve refletir a posição da imagem a ser analisada em relação aos intervalos padrões; por exemplo, uma graduação de 2,5 significa de que a severidade da catarata julgada se encontra na metade entre os padrões 2 e 3. Se a severidade da catarata é igual ao padrão da imagem 3 a graduação é 3,0. Porém, se é menor que a imagem padrão 3, porém maior que a metade entre os intervalos, o valor deve permanecer entre 2,6 e 2,9. Similarmente, o valor de 1,1 a 2,0 é verificado se a catarata é maior que o padrão número 1, porém menor ou igual ao padrão número 2. A mais severa catarata mostrada nos padrões de referência é 5 para C e P e 6 para NC e NO, e os maiores valores para cada graduação deve ser de 5,9 e 6,9 respectivamente. Por isso, os valores variam entre 0,1 e 5,9 ou 6,9.



ANEXO 3:

Tabela de conversão da acuidade visual medida em escala de Snellen para escala LogMar.¹⁴¹

| logMAR | Ângulo (minuto de arco) | Decimal | Imperial | Métrico | Frequência espacial (C/°) |
|---------------|--|----------------|-----------------|----------------|--|
| 1,3 | 20,0 | 0,05 | 20/400 | 6/120 | 600 |
| 1,2 | 15,8 | 0,06 | 20/317 | 6/95 | 475 |
| 1,1 | 12,6 | 0,08 | 20/252 | 6/76 | 378 |
| 1,0 | 10,0 | 0,10 | 20/200 | 6/60 | 300 |
| 0,9 | 7,9 | 0,13 | 20/159 | 6/48 | 238 |
| 0,8 | 6,3 | 0,16 | 20/126 | 6/38 | 189 |
| 0,7 | 5,0 | 0,20 | 20/100 | 6/30 | 150 |
| 0,6 | 4,0 | 0,25 | 20/80 | 6/24 | 119 |
| 0,5 | 3,2 | 0,32 | 20/63 | 6/19 | 95 |
| 0,4 | 2,5 | 0,40 | 20/50 | 6/15 | 75 |
| 0,3 | 2,0 | 0,50 | 20/40 | 6/12 | 60 |
| 0,2 | 1,6 | 0,63 | 20/32 | 6/10 | 48 |
| 0,1 | 1,3 | 0,79 | 20/25 | 6/8 | 38 |
| 0 | 1,0 | 1,00 | 20/20 | 6/6 | 30 |
| -0,1 | 0,8 | 1,26 | 20/16 | 6/5 | 24 |
| -0,2 | 0,6 | 1,58 | 20/13 | 6/4 | 19 |
| -0,3 | 0,5 | 2,00 | 20/10 | 6/3 | 15 |

ANEXO 4

dos pacientes normais incluídos no estudo:

| | Idade | sexo | raça | AV | RX | PIO | n med | MD | PSD | GHT | AT | T | S | N | I | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|-------|------|------|----|-------|-----|-------|-------|------|-----|-----|----|-----|-----|-----|-----|-----|----|----|-----|-----|-----|----|-----|-----|-----|-----|
| 1 | 55 | 2 | 3 | 0 | 1,5 | 21 | 0 | -5,91 | 5,22 | 3 | 110 | 71 | 142 | 83 | 144 | 129 | 117 | 67 | 65 | 117 | 177 | 140 | 68 | 56 | 88 | 150 | 146 |
| 2 | 51 | 1 | 1 | 0 | 3,25 | 12 | 0 | -1,38 | 2,02 | 1 | 89 | 61 | 105 | 80 | 110 | 118 | 89 | 68 | 82 | 96 | 94 | 139 | 64 | 48 | 71 | 105 | 92 |
| 1 | 78 | 2 | 1 | 0 | 3,75 | 17 | 0 | -2,38 | 1,42 | 1 | 84 | 48 | 110 | 76 | 100 | 100 | 93 | 53 | 83 | 106 | 126 | 69 | 43 | 42 | 59 | 95 | 134 |
| 1 | 47 | 1 | 1 | 0 | -0,5 | 16 | 0 | -2,95 | 2,84 | 2 | 77 | 48 | 109 | 66 | 86 | 105 | 83 | 61 | 55 | 60 | 106 | 91 | 50 | 33 | 61 | 115 | 107 |
| 1 | 45 | 2 | 2 | 0 | -0,75 | 14 | 0 | -0,56 | 1,66 | 3 | 112 | 71 | 133 | 78 | 165 | 153 | 106 | 63 | 64 | 140 | 203 | 153 | 71 | 59 | 84 | 111 | 153 |
| 1 | 76 | 1 | 1 | 0 | 2,25 | 16 | 0 | 0,22 | 1,5 | 1 | 86 | 64 | 103 | 81 | 95 | 73 | 93 | 82 | 68 | 66 | 83 | 137 | 74 | 47 | 71 | 128 | 109 |
| 1 | 73 | 1 | 1 | 0 | plano | 13 | 0 | -2,88 | 1,73 | 1 | 84 | 54 | 106 | 66 | 110 | 99 | 81 | 55 | 64 | 99 | 138 | 92 | 54 | 46 | 61 | 91 | 128 |
| 1 | | 2 | 2 | 0 | plano | 12 | 0 | -0,79 | 1,38 | 1 | 77 | 56 | 87 | 62 | 104 | 71 | 85 | 54 | 46 | 84 | 111 | 117 | 59 | 47 | 61 | 93 | 98 |
| 1 | 48 | 1 | 1 | 0 | plano | 16 | 0 | -3,26 | 3,88 | 1 | 93 | 59 | 118 | 76 | 120 | 100 | 83 | 71 | 75 | 96 | 120 | 144 | 62 | 47 | 69 | 141 | 113 |
| 1 | 54 | 2 | 1 | 0 | plano | 10 | 0 | -2,97 | 2,58 | 2 | 80 | 53 | 82 | 74 | 111 | 68 | 68 | 74 | 81 | 102 | 128 | 103 | 46 | 45 | 68 | 117 | 61 |
| 1 | 52 | 2 | 1 | 0 | 2,75 | 16 | 0 | -0,6 | 1,7 | 1 | 88 | 64 | 107 | 69 | 112 | 98 | 84 | 57 | 66 | 79 | 114 | 142 | 66 | 49 | 78 | 129 | 93 |
| 2 | 55 | 1 | 1 | 0 | plano | 16 | 0 | -2,5 | 2,1 | 2 | 91 | 54 | 120 | 70 | 120 | 125 | 105 | 51 | 55 | 91 | 120 | 148 | 65 | 38 | 58 | 117 | 117 |
| 1 | 57 | 2 | 1 | 0 | 1,25 | 13 | 0 | -0,95 | 2 | 3 | 110 | 55 | 139 | 101 | 143 | 158 | 128 | 88 | 88 | 122 | 154 | 152 | 58 | 39 | 69 | 112 | 147 |
| 1 | 52 | 2 | 1 | 0 | -0,75 | 12 | 0 | -2,04 | 2,21 | 1 | 94 | 67 | 11 | 82 | 115 | 95 | 88 | 72 | 86 | 94 | 133 | 119 | 57 | 54 | 89 | 140 | 98 |
| 1 | 49 | 1 | 1 | 0 | plano | 12 | 0 | -1,08 | 1,19 | 1 | 95 | 81 | 110 | 70 | 119 | 101 | 84 | 58 | 68 | 109 | 127 | 121 | 81 | 100 | 62 | 88 | 141 |
| 2 | 64 | 2 | 1 | 0 | 4,25 | 11 | 0 | -4,12 | 1,37 | 4 | 94 | 60 | 127 | 66 | 121 | 115 | 76 | 61 | 62 | 102 | 161 | 101 | 56 | 51 | 74 | 123 | 142 |
| 1 | 59 | 2 | 2 | 0 | plano | 10 | 0 | -3,03 | 2,05 | 2 | 98 | 77 | 132 | 59 | 126 | 124 | 73 | 44 | 58 | 87 | 126 | 164 | 87 | 59 | 86 | 137 | 133 |
| 1 | 49 | 2 | 2 | 0 | plano | 20 | 0 | -3,83 | 1,84 | 1 | 89 | 55 | 116 | 71 | 113 | 86 | 102 | 58 | 54 | 77 | 136 | 127 | 46 | 45 | 74 | 154 | 109 |
| 1 | 48 | 2 | 1 | 0 | 1,5 | 14 | 0 | -1,95 | 1,98 | 1 | 102 | 66 | 134 | 69 | 138 | 108 | 78 | 73 | 55 | 120 | 141 | 153 | 74 | 50 | 73 | 131 | 164 |
| 1 | 54 | 1 | 1 | 0 | plano | 18 | 0 | -2,11 | 1,55 | 1 | 102 | 69 | 125 | 73 | 141 | 100 | 76 | 61 | 83 | 131 | 172 | 122 | 49 | 62 | 95 | 154 | 121 |
| 1 | 45 | 1 | 2 | 0 | 0,75 | 19 | 0 | -1,84 | 1,67 | 1 | 70 | 39 | 93 | 49 | 100 | 75 | 58 | 41 | 48 | 87 | 122 | 91 | 38 | 48 | 32 | 88 | 116 |
| 1 | 52 | 1 | 2 | 0 | -2,75 | 18 | 0 | -2,93 | 1,29 | 1 | 93 | 65 | 126 | 82 | 100 | 110 | 102 | 62 | 81 | 80 | 102 | 119 | 63 | 55 | 79 | 125 | 144 |
| 1 | 54 | 2 | 1 | 0 | 4,25 | 16 | 0 | -0,07 | 1,46 | 1 | 98 | 61 | 131 | 79 | 123 | 117 | 98 | 66 | 74 | 100 | 131 | 138 | 60 | 46 | 77 | 139 | 136 |
| 2 | 51 | 2 | 1 | 0 | 1 | 12 | 0 | 0,02 | 1,05 | 1 | 91 | 67 | 110 | 60 | 126 | 102 | 86 | 52 | 41 | 92 | 130 | 155 | 69 | 58 | 74 | 134 | 95 |
| 2 | 64 | 2 | 1 | 0 | 5 | 16 | 0 | -3,5 | 1,47 | 1 | 93 | 59 | 121 | 75 | 116 | 118 | 71 | 50 | 56 | 104 | 134 | 109 | 64 | 56 | 106 | 113 | 131 |
| 1 | 59 | 1 | 1 | 0 | 3,25 | 16 | 0 | 0,45 | 1,31 | 1 | 95 | 60 | 112 | 65 | 142 | 87 | 90 | 54 | 52 | 114 | 163 | 149 | 60 | 48 | 72 | 130 | 121 |
| 1 | 69 | 2 | 1 | 0 | 5,5 | 11 | 0 | -2,91 | 2,09 | 1 | 80 | 62 | 101 | 56 | 103 | 81 | 61 | 47 | 60 | 81 | 99 | 127 | 62 | 50 | 74 | 127 | 94 |
| 2 | 63 | 1 | 1 | 0 | plano | 15 | 0 | -1,83 | 1,68 | 3 | 82 | 55 | 102 | 60 | 112 | 90 | 85 | 51 | 45 | 93 | 130 | 113 | 57 | 44 | 65 | 93 | 124 |
| 1 | 65 | 2 | 1 | 0 | 3 | 20 | 0 | -2,6 | 2,08 | 1 | 90 | 53 | 117 | 72 | 117 | 102 | 109 | 54 | 53 | 86 | 132 | 132 | 51 | 43 | 64 | 104 | 145 |
| 1 | 54 | 2 | 1 | 0 | 1,5 | 18 | 0 | -0,31 | 1,37 | 1 | 102 | 86 | 120 | 68 | 132 | 82 | 81 | 62 | 61 | 108 | 158 | 129 | 79 | 79 | 101 | 152 | 127 |
| 2 | 65 | 1 | 1 | 0 | plano | 16 | 0 | 1,89 | 1,12 | 1 | 92 | 50 | 123 | 74 | 119 | 112 | 84 | 62 | 71 | 89 | 152 | 115 | 45 | 41 | 66 | 119 | 146 |
| 1 | 82 | 1 | 1 | 1 | plano | 14 | 0 | -2,73 | 1,88 | 1 | 78 | 50 | 105 | 68 | 89 | 100 | 92 | 52 | 60 | 78 | 95 | 95 | 51 | 43 | 58 | 80 | 135 |
| 1 | 65 | 1 | 2 | 0 | plano | 14 | 0 | -2,43 | 2,63 | 1 | 99 | 60 | 118 | 76 | 143 | 106 | 97 | 64 | 66 | 117 | 171 | 141 | 54 | 44 | 82 | 142 | 107 |
| 2 | 46 | 1 | 1 | 0 | -1,75 | 14 | 0 | -0,66 | 1,61 | 1 | 101 | 71 | 141 | 69 | 125 | 121 | 95 | 53 | 58 | 96 | 115 | 165 | 73 | 56 | 84 | 150 | 153 |
| 1 | 57 | 1 | 3 | 0 | 1,25 | 12 | 0 | -1,19 | 1,82 | 1 | 89 | 73 | 104 | 53 | 128 | 98 | 64 | 46 | 48 | 105 | 133 | 147 | 73 | 62 | 83 | 137 | 76 |
| 1 | 62 | 1 | 1 | 0 | plano | 18 | 0 | 1,49 | 1,35 | 1 | 110 | 69 | 142 | 103 | 126 | 101 | 139 | 78 | 92 | 101 | 121 | 154 | 88 | 50 | 68 | 153 | 172 |
| 2 | 61 | 1 | 1 | 0 | 0,25 | 16 | 0 | -2 | 1,45 | 1 | 90 | 77 | 98 | 62 | 124 | 121 | 84 | 66 | 80 | 162 | 132 | 78 | 51 | 58 | 77 | 93 | 80 |

| Nome | Olho | idade | sexo | rgsg | AV | RX | PIO | n med | PSD | GHT | AT | T | S | N | I | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | |
|------|------|-------|------|------|----|-------|-----|-------|-------|------|----|-----|----|-----|----|-----|-----|-----|----|----|-----|-----|-----|----|----|-----|-----|-----|
| JFC | 1 | 48 | 2 | 1 | 0 | pleno | 15 | 0 | -0,46 | 1,13 | 1 | 85 | 71 | 102 | 66 | 99 | 86 | 96 | 95 | 75 | 86 | 137 | 74 | 47 | 93 | 133 | 86 | |
| MS | 1 | 49 | 2 | 1 | 0 | 0,5 | 15 | 0 | 0,86 | 1,06 | 1 | 103 | 73 | 125 | 86 | 127 | 108 | 110 | 74 | 73 | 99 | 138 | 144 | 77 | 56 | 85 | 137 | 130 |
| MAF | 2 | 56 | 2 | 2 | 0 | 1 | 14 | 0 | -0,22 | 1,68 | 1 | 98 | 94 | 122 | 84 | 131 | 130 | 105 | 71 | 75 | 101 | 173 | 130 | 58 | 42 | 62 | 98 | 136 |
| MISA | 1 | 47 | 2 | 1 | 0 | pleno | 12 | 0 | -1,98 | 1,79 | 1 | 98 | 68 | 123 | 67 | 133 | 114 | 82 | 93 | 63 | 108 | 134 | 157 | 83 | 47 | 75 | 112 | 144 |
| MIMP | 1 | 66 | 2 | 1 | 0 | pleno | 13 | 0 | 0,05 | 1,54 | 1 | 89 | 67 | 106 | 74 | 111 | 77 | 80 | 98 | 83 | 85 | 120 | 127 | 72 | 48 | 80 | 132 | 109 |
| MMIG | 2 | 66 | 2 | 3 | 0 | 1,5 | 16 | 0 | -0,09 | 1,63 | 1 | 89 | 84 | 104 | 68 | 99 | 86 | 76 | 64 | 64 | 66 | 109 | 121 | 78 | 59 | 114 | 150 | 76 |
| OCS | 1 | 50 | 1 | 1 | 0 | 3,75 | 10 | 0 | 0,65 | 1,2 | 1 | 109 | 64 | 138 | 80 | 154 | 117 | 99 | 73 | 70 | 139 | 168 | 154 | 71 | 51 | 70 | 137 | 158 |
| RS | 1 | 47 | 1 | 1 | 0 | -2,75 | 13 | 0 | 0,37 | 1,2 | 1 | 102 | 66 | 141 | 77 | 125 | 117 | 104 | 58 | 68 | 83 | 136 | 155 | 73 | 44 | 81 | 165 | 140 |
| SS | 1 | 67 | 1 | 1 | 0 | pleno | 16 | 0 | -0,63 | 1,89 | 1 | 107 | 69 | 143 | 79 | 137 | 113 | 93 | 64 | 79 | 111 | 143 | 156 | 77 | 46 | 83 | 135 | 162 |
| TFM | 1 | 49 | 2 | 1 | 0 | 0,75 | 12 | 0 | -3,29 | 2,49 | 1 | 107 | 61 | 129 | 92 | 145 | 89 | 128 | 78 | 68 | 119 | 153 | 164 | 60 | 46 | 76 | 143 | 154 |
| WA | 2 | 55 | 1 | 1 | 0 | pleno | 19 | 0 | -1,19 | 3,67 | 3 | 95 | 63 | 111 | 76 | 131 | 123 | 101 | 58 | 69 | 126 | 148 | 130 | 60 | 62 | 67 | 97 | 114 |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

ANEXO 5

Dados dos pacientes glaucomatosos incluídos no estudo

| Nome | olho | idade | sexo | raça | AV | rK (EE) | PIO | med | MD | PSD | GHT | AT | T | S | N | I | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------|------|-------|------|------|------|---------|-----|-----|------|------|-----|-----|----|-----|----|-----|-----|-----|----|----|-----|-----|-----|-----|----|----|-----|-----|
| AFBS | 1 | 64 | 2 | 1 | 0,2 | 3,25 | 20 | 2 | -1,7 | 2,58 | 1 | 94 | 67 | 112 | 80 | 118 | 75 | 88 | 68 | 85 | 86 | 125 | 143 | 63 | 51 | 87 | 157 | 102 |
| ASS | 1 | 48 | 2 | 1 | 0 | pleno | 12 | 4 | -1,5 | 2,28 | 1 | 83 | 58 | 102 | 84 | 87 | 66 | 84 | 90 | 79 | 101 | 96 | 65 | 38 | 42 | 94 | 108 | 133 |
| MAS | 2 | 68 | 2 | 2 | 0,3 | 1,75 | 13 | 1 | -3,4 | 6,83 | 3 | 86 | 63 | 110 | 73 | 98 | 108 | 91 | 66 | 63 | 73 | 94 | 139 | 67 | 50 | 73 | 107 | 115 |
| ATOS | 1 | 69 | 2 | 1 | 0,3 | 4 | 18 | 1 | -2,8 | 2,81 | 2 | 74 | 55 | 84 | 51 | 105 | 73 | 63 | 47 | 44 | 79 | 134 | 103 | 50 | 49 | 66 | 91 | 88 |
| ALSS | 1 | 59 | 2 | 1 | 0,15 | pleno | 14 | 1 | -4,8 | 4,47 | 3 | 96 | 83 | 108 | 77 | 116 | 100 | 92 | 70 | 70 | 92 | 119 | 137 | 93 | 70 | 85 | 97 | 128 |
| AAN | 2 | 48 | 2 | 2 | 0,05 | 1,25 | 13 | 3 | -4,3 | 1,81 | 4 | 63 | 42 | 77 | 53 | 82 | 76 | 62 | 50 | 48 | 72 | 89 | 85 | 46 | 37 | 42 | 76 | 77 |
| ACC | 1 | 60 | 1 | 1 | 0 | 0,25 | 14 | 2 | -3,2 | 2,97 | 2 | 68 | 53 | 99 | 51 | 70 | 87 | 57 | 45 | 51 | 67 | 75 | 68 | 51 | 48 | 59 | 96 | 114 |
| AV | 2 | 57 | 1 | 1 | 0,1 | pleno | 13 | 2 | -12 | 11 | 3 | 75 | 48 | 116 | 59 | 75 | 135 | 75 | 53 | 49 | 108 | 69 | 48 | 46 | 41 | 57 | 72 | 139 |
| ASF | 1 | 62 | 2 | 1 | 0,3 | -3,3 | 12 | 4 | -3,8 | 2,27 | 2 | 78 | 51 | 95 | 73 | 95 | 75 | 83 | 65 | 70 | 79 | 116 | 90 | 49 | 47 | 58 | 85 | 125 |
| AG | 2 | 78 | 1 | 1 | 0,3 | pleno | 11 | 4 | -6,8 | 5,81 | 3 | 76 | 74 | 59 | 75 | 94 | 73 | 79 | 80 | 65 | 71 | 110 | 102 | 69 | 65 | 90 | 60 | 45 |
| AC | 2 | 73 | 1 | 1 | 0,15 | 3 | 12 | 4 | -12 | 12,6 | 3 | 72 | 50 | 95 | 70 | 72 | 96 | 85 | 72 | 54 | 74 | 63 | 80 | 44 | 44 | 64 | 99 | 89 |
| BIBR | 1 | 71 | 2 | 1 | 0,3 | 1,25 | 16 | 2 | -1,9 | 3,5 | 3 | 80 | 67 | 78 | 78 | 97 | 78 | 66 | 74 | 93 | 88 | 104 | 98 | 53 | 62 | 84 | 106 | 51 |
| ESS | 1 | 69 | 2 | 1 | 0 | 1 | 11 | 2 | 0,15 | 4,41 | 1 | 89 | 63 | 110 | 68 | 117 | 93 | 93 | 53 | 58 | 101 | 125 | 124 | 78 | 47 | 64 | 103 | 134 |
| ECS | 1 | 60 | 1 | 2 | 0 | pleno | 14 | 1 | -0,7 | 2,67 | 2 | 71 | 40 | 82 | 70 | 91 | 75 | 79 | 66 | 64 | 117 | 89 | 67 | 35 | 39 | 47 | 99 | 72 |
| ES | 1 | 60 | 1 | 2 | 0,2 | -1,5 | 14 | 2 | -3,3 | 2,27 | 3 | 71 | 37 | 88 | 85 | 96 | 74 | 67 | 55 | 72 | 86 | 120 | 81 | 35 | 28 | 49 | 92 | 96 |
| EM | 1 | 48 | 1 | 1 | 0 | pleno | 21 | 2 | -5,7 | 7,81 | 3 | 86 | 71 | 111 | 62 | 102 | 110 | 77 | 61 | 48 | 74 | 93 | 138 | 80 | 61 | 72 | 92 | 129 |
| ESF | 1 | 67 | 2 | 1 | 0,3 | -0,5 | 12 | 3 | -1,5 | 3,09 | 2 | 71 | 47 | 83 | 59 | 94 | 80 | 62 | 63 | 52 | 81 | 114 | 86 | 51 | 38 | 53 | 87 | 83 |
| FCS | 1 | 60 | 1 | 3 | 0,2 | pleno | 12 | 4 | -6,2 | 6,37 | 3 | 81 | 55 | 91 | 69 | 107 | 74 | 73 | 61 | 75 | 97 | 120 | 103 | 54 | 47 | 64 | 106 | 93 |
| GB | 2 | 52 | 1 | 1 | 0 | 0,5 | 16 | 0 | -2,1 | 2,87 | 3 | 104 | 85 | 138 | 60 | 135 | 153 | 67 | 55 | 58 | 87 | 144 | 174 | 107 | 64 | 83 | 112 | 148 |
| GCS | 2 | 70 | 1 | 2 | 0,15 | 1 | 15 | 3 | -3,2 | 3,63 | 3 | 64 | 48 | 91 | 33 | 65 | 79 | 71 | 39 | 50 | 79 | 76 | 41 | 43 | 42 | 58 | 100 | 95 |
| GM | 1 | 69 | 1 | 1 | 0,2 | 1 | 14 | 2 | -5,9 | 3,99 | 2 | 92 | 65 | 111 | 72 | 122 | 101 | 75 | 64 | 76 | 114 | 148 | 103 | 61 | 54 | 79 | 109 | 124 |
| ISA | 2 | 70 | 1 | 2 | 0,05 | 1,5 | 10 | 2 | -1,9 | 3,15 | 3 | 72 | 42 | 83 | 69 | 96 | 63 | 103 | 50 | 53 | 82 | 102 | 104 | 39 | 37 | 49 | 76 | 110 |
| JMT | 1 | 64 | 2 | 1 | 0,2 | pleno | 14 | 1 | -9,3 | 8,47 | 3 | 100 | 82 | 130 | 70 | 118 | 108 | 83 | 54 | 73 | 99 | 97 | 157 | 94 | 58 | 95 | 154 | 129 |
| IBL | 1 | 57 | 1 | 2 | 0 | pleno | 13 | 3 | -2,6 | 3,8 | 3 | 74 | 40 | 85 | 71 | 101 | 97 | 97 | 58 | 57 | 82 | 131 | 90 | 45 | 34 | 42 | 60 | 96 |
| JCP5 | 2 | 58 | 1 | 2 | 0 | 0,75 | 12 | 2 | -1,3 | 1,9 | 3 | 76 | 38 | 107 | 59 | 101 | 100 | 68 | 51 | 57 | 98 | 122 | 82 | 43 | 31 | 40 | 81 | 140 |
| JRDF | 2 | 66 | 1 | 3 | 0,05 | pleno | 14 | 1 | -5,8 | 3,89 | 3 | 87 | 53 | 121 | 66 | 108 | 112 | 86 | 55 | 58 | 102 | 120 | 101 | 62 | 45 | 51 | 144 | 107 |
| JTO | 1 | 53 | 1 | 1 | 0 | 1 | 13 | 4 | -4,4 | 4,44 | 3 | 71 | 54 | 84 | 57 | 88 | 71 | 68 | 49 | 53 | 71 | 100 | 95 | 47 | 50 | 67 | 105 | 78 |
| JASF | 2 | 63 | 1 | 1 | 0,1 | 0,75 | 14 | 3 | -2,7 | 3,49 | 3 | 82 | 52 | 94 | 65 | 117 | 93 | 93 | 50 | 52 | 105 | 130 | 115 | 50 | 49 | 58 | 113 | 76 |

| Continuação | | olho | idade | sexo | raça | AV | RK (EE) | PIO | med | MD | PSD | GHT | AT | T | S | N | I | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-------------|---|------|-------|------|------|-------|------------|-----|------|------|-----|-----|----|-----|----|-----|-----|-----|----|----|-----|-----|-----|----|-----|----|-----|-----|----|
| JAR | 2 | 51 | 1 | 3 | 0,15 | -4,8 | 10 | 3 | -3,4 | 2,76 | 2 | 78 | 46 | 103 | 69 | 95 | 91 | 78 | 67 | 61 | 67 | 128 | 90 | 37 | 40 | 62 | 126 | 91 | |
| JSN | 1 | 67 | 1 | 1 | 0 | 2,25 | 13 | 4 | -2,6 | 2,43 | 3 | 67 | 43 | 77 | 56 | 90 | 74 | 69 | 43 | 55 | 82 | 114 | 74 | 49 | 44 | 35 | 77 | 81 | |
| JJA | 2 | 67 | 1 | 1 | 0 | 1,5 | 12 | 4 | -6,4 | 4,93 | 3 | 74 | 49 | 104 | 64 | 79 | 96 | 59 | 64 | 69 | 68 | 54 | 113 | 68 | 30 | 48 | 90 | 126 | |
| JRF | 1 | 57 | 1 | 1 | 0,3 | -2,5 | 13 | 3 | -2,8 | 1,78 | 3 | 69 | 48 | 89 | 63 | 77 | 82 | 66 | 49 | 73 | 100 | 89 | 41 | 45 | 44 | 54 | 90 | 94 | |
| JGS | 1 | 56 | 2 | 1 | 0,2 | 0,75 | 16 | 2 | -1,9 | 4,73 | 3 | 73 | 36 | 103 | 69 | 84 | 84 | 83 | 53 | 68 | 109 | 125 | 18 | 45 | 2 | 62 | 113 | 113 | |
| JDS | 1 | 53 | 2 | 1 | 0 | 1,5 | 16 | 1 | -3,7 | 6,16 | 3 | 95 | 62 | 128 | 76 | 115 | 118 | 92 | 63 | 73 | 107 | 109 | 130 | 59 | 50 | 77 | 117 | 148 | |
| LP | 2 | 54 | 2 | 1 | 0,2 | plano | 14 | 1 | -3,6 | 3,89 | 2 | 92 | 63 | 112 | 81 | 113 | 93 | 99 | 64 | 79 | 93 | 117 | 131 | 64 | 49 | 75 | 123 | 118 | |
| LES | 1 | 45 | 2 | 1 | 0 | -1,3 | 18 | 2 | -3,8 | 3,1 | 3 | 83 | 89 | 96 | 66 | 82 | 94 | 87 | 54 | 57 | 71 | 78 | 97 | 91 | 112 | 65 | 94 | 100 | |
| ML | 1 | 64 | 1 | 1 | 0,15 | 2 | 14 | 3 | -6,3 | 5,66 | 3 | 73 | 38 | 90 | 66 | 98 | 67 | 66 | 66 | 68 | 70 | 128 | 97 | 33 | 30 | 50 | 108 | 95 | |
| MSC | 1 | 78 | 2 | 1 | 0,3 | 4,25 | 14 | 3 | -7,1 | 5,79 | 3 | 77 | 46 | 108 | 81 | 75 | 100 | 103 | 65 | 76 | 96 | 80 | 48 | 45 | 38 | 54 | 100 | 123 | |
| MINOS | 1 | 52 | 2 | 1 | 0,05 | 2,75 | 18 | 1 | -2,8 | 2,36 | 2 | 79 | 50 | 92 | 65 | 108 | 81 | 86 | 53 | 55 | 92 | 141 | 92 | 49 | 35 | 65 | 123 | 72 | |
| MLSC | 1 | 46 | 2 | 3 | 0 | plano | 13 | 4 | -3,6 | 1,97 | 3 | 63 | 54 | 69 | 48 | 80 | 68 | 52 | 48 | 45 | 75 | 72 | 92 | 61 | 42 | 60 | 65 | 73 | |
| MIOR | 1 | 67 | 2 | 1 | 0 | 4 | 12 | 1 | -2,7 | 2,33 | 2 | 78 | 69 | 94 | 65 | 84 | 70 | 76 | 58 | 60 | 81 | 71 | 101 | 79 | 58 | 70 | 106 | 105 | |
| MIART | 2 | 59 | 2 | 1 | 0,05 | 1,25 | 10 | 1 | -4,4 | 4,39 | 3 | 85 | 52 | 122 | 62 | 104 | 105 | 72 | 53 | 61 | 97 | 122 | 92 | 53 | 38 | 65 | 151 | 108 | |
| MNSR | 2 | 52 | 2 | 1 | 0 | 2,25 | 11 | 3 | -2,2 | 2,77 | 1 | 75 | 41 | 115 | 53 | 92 | 106 | 67 | 48 | 43 | 80 | 114 | 83 | 34 | 38 | 52 | 104 | 134 | |
| MF | 2 | 43 | 2 | 1 | 0,3 | plano | 12 | 1 | -4,6 | 4,13 | 3 | 90 | 62 | 114 | 63 | 123 | 106 | 82 | 50 | 58 | 104 | 126 | 138 | 79 | 44 | 62 | 100 | 137 | |
| MPS | 1 | 50 | 2 | 1 | 0 | 2 | 16 | 2 | -4,5 | 7,27 | 3 | 75 | 62 | 85 | 71 | 82 | 75 | 84 | 69 | 60 | 61 | 77 | 107 | 58 | 51 | 76 | 112 | 67 | |
| NLM | 1 | 78 | 2 | 1 | 0,15 | plano | 12 | 2 | -5,9 | 5,46 | 3 | 86 | 75 | 105 | 58 | 104 | 77 | 73 | 56 | 46 | 84 | 120 | 108 | 77 | 63 | 84 | 122 | 117 | |
| NCM | 1 | 50 | 2 | 1 | 0 | 0,75 | 18 | 3 | -2,7 | 1,86 | 3 | 86 | 56 | 104 | 69 | 114 | 91 | 89 | 62 | 56 | 91 | 128 | 123 | 66 | 44 | 58 | 87 | 133 | |
| NSS | 2 | 58 | 2 | 1 | 0,05 | plano | 14 | 4 | -6,1 | 7,39 | 3 | 86 | 56 | 100 | 62 | 127 | 78 | 68 | 59 | 57 | 104 | 140 | 138 | 58 | 47 | 63 | 95 | 127 | |
| NTSS | 2 | 50 | 2 | 1 | 0,05 | 1,25 | 12 | 1 | -1,2 | 1,47 | 1 | 89 | 68 | 121 | 52 | 114 | 116 | 56 | 59 | 40 | 78 | 121 | 143 | 78 | 51 | 75 | 122 | 125 | |
| ON | 1 | 73 | 1 | 1 | 0 | 3,5 | 18 | 1 | -7,4 | 11,8 | 3 | 71 | 50 | 87 | 55 | 93 | 91 | 68 | 40 | 56 | 69 | 81 | 130 | 51 | 33 | 66 | 76 | 95 | |
| OB | 1 | 55 | 1 | 1 | 0,05 | plano | 16 | 1 | -4,1 | 3,47 | 2 | 86 | 53 | 116 | 70 | 104 | 122 | 88 | 48 | 76 | 116 | 127 | 68 | 49 | 46 | 65 | 108 | 118 | |
| PF | 2 | 71 | 1 | 1 | 0,1 | 3,5 | 12 | 0 | -3,2 | 3,97 | 3 | 74 | 51 | 90 | 59 | 96 | 85 | 72 | 45 | 59 | 70 | 98 | 121 | 59 | 37 | 58 | 99 | 84 | |
| PDL | 1 | 53 | 1 | 3 | 0 | plano | 16 | 3 | -2,2 | 3,13 | 3 | 53 | 43 | 74 | 34 | 60 | 85 | 46 | 45 | 11 | 35 | 74 | 71 | 43 | 35 | 51 | 54 | 84 | |
| RFG | 1 | 46 | 2 | 1 | 0 | -0,3 | 16 | 2 | -4,4 | 3,46 | 3 | 79 | 63 | 86 | 72 | 94 | 69 | 79 | 65 | 73 | 101 | 110 | 73 | 62 | 54 | 74 | 83 | 105 | |
| RAF | 1 | 58 | 1 | 1 | 0,05 | 1,5 | 13 | 3 | -3,4 | 3,73 | 3 | 75 | 58 | 87 | 63 | 92 | 75 | 69 | 57 | 62 | 85 | 91 | 100 | 70 | 48 | 56 | 92 | 95 | |
| SCL | 1 | 61 | 1 | 1 | 0 | 4 | 13 | 1 | -2,4 | 2,18 | 3 | 86 | 53 | 107 | 71 | 114 | 90 | 85 | 65 | 62 | 100 | 128 | 114 | 61 | 42 | 56 | 106 | 126 | |
| SO | 1 | 75 | 1 | 4 | 0,1 | 0,25 | 10 | 2 | -2,7 | 2,07 | 2 | 105 | 76 | 128 | 68 | 149 | 96 | 78 | 61 | 66 | 126 | 180 | 143 | 77 | 68 | 83 | 144 | 143 | |
| SAS | 1 | 54 | 1 | 2 | 0,05 | 0,5 | 12 | 1 | -5,5 | 5,96 | 3 | 99 | 79 | 128 | 71 | 118 | 94 | 67 | 68 | 79 | 84 | 112 | 160 | 81 | 61 | 95 | 186 | 104 | |
| SABT | 1 | 50 | 2 | 1 | 0,2 | 2 | 19 | 0 | -4,1 | 6,83 | 3 | 97 | 71 | 139 | 72 | 104 | 116 | 93 | 63 | 60 | 59 | 115 | 138 | 63 | 53 | 96 | 143 | 139 | |
| TMO | 2 | 59 | 2 | 3 | 0,1 | 2,75 | 10 | 3 | -4,8 | 7,07 | 3 | 95 | 43 | 126 | 94 | 115 | 144 | 111 | 90 | 82 | 113 | 139 | 92 | 35 | 38 | 56 | 104 | 130 | |

ANEXO 6

Tabela das aROCs dos MLCs com todos os parâmetros e com a redução de dimensionalidade (exclusão dos parâmetros com menores aROCs):

| Atributos | Qtde | NB | MLP | SVMl | SVMg | RBF | ADA | BAG | ENS | CTREE | RAN | Maior | Menor | Média |
|--------------|------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|-------|-------|
| 1 a 20 | 20 | 0,912 | 0,845 | 0,929 | 0,913 | 0,857 | 0,932 | 0,893 | 0,910 | 0,777 | 0,933 | 0,933 | 0,777 | 0,890 |
| 1 a 19 | 19 | 0,914 | 0,839 | 0,934 | 0,913 | 0,854 | 0,932 | 0,892 | 0,908 | 0,777 | 0,940 | 0,940 | 0,777 | 0,890 |
| 1 a 18 | 18 | 0,915 | 0,873 | 0,935 | 0,914 | 0,866 | 0,933 | 0,892 | 0,907 | 0,795 | 0,935 | 0,935 | 0,795 | 0,897 |
| 1 a 17 | 17 | 0,920 | 0,860 | 0,935 | 0,915 | 0,867 | 0,933 | 0,896 | 0,906 | 0,792 | 0,939 | 0,939 | 0,792 | 0,896 |
| 1 a 16 | 16 | 0,919 | 0,855 | 0,935 | 0,920 | 0,865 | 0,933 | 0,896 | 0,907 | 0,792 | 0,939 | 0,939 | 0,792 | 0,896 |
| 1 a 15 | 15 | 0,920 | 0,858 | 0,935 | 0,920 | 0,859 | 0,933 | 0,906 | 0,907 | 0,792 | 0,938 | 0,938 | 0,792 | 0,897 |
| 1 a 14 | 14 | 0,921 | 0,888 | 0,936 | 0,923 | 0,865 | 0,930 | 0,906 | 0,907 | 0,792 | 0,939 | 0,939 | 0,792 | 0,901 |
| 1 a 13 | 13 | 0,922 | 0,896 | 0,936 | 0,926 | 0,871 | 0,930 | 0,905 | 0,906 | 0,792 | 0,941 | 0,941 | 0,792 | 0,903 |
| 1 a 12 | 12 | 0,922 | 0,881 | 0,938 | 0,926 | 0,882 | 0,910 | 0,896 | 0,903 | 0,792 | 0,944 | 0,944 | 0,792 | 0,899 |
| 1 a 11 | 11 | 0,924 | 0,906 | 0,936 | 0,928 | 0,887 | 0,920 | 0,901 | 0,897 | 0,792 | 0,943 | 0,943 | 0,792 | 0,903 |
| 1 a 10 | 10 | 0,924 | 0,892 | 0,928 | 0,929 | 0,878 | 0,927 | 0,894 | 0,890 | 0,777 | 0,942 | 0,942 | 0,777 | 0,898 |
| 1 a 9 | 9 | 0,918 | 0,900 | 0,925 | 0,928 | 0,865 | 0,913 | 0,895 | 0,888 | 0,781 | 0,933 | 0,933 | 0,781 | 0,895 |
| 1 a 8 | 8 | 0,920 | 0,918 | 0,922 | 0,930 | 0,883 | 0,914 | 0,913 | 0,888 | 0,781 | 0,938 | 0,938 | 0,781 | 0,901 |
| 1 a 7 | 7 | 0,924 | 0,920 | 0,925 | 0,929 | 0,907 | 0,918 | 0,912 | 0,883 | 0,804 | 0,938 | 0,938 | 0,804 | 0,906 |
| 1 a 6 | 6 | 0,926 | 0,910 | 0,923 | 0,926 | 0,903 | 0,930 | 0,907 | 0,890 | 0,840 | 0,939 | 0,939 | 0,840 | 0,909 |
| 1 a 5 | 5 | 0,928 | 0,934 | 0,924 | 0,924 | 0,899 | 0,926 | 0,907 | 0,878 | 0,837 | 0,942 | 0,942 | 0,837 | 0,910 |
| 1 a 4 | 4 | 0,925 | 0,919 | 0,916 | 0,919 | 0,934 | 0,909 | 0,911 | 0,888 | 0,836 | 0,946 | 0,946 | 0,836 | 0,910 |
| 1 a 3 | 3 | 0,906 | 0,906 | 0,860 | 0,884 | 0,919 | 0,888 | 0,899 | 0,882 | 0,816 | 0,927 | 0,927 | 0,816 | 0,887 |
| 1 a 2 | 2 | 0,899 | 0,916 | 0,858 | 0,880 | 0,900 | 0,899 | 0,889 | 0,892 | 0,879 | 0,888 | 0,916 | 0,858 | 0,890 |
| 1 | 1 | 0,897 | 0,888 | 0,908 | 0,908 | 0,888 | 0,856 | 0,870 | 0,831 | 0,807 | 0,842 | 0,908 | 0,807 | 0,870 |
| Maior | | 0,928 | 0,934 | 0,938 | 0,930 | 0,934 | 0,933 | 0,913 | 0,910 | 0,879 | 0,946 | 0,946 | | |
| Menor | | 0,897 | 0,839 | 0,858 | 0,880 | 0,854 | 0,856 | 0,870 | 0,831 | 0,777 | 0,842 | | 0,777 | |
| Média | | 0,918 | 0,890 | 0,922 | 0,918 | 0,882 | 0,918 | 0,899 | 0,893 | 0,803 | 0,931 | | | 0,897 |
| | | | | | | | | | | | | | | |

Qtde= quantidade de atributos ; MLC- classificador de aprendizagem de máquina; aROC – área abaixo da curva ROC ; BAG – Bagging; NB – Naive-Bayes; MLP – Multilayer Perceptron; RBF – Radial Basis Function; RAN – Random Forest; ENS – Ensemble Selection; CTREE – Classification Tree (J48); ADA –Ada Boost M1; SVMl – Support Vector Machine Linear; SVMg support vector machine Gaussian.

