

DANIEL LOUZADA FERNANDES

**FRAMEWORK DE MINERAÇÃO DE DADOS
PARA ANÁLISE DE EXPERIMENTOS COM
EYE TRACKING: UMA APLICAÇÃO EM
BIOÉTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2017

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

F363f
2017
Fernandes, Daniel Louzada, 1993-
Framework de mineração de dados para análise de
experimentos com eye tracking : uma aplicação em bioética /
Daniel Louzada Fernandes. – Viçosa, MG, 2017.
xvii, 104f. : il. (algumas color.) ; 29 cm.

Inclui apêndice.

Orientador: Fabio Ribeiro Cerqueira.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f.81-93.

1. Mineração de dados (Computação). 2. Processamento de
dados oculares. 3. Bioética. 4. Medicina - Processo decisório.
5. Estatística. I. Universidade Federal de Viçosa. Departamento
de Informática. Programa de Pós-graduação em Ciência da
Computação. II. Título.

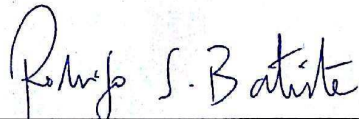
CDD 22 ed. 005.74

DANIEL LOUZADA FERNANDES

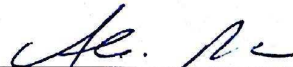
**FRAMEWORK DE MINERAÇÃO DE DADOS PARA ANÁLISE DE
EXPERIMENTOS COM EYE TRACKING: UMA APLICAÇÃO EM
BIOÉTICA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

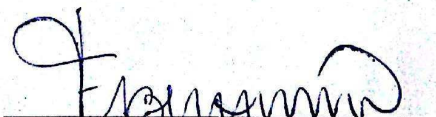
APROVADA: 18 de julho de 2017.



Rodrigo Siqueira Batista



Alcione de Paiva Oliveira



Fábio Ribeiro Cerqueira
(Orientador)

Aos meus pais, Jocimar e Ana Lucia, que sempre me guiaram pelo labirinto chamado vida.

“Em tempos de grandes mudanças, aqueles com a capacidade de aprender herdarão a terra, enquanto aqueles que já sabem tudo estarão preparados para um mundo que não existe mais.”

(Eric Hoffer)

Agradecimentos

Agradeço primeiramente a Deus por sempre ter estado comigo, me dando forças e coragem para superar os obstáculos do dia a dia. Por nunca ter me abandonado, tanto nos momentos tristes e difíceis quanto nos felizes da vida, iluminando sempre o meu caminho com a sua luz. Obrigado Senhor pela minha vida, saúde, família, amigos e por ter me permitido chegar ao término desta longa jornada.

Aos meus pais, irmãos e sobrinhos pela paciência, atenção, carinho e apoio incondicional que tiveram por mim durante essa trajetória. Agradeço imensamente a eles por sempre estarem comigo, buscando, mesmo de longe, meu bem estar, saúde e educação. Agradeço também a todos os meus familiares por sempre me motivarem a seguir meus sonhos e por torcerem pelas minhas conquistas.

Agradeço especialmente ao meu orientador, professor Fabio Ribeiro Cerqueira, por ter me aceitado como orientando desde o início, por ter confiado em mim para o desenvolvimento do nosso desafiante projeto e pela imensa ajuda na minha formação como pesquisador. Obrigado também por toda orientação, apoio, paciência, dedicação, ensinamentos e companheirismo durante toda essa caminhada acadêmica!

Aos demais professores e funcionários do Departamento de Informática que contribuíram de alguma forma em minha formação e que sempre proporcionaram a nós, estudantes, um excelente ambiente de ensino e pesquisa. Em especial, ao secretário do mestrado, Altino Alves de Sousa Filho, e aos professores André Gustavo dos Santos, Jugurta Lisboa Filho e Mauro Nacif Rocha por todos ensinamentos, esclarecimentos, conversas e risadas.

Agradeço também aos funcionários da Universidade Federal de Viçosa (UFV) em geral, por toda dedicação e empenho em fazer com que a universidade funcione tão bem para nós.

Aos meus amigos e colegas do mestrado, por todo companheirismo e troca de saberes. Em especial, a minha turma de 2015, sendo eles: Alba; Aly; Charles; Fabio; Geraldo; Gilson; Jonatas; Liliane; Marco; Michele; Rubens e Victor, por toda ajuda durante a realização das disciplinas e pesquisa! Além disso, pelas ótimas e

duradouras conversas, conselhos, comida, saídas e sociais, partidas de peteca e de imagem&ação, momentos bons de alegria e ruins de desespero compartilhados, entre outros. Saibam que este mestrado não seria o mesmo sem vocês e espero que nossa amizade continue além do campus da UFV!

Agradeço aos meus amigos pesquisadores dos departamentos de Medicina e Enfermagem e de Educação Física pelo auxílio, principalmente, no início do projeto. A minha grande amiga, Angélica Silva, pelas ótimas conversas com café e pão de queijo em sua casa. E aos demais amigos da UFV, como, da graduação em Ciência da Computação, da Associação Atlética Acadêmica das Ciências Exatas, do programa Segundo Tempo, entre outros, pela boa amizade.

Não posso deixar de agradecer também aos meus companheiros de república que me acolheram tão bem e que ajudaram tanto no meu desenvolvimento pessoal.

Aos meus amigos de Cachoeiro de Itapemirim, que mesmo com a distância, continuaram me incentivando a buscar meus sonhos e nunca desistir deles.

Por fim, mas não menos importante, quero agradecer as agências de pesquisa e financiamento Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) por todo suporte financeiro para a realização do mestrado com dedicação exclusiva.

Sumário

Lista de Figuras	viii
Lista de Tabelas	xi
Lista de Abreviaturas e Siglas	xii
Resumo	xiv
Abstract	xvi
1 Introdução	1
1.1 O problema e sua importância	3
1.2 Justificativa	5
1.3 Hipótese	5
1.4 Objetivos	6
1.5 Aspectos éticos de pesquisa	6
1.6 Organização do trabalho	7
2 Eye Tracking	8
2.1 Atenção visual	8
2.2 Rastreamento ocular	9
2.2.1 Métricas básicas e suas representações	11
2.2.2 Aplicações e trabalhos relacionados	14
2.3 Carga cognitiva	18
2.3.1 Pupilometria	18
3 Métodos Computacionais e Matemáticos	21
3.1 Mineração de dados	21
3.1.1 Algoritmos de aprendizagem máquina	24
3.2 Processo de extração de conhecimento	31

3.2.1	Transformação de dados	33
3.2.2	Avaliação dos modelos de dados	35
3.3	Análise estatística	39
3.3.1	Testes de hipótese estatística	39
4	Abordagem de Aprendizagem de Máquina	46
4.1	Framework de aprendizagem de máquina	47
4.2	Materiais e métodos	48
4.2.1	Conjunto de dados	48
4.2.2	Atributos	50
4.2.3	Ferramentas de aprendizagem de máquina	53
4.2.4	Algoritmos de aprendizagem	53
4.3	Resultados e discussão	55
4.3.1	Avaliação do desempenho do modelo	55
4.3.2	Extraindo padrões por meio de simulações com modelos MLP	56
4.3.3	Extraindo padrões por meio de regras de árvores de decisão	60
4.4	Conclusão	61
5	Abordagem de Estatística e Visualização de Dados	63
5.1	Framework de análise de séries temporais	64
5.2	Materiais e métodos	65
5.2.1	Conjunto de dados	65
5.2.2	Extração de séries temporais	67
5.2.3	Testes estatísticos	68
5.2.4	Ferramentas de visualização e estatística	69
5.3	Resultados e discussão	71
5.3.1	Análise estatística	71
5.3.2	Análise exploratória	72
5.4	Conclusão	78
6	Conclusões	79
	Referências Bibliográficas	81
	Apêndice A Artigo publicado no International Conference on Computational Science 2017 (Qualis A1)	94

Lista de Figuras

2.1	Aparelho utilizado para registrar os movimentos dos olhos. Fonte: Yarbus [1967].	10
2.2	Atuais dispositivos de rastreamento ocular. A subfigura (a) demonstra um modelo de rastreador do tipo fixo e a subfigura (b) exhibe um modelo do tipo móvel. Fonte: Lupu & Ungureanu [2013].	11
2.3	Distinção dos diferentes tipos de métricas oculares. Fonte: Blascheck et al. [2014].	13
2.4	Exemplos de relatórios visuais: (a) Gaze plots, (b) Heat maps, (c) Clusters, (d) AOI. Fonte: Strandvall [2008].	14
2.5	Ilustração de respostas pupilares na execução de tarefas cognitivas. Fonte: Hossain & Yeasin [2014].	20
2.6	Músculos lisos responsáveis pela dilatação e constrição da pupila. Fonte: Beatty & Lucero-Wagoner [2000].	20
3.1	Construção de um modelo de classificação. Fonte: Tan et al. [2006]. . . .	25
3.2	Modelo de perceptron.	26
3.3	Tipos de função de ativação: (a) linear, (b) limiar, (c) sigmóide e (d) tangente hiperbólica.	28
3.4	Representação simplificada do funcionamento de uma rede MLP. Fonte: Tan et al. [2006].	29
3.5	Processo de classificação usando árvore de decisão. Fonte: Tan et al. [2006].	30
3.6	Visão geral das etapas que constituem o processo KDD. Fonte: Fayyad et al. [1996].	31
3.7	Visão geral da validação cruzada 10-partes. Fonte: Raschka [2015]. . . .	36
3.8	Matriz de confusão para um problema de classificação binária. Fonte: Tan et al. [2006].	37

3.9	Comparação do desempenho de dois métodos por meio da área sob o curva ROC.	39
3.10	Fluxograma para escolher o teste estatístico.	41
3.11	Verificação da normalidade por meio do histograma e Q-Q plots. As sub- figuras (a) e (b) demonstram o comportamento de dados que não seguem uma distribuição normal, enquanto as subfiguras (c) e (d) apresentam o formato de sino e a linearidade dos pontos, sugerindo, portanto, que os dados são normalmente distribuídos.	42
4.1	Pipeline de extração de conhecimento.	47
4.2	Capturas de tela do filme que os participantes assistiram, sendo: (a) cena da Maggie e (b) cena do Frankie. Fonte: Eastwood & Hagglund [2005]. .	49
4.3	Curva ROC do modelo de MLP gerado. A AUC média é equivalente a 0,903.	56
4.4	Probabilidade prevista de aceitar a participar da prática da eutanásia versus tempo total de duração das fixações gerais.	57
4.5	Probabilidade prevista de aceitar a participar da prática da eutanásia versus tempo de duração das fixações no tubo traqueal.	57
4.6	Probabilidade prevista de aceitar a participar da prática da eutanásia versus quantidade de fixações no peito de Maggie (que contém um cordão com um crucifixo).	58
4.7	Probabilidade prevista de aceitar a participar da prática da eutanásia versus tamanho da dilatação vertical média da pupila.	59
4.8	Probabilidade prevista de aceitar a participar da prática da eutanásia versus tamanho da dilatação geral média da pupila.	59
4.9	Padrões gerados por um modelo C4.5 sensível ao custo. As subfiguras (a) e (b) demonstram as regras extraídas e a árvore de decisão pelo modelo C4.5 sensível ao custo.	60
5.1	Pipeline de extração de conhecimento.	65
5.2	Procedimento para obtenção de dados pupilares.	68
5.3	Comportamento médio do diâmetro pupilar geral dos grupos de partici- pantes ao longo do experimento por meio de (a) boxplot e (b) gráfico de linhas.	73
5.4	Múltiplos boxplots dos dados do diâmetro pupilar dos grupos de parti- cipantes durante os 13 bins.	75

5.5	Múltiplos gráficos de linhas de séries temporais do comportamento do diâmetro pupilar dos grupos participantes durante os 13 bins.	76
-----	--------------------------------------------------------------------------------------------------------------------------------------------	----

Lista de Tabelas

4.1	Atributos restantes após o passo de seleção manual.	51
4.2	Nível de importância dos atributos com base na medida F.	52
4.3	Avaliação estatística (validação cruzada 10-partes) dos modelos gerados a partir de vários algoritmos de aprendizagem com e sem CSC.	55
5.1	Resumo das 13 cenas do trecho filme.	66
5.2	Resultados do teste de Mann Whitney U para os 13 bins.	72

Lista de Abreviaturas e Siglas

Ac:	Acurácia
ANOVA:	<i>Analysis of Variance</i>
AOI:	<i>Area of Interest</i>
ARFF:	<i>Attribute-Relation File Format</i>
ASL:	<i>Applied Science Laboratories</i>
AUC:	<i>Area Under the Curve</i>
CAPES:	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CFS:	<i>Correlation-based Feature Selection</i>
CSC:	Classificação Sensível ao Custo
CSV:	<i>Comma-separated Values</i>
CAAE:	Certificado de Apresentação para Apreciação Ética
EEG:	Eletroencefalograma
FAPEMIG:	Fundação de Amparo a Pesquisa do Estado de Minas Gerais
fMRI:	<i>Functional Magnetic Resonance Imaging</i>
FN:	<i>False Negatives</i>
FP:	<i>False Positives</i>
GNU:	<i>GNU is Not Unix</i>
Hz:	<i>Hertz</i>
H_0 :	Hipótese Nula
H_1 :	Hipótese Alternativa
IBM:	<i>International Business Machines</i>
IC:	Intervalo de Confiança
ICCS:	<i>International Conference on Computational Science</i>

ID:	<i>Identifier</i>
IG:	<i>InfoGain</i>
KDD:	<i>Knowledge-Discovery in Databases</i>
<i>k</i> NN :	<i>k-Nearest Neighbour</i>
MET:	<i>Mobile Eye Tracking</i>
MLP:	<i>Multilayer Perceptron</i>
ms:	Milissegundos
Pr:	Precisão
RNA:	Rede Neural Artificial
ROC:	<i>Receiver Operating Characteristic</i>
SMO:	<i>Sequential Minimal Optimization</i>
Sn:	Sensibilidade
SNS:	Sistema Nervoso Simpático
SNP:	Sistema Nervoso Parassimpático
SVM:	<i>Support Vector Machine</i>
TN:	<i>True Negatives</i>
TNR:	<i>True Negative Rate</i>
TP:	<i>True Positives</i>
TPR:	<i>True Positive Rate</i>
TEPR:	<i>Task-Evoked Pupillary Responses</i>
UFV:	Universidade Federal de Viçosa
WEKA:	<i>Waikato Environment for Knowledge Analysis</i>

Resumo

FERNANDES, Daniel Louzada, M.Sc., Universidade Federal de Viçosa, julho de 2017. **Framework de mineração de dados para análise de experimentos com eye tracking: Uma aplicação em bioética.** Orientador: Fábio Ribeiro Cerqueira.

A capacidade de tomar decisões e de resolver problemas representa uma habilidade cognitiva importante para o cotidiano das pessoas. Na prática clínica, principalmente nas situações que envolvem questões bioéticas, a tomada de decisão do profissional da saúde advém de uma complexa interação entre o ambiente de trabalho e os processos neurais subjacentes a cognição e a emoção, resultando em escolhas e julgamentos que levam em consideração diversos aspectos. Com o avanço da tecnologia, pesquisadores têm desenvolvido novas técnicas que investigam a tomada de decisão por meio da mensuração da carga cognitiva durante a realização de tarefas. Nos últimos anos, houve um aumento do interesse em pesquisas relacionadas com a atenção visual por meio de dispositivos de rastreamento ocular. Por consequência, este tornou-se um importante tema de investigação, aplicado em vários campos para estudar e compreender os padrões entre sujeitos. No entanto, a grande quantidade de dados brutos derivados pelo rastreador torna a tarefa de análise de dados complexas, dificultando, o processo de descoberta de padrões. Além disso, uma vez que as ferramentas existentes apresentam limitações para suportar grandes conjuntos de dados e há uma ausência de registros na literatura para auxiliar na condução desse tipo de pesquisa, este estudo tornou-se um desafio. Portanto, torna-se recomendado a combinação de metodologias para aprimorar a apreciação dos dados produzidos, otimizando a avaliação da atenção visual no processo cognitivo. Nesta dissertação, é proposta duas abordagens que além de funcionar como um arcabouço computacional para processamento dos dados oculares e extração de importantes padrões, também direciona para futuras pesquisas na área. A primeira abordagem

é baseada em algoritmos de aprendizagem de máquina (redes multilayer perceptron e árvores de decisão C4.5) para construir modelos capazes de prever a decisão bioética por meio de padrões extraídos de dados oculares. Já a segunda abordagem se baseia na utilização de métodos de visualização de dados e estatística para facilitar a interpretação e avaliação dos dados por meio da validação estatística e análise exploratória. Os resultados de ambas as abordagens mostraram uma conexão entre comportamento cognitivo e a tomada de decisão. Desta forma, infere-se que os pipelines dos procedimentos propostos são eficientes para testar hipóteses científicas relacionadas com o papel da atenção visual e tomadas de decisões clínicas.

Abstract

FERNANDES, Daniel Louzada, M.Sc., Universidade Federal de Viçosa, July, 2017.
**Data mining framework for analysis of experiments with eye tracking:
An application in bioethics.** Adviser: Fábio Ribeiro Cerqueira.

The ability to make decisions and solve problems represents an important cognitive ability for people's daily lives. In clinical practice, especially in situations involving bioethical issues, decision making by healthcare professionals comes from a complex interaction between the work environment and the neural processes underlying cognition and emotion, resulting in choices and judgments that take into account different aspects. With the advancement of technology, researchers have developed new techniques that investigate decision making by measuring cognitive load while performing tasks. In recent years, there has been an increased interest in research related to visual attention through eye tracking devices. Consequently, this has become an important research topic, applied in various fields to study and understand the patterns between subjects. However, the large amount of raw data derived from the device makes the task of data analysis complex, making it difficult to discover patterns. In addition, as existing tools have limitations to support large datasets and due to a lack of records in the literature to assist in conducting this type of research, studies with eye tracking data have become a challenge. Therefore, it is recommended the combination of methodologies to improve the appreciation of the produced data, optimizing the evaluation of visual attention in the cognitive process. In this dissertation, two approaches are proposed to serve as a computational framework for processing eye data and extracting important patterns, as well as to direct future research in the area. The first approach is based on machine learning algorithms (multilayer perceptron networks and decision trees C4.5) to construct models capable of predicting bioethical decision using patterns extracted from eye data. The second approach is based on the use of data visualization and statistical

methods to facilitate the interpretation and evaluation of the data through statistical validation and exploratory analysis. The results of both approaches showed a connection between cognitive behavior and decision making. In this way, it is inferred that the pipelines of the proposed procedures are efficient to test scientific hypotheses related to the role of visual attention and clinical decision making.

Capítulo 1

Introdução

Diante de constantes avanços tecnológicos das últimas décadas, fortes laços de relacionamento entre a ciência da computação e as demais áreas foram criados, resultando em um nascimento rico e progressivo de novas relações que garantem um futuro brilhante para a ciência [Denning, 2005].

Nos anos 70, visando alcançar melhorias na produtividade médica, a área da saúde explorou as capacidades extraordinárias da computação por meio de seu uso intensivo tanto nas abordagens dos processos biológicos quanto médicos. Logo, possibilitou grandes avanços e resultados eficientes na prestação de cuidados com a saúde [Collen, 1986; Patel & Kaufman, 1998]. Essa relação entre as respectivas áreas acarretou na origem de um novo ramo da ciência, conhecido como Informática Biomédica.

A Informática Biomédica é definida como a ciência da informação de cunho interdisciplinar que desenvolve: o estudo, a compreensão e gerenciamento de dados, informações e conhecimentos biomédicos por meio da aplicação das tecnologias de informação e comunicação na saúde [Collen, 1986; Patel & Kaufman, 1998].

No entanto, segundo registros encontrados na literatura, alguns cientistas alegam que as definições geralmente utilizadas para este ramo da ciência são estreitas, ou seja, que a ideia apresentada é que a computação e as tecnologias são utilizadas apenas como ferramentas de apoio para medicina. Na concepção destes pesquisadores, a Informática Biomédica tem uma contribuição mais profunda do que aparenta para a comunidade científica, não limitando-se apenas a informatização dos processos e do fluxo de trabalho, mas, concentrando-se na modelagem da formação da prática clínica para a constante melhoria dos cuidados da saúde por meio do entendimento, da expansão e da organização do conhecimento humano [Shortliffe, 1984; Patel & Kaufman, 1998; Musen et al., 2002; Bernstam et al., 2010].

Neste sentido, a Ciência Cognitiva, campo científico multidisciplinar que estava ocupando há muitos anos um lugar na periferia da Informática Biomédica, começou a ser mais explorada pelas funções desempenhadas no progresso de compreensão dos processos cognitivos [Tourassi et al., 2013]. A Ciência Cognitiva tem como objetivo o estudo científico da cognição e de seu papel na agência inteligente por meio da interação entre os campos da Linguística, Psicologia, Filosofia, Neurociência, Antropologia e Inteligência Artificial [Patel & Kannampallil, 2015].

As abordagens teóricas e metodológicas da Ciência Cognitiva permitiram a realização de estudos eficazes e detalhados a respeito de uma ampla gama de domínios práticos, como: usabilidade de sistemas médicos; formação e orientação médica e educativa; racionalização e melhoria do fluxo de trabalhos e dos processos clínicos; e compreensão do processo de julgamento, raciocínio e tomada de decisão clínica [Evans & Patel, 2013].

Em particular, há um aumento significativo de interesse a respeito da compreensão do processo decisório clínico. A capacidade dos profissionais da saúde em tomar decisões consiste em uma habilidade cognitiva importante, responsável pela análise de diversas opções existentes e a escolha de uma alternativa, com base em caminhos de ações antecipadamente estabelecidos, os quais uma pessoa deverá seguir para a resolução de um problema [Chiavenato, 2004; Primak, 2008]. Na prática clínica, principalmente quando se trata de questões bioéticas, – mormente em momentos nos quais não existem consenso moral, como, por exemplo, na prática: da eutanásia; do aborto; da fertilização *in-vitro*; entre outros –, a tomada de decisão deriva de uma complexa interação entre o ambiente de trabalho e os processos neurais subjacentes à cognição e à emoção, resultando em escolhas e julgamentos que levam em consideração o conhecimento, o risco, a incerteza e o arrependimento [Gutnik et al., 2006; Reyna & Rivers, 2008; Sugarman & Sulmasy, 2010; Brosch et al., 2013].

A bioética é uma questão fundamental para a atuação prática dos profissionais da saúde. Segundo Goldim [2006], é uma reflexão compartilhada, complexa e interdisciplinar sobre a adequação das ações que envolvem a vida e o viver. Desde sua emergência, a bioética ou ética da vida, tem abrangido uma gama completa de questões que vão desde decisões clínicas, condutas profissionais e políticas públicas das ciências da vida e cuidados da saúde até direitos humanos, obrigações éticas e dimensões morais [Figueiredo, 2011; Junqueira, 2011]. A bioética clínica, em particular, é um dos ramos mais complexos da bioética, pois, está diretamente relacionada a análise de conflitos, problemas ou questões de ordem moral ou ética polêmicas que surgem na prática clínica e assistencial e em pesquisas com seres humanos [Figueiredo, 2011].

Compreender e elucidar o processo decisório em bioética clínica apresenta alguns desafios interessantes e valiosos. Observa-se que a caracterização das relações causais entre os fatores que desencadeiam a tomada de decisão é uma tarefa árdua devido a complexidade de articulação dos aspectos neuropsicológicos e neurobiológicos subjacentes dos seres humanos que são considerados no momento da ação. Visto isso, cientistas buscam maneiras de preencher a lacuna a respeito da influência dos processos mentais nas tomadas de decisões, e uma das formas encontradas é por meio da investigação da atenção visual [Görgen, 2010; Tourassi et al., 2013].

Os aspectos neurofisiológicos da visão são fundamentais para os processos perceptivos e cognitivos, por causa da forma analítica que os olhos conjugam as informações revelando padrões sobre a pesquisa visual [Siqueira-Batista et al., 2015]. Sendo, portanto, útil para os estudos de hipóteses científicas que desejam relacionar a visão com as atividades cognitivas evocadas por tarefa.

Consequentemente, esses estudos serviram como chave para moldar e estruturar o desenvolvimento de sistemas, ferramentas tecnológicas e mecanismos viáveis que realizam a mensuração e caracterização de certas regularidades do complexo sistema de processamento de informação humana por meio da carga cognitiva [Shorliffe, 1984; Patel & Kannampallil, 2015; Zagermann et al., 2016].

1.1 O problema e sua importância

O mapeamento da atenção visual é realizado por meio da utilização de tecnologias que possibilitam a investigação dos processos cognitivos. Atualmente, encontra-se na literatura diferentes técnicas e tecnologias que desempenham essa função, como, o eletroencefalograma (EEG); imagem por ressonância magnética funcional, do inglês, functional magnetic resonance imaging (fMRI); e o rastreador ocular [Görgen, 2010; Plöchl et al., 2012]. Este último, em especial, também conhecido como eye tracking, é um dispositivo tecnológico que vem ganhando popularidade na comunidade científica graças a sua aplicação em diversos campos de pesquisa e aos resultados alcançados [Coco, 2009; Pärnamets et al., 2015]. Este dispositivo possibilita o mapeamento da atenção visual por meio da medição e do registro dos movimentos dos olhos de acordo com estímulos visuais apresentados [Galgani et al., 2009].

Para analisar e interpretar os dados providos pelo dispositivo de rastreamento ocular, é necessário escolher quais métricas oculares quantitativas serão averiguadas dentre uma ampla variedade de medidas existentes, como, por exemplo, duração do

olhar, sequência e número fixações, tamanho da pupila, taxa de intermitência, amplitude de sacada, entre outros [Wu et al., 2008; Barreto, 2012]. Essas métricas são utilizadas para examinar a relação entre a carga cognitiva e a carga de visualização, possibilitando a obtenção de discernimentos sobre os modos de raciocínio e resolução de problemas dos sujeitos [Andrienko et al., 2012; Tourassi et al., 2013]. Entretanto, a compreensão dos efeitos do esforço cognitivo por meio dos movimentos dos olhos requer uma análise robusta das respostas obtidas pelos dados oculares [Hossain & Yeasin, 2014].

Em razão disso, a parte mais desafiadora das pesquisas de rastreamento ocular está na análise de seus dados. A grande quantidade de dados brutos gerados pelo dispositivo e o vínculo desses dados aos processos cognitivos torna a tarefa de análise complexa e dispendiosa, dificultando assim, o processo de descoberta de conhecimento [Privitera et al., 2008; Andrienko et al., 2012; Blascheck et al., 2014; Muthumanickam et al., 2016]. Além do mais, os padrões obtidos nos registros podem variar entre os sujeitos avaliados devido possivelmente serem afetados por algum tipo de interferência ou perturbação, como, por exemplo, a ocorrência de imperfeições durante o processo de aquisição dos dados; movimentos não intencionais dos olhos e o estado de espírito, o que torna a análise um desafio inviável de ser resolvido de forma manual [Wu et al., 2008].

Por consequência disso, ferramentas computacionais têm sido desenvolvidas para analisar os movimentos dos olhos e extrair conclusões sobre os comportamentos dos sujeitos. Porém, as ferramentas disponíveis no presente momento são limitadas e sozinhas não possibilitam uma fácil interpretação dos dados [Räihä et al., 2005; Bednarik, 2007].

Assim, para determinar qual abordagem utilizar, deve-se selecionar as técnicas que melhor suportam a análise desejada. Entretanto, é necessário a realização de um levantamento exaustivo da literatura com intenção de encontrar as técnicas mais adequadas para o estudo a ser investigado [Blascheck et al., 2014]. Pesquisas atuais recomendam a combinação de metodologias com o objetivo de aprimorar a apreciação dos dados produzidos, otimizando assim, a análise da influência da atenção visual [Andrienko et al., 2012; Barreto, 2012; Yu et al., 2012; Blascheck et al., 2014].

Observando essa perspectiva, este trabalho propõe o uso de duas abordagens computacionais distintas que permitem inferir em padrões que possam ser utilizados para testar hipóteses científicas correlacionadas à atenção visual, cognição e tomada de decisão. As abordagens, (i) baseada em algoritmos de aprendizagem de máquina e (ii) em testes estatísticos e técnicas de visualização de dados, são capazes de apoiar o

estudo da investigação médica no processo de tomada de decisão em bioética clínica, particularmente neste trabalho, no caso de eutanásia¹.

1.2 Justificativa

A ideia de estudar os dados derivados de rastreadores oculares para associar as características visuais à tomada de decisões em uma questão moral tão complexa foi proposta pela primeira vez por Siqueira-Batista et al. [2015], e recomendada por Pärnamets et al. [2015].

Os estudos presentes na literatura apontam que as pesquisas desenvolvidas com sistemas de rastreamento ocular motivam o interesse profissional e acadêmico principalmente: pela proliferação de soluções alcançadas; pelo preço relativamente acessível; por ser uma abordagem poderosa para a descoberta de padrões e por permitir estudar a dinâmica em tempo real dos sistemas cognitivos [Barreto, 2012; Sasson & Elison, 2012].

Por ser um campo jovem, aberto e com múltiplas fronteiras disciplinares, ainda não existem estudos aprofundados na literatura sobre o auxílio da computação para tomada de decisão em bioética, uma vez que pesquisas e estudos abordam fragmentos do assunto, sem entretanto, chegar a uma sugestão integrativa e concreta o suficiente para explicar e antecipar decisões [Pärnamets et al., 2015; Siqueira-Batista et al., 2015].

Este trabalho se torna relevante à medida que apresenta um estudo novo e interessante que envolve múltiplas áreas e por contribuir positivamente para o conhecimento acerca do papel da atenção visual no processo decisório, em especial, na bioética clínica. Além disso, é o primeiro trabalho a levar a cabo a proposta de Siqueira-Batista et al. [2015], e Pärnamets et al. [2015].

1.3 Hipótese

Com o uso de técnicas de aprendizagem de máquina, estatística e de visualização de dados é possível elaborar um arcabouço computacional matemático que auxilie a análise e a descoberta de conhecimento em dados de eye tracking, auxiliando o processo decisório.

¹Definida como uma prática para abreviar a vida, a fim de aliviar ou evitar o sofrimento dos pacientes.

1.4 Objetivos

Nota-se na literatura, a ausência de trabalhos que conduzem estudos de investigação da atenção visual por meio de dados oculares. Visto isso, os métodos descritos neste trabalho estão focados no problema de análise da grande quantidade de dados derivados pelo eye tracking. O objetivo deste trabalho é apresentar um framework de mineração de dados, constituído por duas abordagens distintas e complementares, que possibilite uma investigação mais eficaz dos conjuntos de dados fornecidos pelo rastreador ocular, a fim de realizar uma análise comparativa e retornar as diferenças entre categorias decisórias.

Além do objetivo principal, este trabalho propõe como objetivos específicos:

- Investigar e levantar as técnicas e sistemas computacionais existentes no estado da arte para análise de dados oculares juntamente com suas limitações e métricas utilizadas;
- Estruturar e apresentar dois pipelines que descreverão o processo completo da extração de conhecimento sobre os aspectos cognitivos que possam ser rastreados pelos dados oculares;
- Construir e validar modelos preditivos capazes de prever a decisão médica com base nos dados e atributos oculares por meio de algoritmos de aprendizagem de máquina;
- Definir métodos estatísticos e de visualização de dados de maneira sistemática para verificar a existência de diferenças significativas entre o comportamento da variabilidade do diâmetro pupilar com a tomada de decisão em bioética;
- Extrair padrões que nos permitam entender a conexão entre o que vemos e como reagimos com base na informação que processamos.

1.5 Aspectos éticos de pesquisa

O estudo abordado nesta dissertação envolveu a participação de seres humanos. Desta forma, o projeto de pesquisa que deu origem a esta pesquisa foi submetido e aprovado pelo Comitê de Ética em Pesquisa da UFV (CAAE N° 25353714.8.0000.5153), em conformidade com as resoluções 196/1996 e 466/2012 do Conselho Nacional de Saúde do Brasil. Todas as normas e regulamentações propostas no Brasil, foram atendidas.

1.6 Organização do trabalho

Além do capítulo introdutório já apresentado, esta dissertação foi estruturada em mais cinco capítulos, que serão descritos a seguir.

O Capítulo 2, apresenta uma breve descrição sobre os fundamentos teóricos envolvidos no estudo, como, a atenção visual, a tecnologia eye tracking, o contexto histórico dessa ferramenta tecnológica, as métricas oculares usadas, os trabalhos da literatura correlacionados e carga cognitiva.

No Capítulo 3, são apresentadas as informações referentes as principais técnicas de mineração de dados e estatística utilizadas nos procedimentos propostos.

Nos Capítulos 4 e 5, serão descritas, detalhadamente e distintamente, as abordagens desenvolvidas para análise dos dados oculares, incluindo: introdução as abordagens, os materiais e métodos, os resultados obtidos durante os testes realizados, as discussões promovidas e por último, as conclusões inferidas.

Por fim, no Capítulo 6, são apresentados as conclusões da dissertação, juntamente com as suas contribuições e futuras orientações para a pesquisa.

Após as conclusões, há o apêndice de anexos, onde o artigo, publicado pode ser visto.

Capítulo 2

Eye Tracking

Este capítulo tem como objetivo apresentar em forma de revisão os conceitos-chaves, contexto histórico e pesquisas relacionadas com o tema deste estudo, contextualizando os tópicos necessários para a compreensão do problema que será abordado.

2.1 Atenção visual

A visão é o sentido mais desenvolvido e importante do ser humano, devido ser o responsável pela obtenção de grande parte do conhecimento do mundo externo. O local onde o olhar é fixado e seus movimentos estão associados com o foco de atenção dos sujeitos, possibilitando, portanto, a revelação das características sobre seu comportamento e a reflexão do seu estado mental [Beatty, 1982].

A atenção visual, particularmente, é o domínio da atenção com o maior número de estudos empíricos presentes na literatura neurocientífica, por causa da simplicidade em apresentar e manipular estímulos visuais [Araujo & Carreiro, 2009; Haladjian & Montemayor, 2015]. A atenção visual consiste em um conjunto de operações cognitivas altamente flexíveis, responsáveis por filtrar seletivamente informações visuais relevantes de forma automática ou voluntária durante situações em que as ações são baseadas nas interações com o ambiente [Chun & Wolfe, 2001; Görden, 2010].

A atenção visual pode ser classificada em diferentes tipos que variam conforme vários aspectos, como: controle atencional; estímulos e sujeitos. Os principais tipos encontrados na literatura, segundo Görden [2010] e Haladjian & Montemayor [2015], são:

- Atenção ativa: o sujeito focaliza sua atenção de forma dirigida. Portanto,

ocorre de maneira voluntária. Esta atenção também é conhecida como atenção de cima para baixo ou endógena;

- Atenção passiva: ocorre de forma automática ou involuntária por meio de estímulos apresentados no ambiente. Também conhecida como atenção de baixo pra cima ou exógena;
- Atenção aberta: também conhecida como atenção mensurável, isto é, quando é possível inferir a localização do direcionamento dos olhos de uma pessoa;
- Atenção secreta: é o tipo de atenção na qual uma pessoa pode fingir que está olhando para um lugar enquanto simultaneamente está atento ou pensando em outro. É o tipo que não é visível externamente, também conhecida como atenção da mente.

A respeito dos tipos de atenção aberta e secreta, particularmente, existem diferentes técnicas que permitem a investigação desses dois tipos de atenção. No caso da atenção secreta, deseja-se descobrir como funciona a atenção interna de uma pessoa. Para isso, pode-se utilizar o EEG ou o fMRI para fazer o mapeamento dos processos neuronais. Já quando se trata da atenção aberta, espera-se saber o que realmente guia a atenção da pessoa. Neste caso, a técnica mais recomendável é o dispositivo de rastreamento ocular [Görge, 2010].

Conforme os conceitos dos tipos de atenção listados acima, pode-se alegar que o comportamento visual humano, em geral, é o resultado da interação entre esses tipos de atenção. Visto isso, nesta pesquisa, foi utilizada a tecnologia de rastreamento ocular para investigar se o foco atencional dos profissionais da saúde ao presenciarem um momento delicado que envolve a questão de fim de vida, tem alguma relação com a tomada de decisão realizada.

2.2 Rastreamento ocular

Rastreamento ocular, do inglês, eye tracking, é um conjunto de tecnologias que possibilita a medição e o registro da posição e dos movimentos dos olhos de um indivíduo com intuito de mapear a atenção visual em um ambiente real ou controlado. Ele determina em quais áreas uma pessoa fixa a sua atenção, por quanto tempo e em qual ordem segue a sua exploração visual [Barreto, 2012]. Em outras palavras, é um dispositivo que pode capturar com alta resolução espacial e temporal o cursor visual [Yu et al., 2012].

O uso de tecnologias com objetivo similar à tecnologia eye tracking não é uma novidade. Os estudos sobre os movimentos dos olhos tiveram seu início no final do século XIX, pelo francês Louis Émile Javal, sendo o primeiro a descrever sobre os movimentos dos olhos [Barreto, 2012]. Com o passar dos anos, diversos métodos de estudos foram desenvolvidos e aperfeiçoados graças ao avanço tecnológico, contribuindo assim, com conhecimentos sobre os movimentos oculares.

Os primeiros rastreadores oculares eram dispositivos mecânicos e causavam desconforto nos usuários devido ao seu tamanho e sua capacidade de invasão (Figura 2.1). Nos dias de hoje, os dispositivos são mais compactos, não invasivos e não exigem chapelaria, somente alguns pequenos anexos físicos a serem usados pelos usuários (Figura 2.2) [Mansour & Flowe, 2010].

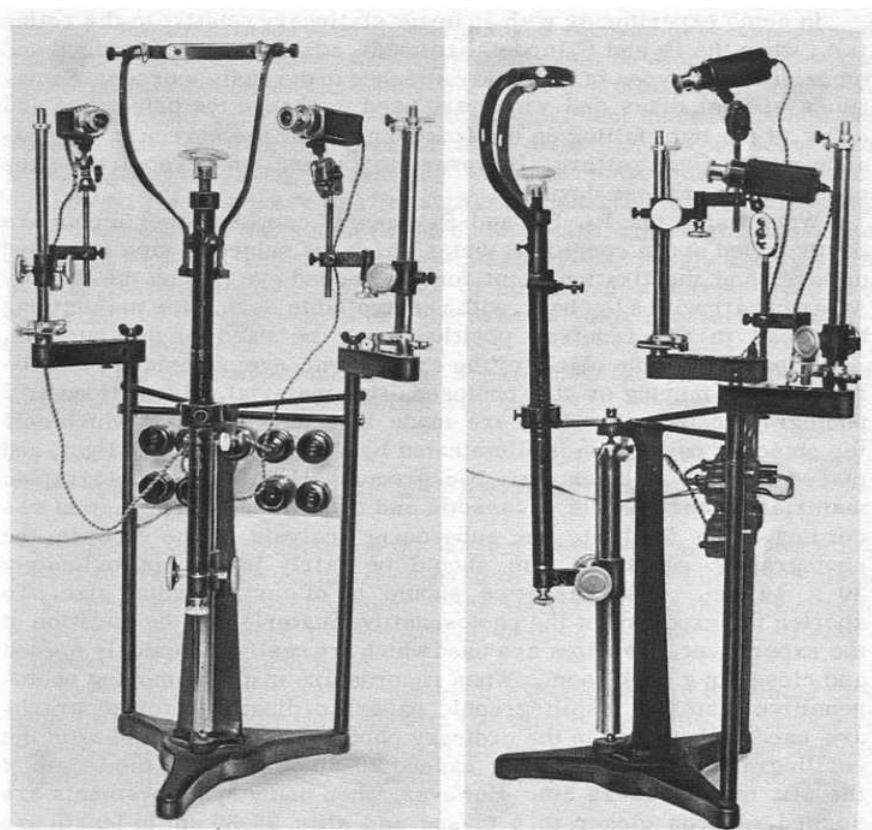


Figura 2.1. Aparelho utilizado para registrar os movimentos dos olhos. Fonte: Yarbus [1967].

Segundo Barreto [2012], é comum confundir a data em que o estudo começou a ser realizado com a data em que a técnica por meio da tecnologia começou a ser aplicada. O primeiro sistema de eye tracking comercial para computadores foi produzido em 1988, nos Estados Unidos, pela empresa LC Technologies. Esse sistema era composto por algoritmos avançados de processamento digital de imagens e por

métodos de focalização automática que possibilitavam a localização da pupila, a reflexão da córnea e a calibração da fixação ocular com maior precisão e consistência, mesmo com a cabeça do usuário em movimento.

Existem dois tipos de dispositivos de rastreamento ocular, sendo eles: o fixo e o móvel. O tipo fixo é responsável por registrar os movimentos dos olhos a uma certa distância, geralmente, são integrados a um monitor. Estes dispositivos requerem algoritmos mais avançados de processamento de imagens, por realizar a captura dos movimentos da cabeça, a detecção de faces e a obtenção da orientação dos rostos em 3D. O tipo móvel é transportado com os usuários e são ideais para as atividades que envolvem a mobilidade, principalmente da cabeça. Esse tipo é baseado em sistemas embarcados, são de baixo custo, não necessitam de algoritmos sofisticados para processamento de imagem e possuem bons desempenhos em operações de tempo real [Galgani et al., 2009; Barreto, 2012; Yu et al., 2012].

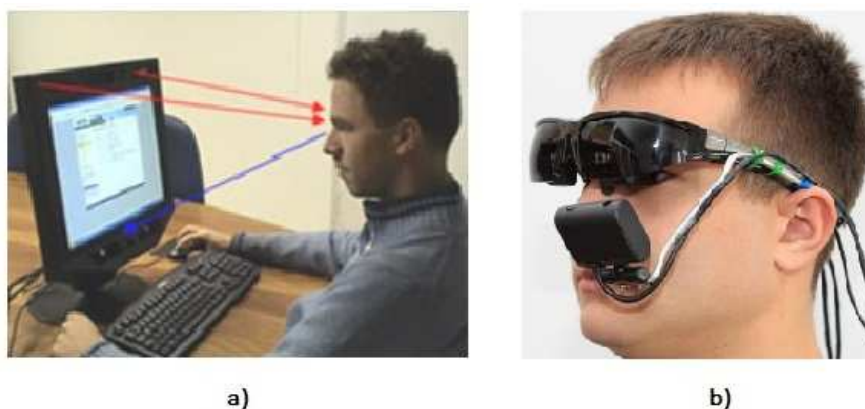


Figura 2.2. Atuais dispositivos de rastreamento ocular. A subfigura (a) demonstra um modelo de rastreador do tipo fixo e a subfigura (b) exibe um modelo do tipo móvel. Fonte: Lupu & Ungureanu [2013].

2.2.1 Métricas básicas e suas representações

Os dados de rastreamento ocular consistem em registros sobre as posições e os horários das fixações dos olhos. Cada registro inclui o identificador de usuário, o tempo, a posição no espaço de exibição (coordenadas x e y) e a duração da fixação. Os registros também podem incluir outros atributos, por exemplo, identificador de estímulo quando diferentes estímulos são usados na coleta de dados [Andrienko et al., 2012].

De acordo com Barreto [2012] e Wu & Munzner [2015], a análise dos dados providos pelo dispositivo de rastreamento ocular geralmente faz uso de um conjunto

de métricas quantitativas, como, por exemplo:

- Fixações: momento em que os olhos estão relativamente fixos em alguma área de interesse (area of interest - AOI) compreendendo ou decifrando informações, tendo uma duração mínima de 120 milissegundos (ms);
- Duração do olhar: duração acumulativa de uma série de fixações consecutivas dentro de uma AOI;
- Sacada: movimento ocular que ocorre entre fixações.
- Trajetória ocular (scanpaths): são os caminhos dos olhos ao examinar o campo visual e analisar qualquer tipo de informação visual;
- Tamanho da pupila: aumento ou redução do tamanho do diâmetro pupilar;
- Taxa de intermitência: piscar dos olhos; entre outros.

A Figura 2.3 a seguir, ilustra um exemplo que distingue algumas das métricas citadas. Os pontos de observação (gaze point) são incorporados espacialmente e temporariamente em fixações (fixation). As fixações são interligadas por sacadas (saccade) e possuem uma certa duração representada pelo raio, isto é, o tamanho do círculo que o número está contido. Uma ordem temporal de fixações é um olhar (gaze - círculo azul), no entanto, somente se as fixações estiverem dentro de uma AOI. Uma AOI é uma região de interesse específica sobre o estímulo. A sacada de uma AOI para a próxima é chamado de transição (transition). Uma sequência completa de fixações e sacadas é chamada de trajetória ocular [Blascheck et al., 2014].

Com a análise das métricas, os pesquisadores conseguem determinar uma vasta quantidade de informações e a obtenção de discernimentos sobre os modos de raciocínio e resolução de problemas dos sujeitos [Andrienko et al., 2012; Tourassi et al., 2013]. No entanto, mesmo que as métricas apresentadas sejam as mais comuns na literatura, elas nem sempre são as mais aptas a um estudo. Portanto, é necessário realizar uma análise das mesmas sobre o estudo que será feito para decidir quais são as mais adequadas.

Uma forma de demonstrar e resumir os possíveis padrões visuais obtidos pelo uso das métricas oculares é por meio dos relatórios visuais ou das representações gráficas. Essas representações em forma de mapas, são capazes de informar de maneira exploratória, a trajetória ocular dos usuários [Barreto, 2012]. Uma das técnicas mais conhecida é o fixation map. Esta técnica visual, em forma de registro

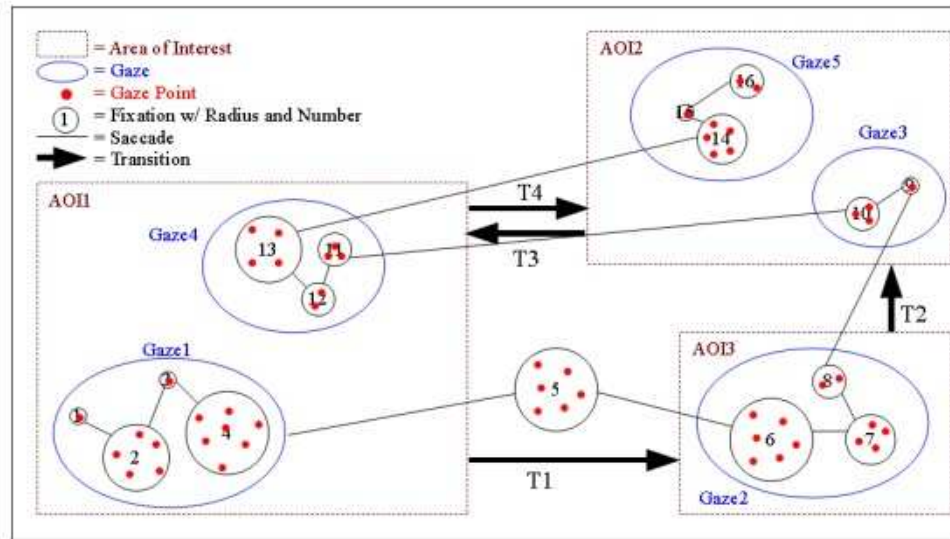


Figura 2.3. Distinção dos diferentes tipos de métricas oculares. Fonte: Blascheck et al. [2014].

bidimensional, representa todas as fixações que estão sendo analisadas por meio de círculos sobrepostos na imagem que foi observada [Wooding, 2002]. Entretanto, devido a sobreposição dos círculos, uma forte desordem visual é produzida quando as fixações de múltiplos sujeitos estão sendo analisadas, dificultando assim a análise [Ristovski et al., 2013].

Outro exemplo de relatório visual bem popular que exibe a distribuição da atenção dos usuários sobre o espaço de exibição é o heat map [Bojko, 2009; Andrienko et al., 2012]. Nesta representação, as áreas mais intensas ou “quentes” indicam o local na imagem em que os sujeitos analisados fixaram com maior frequência a sua atenção [Bojko, 2009; Barreto, 2012]. Heat maps são eficazes, intuitivos e fornecem uma rápida visão da distribuição dos dados [Ristovski et al., 2013]. Contudo, assim como os fixation maps, os heat maps não suportam a ordenação temporal dos dados. Desta forma, ambas abordagens não são ideais quando o objetivo é analisar temporalmente a sequência dos dados observados.

Para resolver o problema de sequenciamento, foi criada a técnica de visualização conhecida como gaze plot. Esta abordagem exibe o movimento do olhar de cada sujeito de maneira individual. O gaze plot mostra a sequência de movimento, ordem e duração correspondente às fixações do olhar [Barreto, 2012]. A sequência é representada por meio de uma série ordenada de círculos, onde o tamanho de cada círculo representa a duração da fixação e o seu número apresentado por um rótulo indica a ordem temporal [Ristovski et al., 2013].

Além dessas três técnicas visuais mencionadas, existem outras que visam investigar os dados oculares de formas distintas, como, por exemplo: shaded area maps; cluster; bee swarm; técnicas híbridas; entre outras [Barreto, 2012]. A Figura 2.4 ilustra alguns exemplos de relatórios visuais.

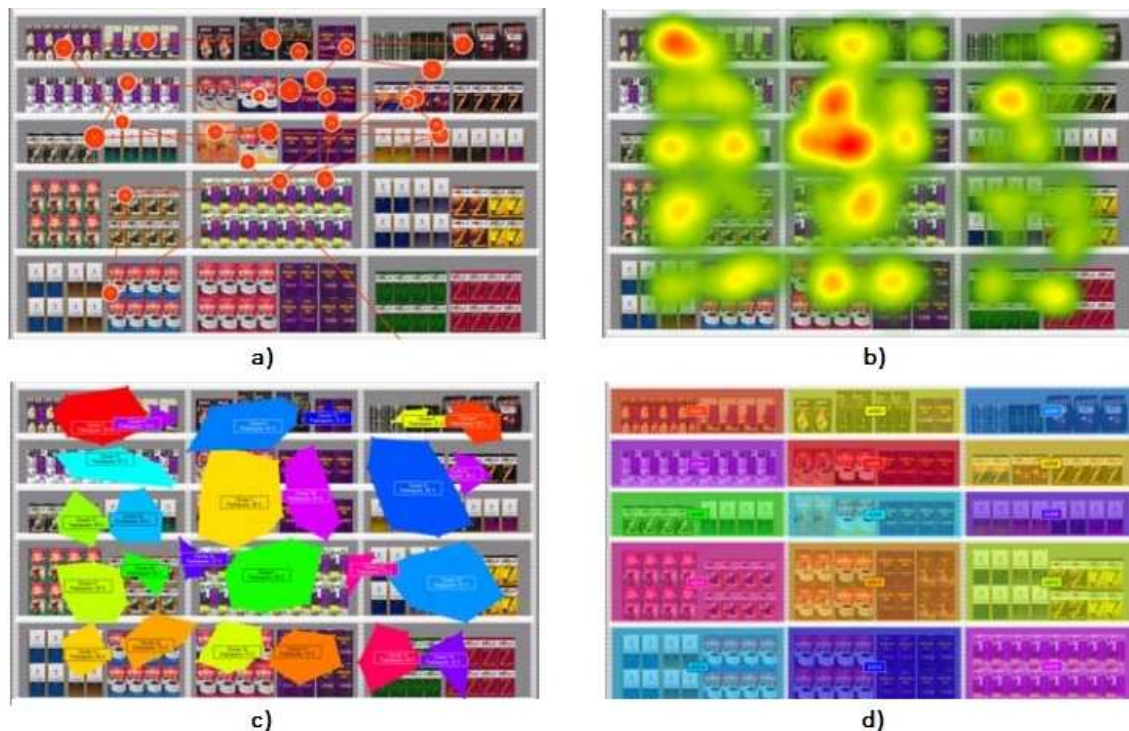


Figura 2.4. Exemplos de relatórios visuais: (a) Gaze plots, (b) Heat maps, (c) Clusters, (d) AOI. Fonte: Strandvall [2008].

2.2.2 Aplicações e trabalhos relacionados

A tecnologia eye tracking tem apresentado um grande potencial de aplicação em uma ampla variedade de áreas do conhecimento, tanto na análise quanto na interpretação dos dados oculares para a compreensão do comportamento humano [Coco, 2009; Pärnamets et al., 2015]. Portanto, vem sendo um canal promissor para os pesquisadores .

Na literatura, recentemente, vem de destacando-se com o uso do eye tracking as pesquisas nas seguintes áreas:

- Saúde mental [Galgani et al., 2009; Sasson & Elison, 2012];
- Radiologia e diagnóstico por imagem [Tourassi et al., 2013; Kübler et al., 2015];
- Investigação criminal [Mansour & Flowe, 2010];

- Ensino-aprendizagem [Yu et al., 2012; André et al., 2015];
- Esporte [Williams & Davids, 1998];
- Interação homem-computador [Granka et al., 2004; Wu & Munzner, 2015; Zhang et al., 2015];
- Avaliação de comportamento e conhecimento [Bednarik, 2007; Cardoso & Costa, 2013];
- Marketing [Ristovski et al., 2013];
- Psicologia cognitiva [Klingner et al., 2008; Hossain & Yeasin, 2014];
- Robótica [Shic & Scassellati, 2007];
- Tomadas de decisões morais e éticas [Pärnamets et al., 2015]; entre outras.

Embora as pesquisas mencionadas acima e outros presentes na literatura tenham demonstrado aplicações do rastreamento ocular, nem todas foram arquitetadas com apoio computacional. Assim, com a necessidade de analisar os dados oculares, obter informações e proporcionar aos analistas uma compreensão rápida e eficaz dos resultados obtidos nos diversos experimentos, foram desenvolvidas técnicas computacionais que visam auxiliar a investigação [Blascheck et al., 2014]. A seguir, serão apresentados alguns exemplos de trabalhos propostos para o estudo dos dados oculares por meio de abordagens computacionais de aprendizagem de máquina e visualização de dados.

2.2.2.1 Pesquisas sobre tomadas de decisões éticas

McLaren [2006] aponta os desafios, as etapas iniciais e as orientações futuras dos modelos computacionais de raciocínio ético. O autor explica as dificuldades enfrentadas pelos pesquisadores no campo da investigação ética ao tentar criar softwares para auxiliar na tomada de decisões. Truth-Teller, SIROCCO e MedEthEx são alguns programas conhecidos, desenvolvidos usando algoritmos de aprendizagem de máquina, que podem ser úteis para lidar com questões éticas. Esses sistemas computacionais foram baseados em estruturas teóricas previamente conhecidas e em raciocínio sobre casos.

Um estudo que investigou por meio de algoritmos de aprendizagem de máquina, a relação entre os dados oculares de radiologistas com a decisão diagnóstica na detecção do câncer de mama, foi publicado por Tourassi et al. [2013]. Na pesquisa,

que não enfocou aspectos éticos, os autores foram capazes de produzir e validar, por meio do uso de um software de mineração de dados, modelos preditivos altamente precisos que apontam para as hipóteses que ligam o conteúdo de uma imagem, o olhar e a cognição com o erro de diagnóstico médico. Os autores demonstraram, por meio de modelos de algoritmos de aprendizagem de máquina, a existência de uma interação entre o olhar com a tomada de decisão.

O estudo de Pärnamets et al. [2015] reforça a influência certas variáveis relacionadas ao movimento dos olhos na tomada de decisões éticas. Neste estudo, os autores contextualizaram a influência do rastreamento ocular na tomada de decisão ética, sendo o primeiro estudo a mostrar relações causais entre o olhar e as escolhas morais. No entanto, nenhuma ferramenta computacional foi desenvolvida para auxiliar na análise e verificação de dados.

2.2.2.2 Pesquisas sobre representação de dados

Mesmo que o problema de sequenciamento tenha sido resolvido com a técnica de gaze plot, a comparação e interpretação dos dados com ordenação temporal ainda é uma tarefa complexa e não trivial. Desta forma, segundo Ristovski et al. [2013], Blascheck et al. [2014] e Wu & Munzner [2015], pesquisadores têm desenvolvido diversos métodos e softwares que tentam identificar padrões de atenção visual em dados de rastreamento ocular com ordenação temporal baseados em diferentes fatores e análises, a fim de obter meios que auxiliem a responder as seguintes perguntas típicas levantadas quando se deseja estudar os movimentos dos olhos:

- quais são as áreas de interesse na cena?
- em que o observador se fixa?
- quando um observador observa uma área de interesse?
- quanto tempo ou quantas vezes um observador se fixa em cada objeto?
- em que ordem um observador observa diferentes áreas de interesse?
- existe um padrão particular na ordem das fixações em áreas de interesse?
- existe uma semelhança entre os padrões de diferentes sujeitos?
- os observadores respondem de forma diferente se a tarefa dada é diferente nas mesmas circunstâncias?

Ristovski et al. [2013] desenvolveram um sistema de visualização interativo conhecido como EyeC, que permite a análise e comparação de padrões espaço-temporais de sujeitos, ou grupos de sujeitos, sobre um número limitado de AOIs. O sistema foi composto por meio do uso de medidas de distância e pela combinação das técnicas de visualização clássicas e gráficos estatísticos. A abordagem proposta permite ao analista a realização de consultas complexas e a extração de informações para formulação de hipóteses.

Assim como Ristovski et al. [2013], Wu & Munzner [2015] propuseram um sistema de visualização que facilita a descoberta e análise de padrões sequenciais em grandes volumes de dados de rastreamento ocular. O SEQIT, assim denominado, é um sistema web open source com um design limpo, ágil e responsivo que realiza a análise interativa de padrões espaço-temporais por meio de múltiplas comparações sequenciais associadas às características dos sujeitos. O sistema agrega as fixações das AOIs visitadas e retorna o sequenciamento dos movimentos dos olhos na forma de um cronograma.

Recentemente, Muthumanickam et al. [2016] propuseram uma abordagem híbrida de visualização interativa para resolver o problema de desordem visual ocasionada pela sobreposição do grande volume de dados de rastreamento ocular. A abordagem é baseada em métodos de agrupamento e na exploração interativa orientada pelo usuário em árvores de sequências de AOIs por meio do método ActiviTree.

Na literatura, há outras técnicas e ferramentas que visam possibilitar a visualização da sequência dos dados oculares, como, por exemplo: o eyePatterns [West et al., 2006]; o eSeeTrack [Tsang et al., 2010]; o ISeeCube [Kurzahls et al., 2015]; a técnica de gráficos de tempo de Rähä et al. [2005]; entre outros. Para um estudo mais aprofundado sobre as técnicas de visualização de dados oculares existentes e suas funcionalidades, aconselha-se a leitura dos trabalhos de Andrienko et al. [2012] e Blascheck et al. [2014].

No entanto, os trabalhos citados acima se concentram no uso das métricas oculares referentes a quantidade e duração de fixações sobre as AOIs, não sobre a variabilidade do diâmetro pupilar. Nota-se na literatura, uma carência de trabalhos sobre técnicas de visualização que permitam a análise dos dados segundo o tamanho da pupila. Um exemplo deste tipo de trabalho é o de Hossain & Yeasin [2014], que propuseram um método baseado na transformação de Hilbert para calcular padrões temporais a partir de respostas pupilares.

2.3 Carga cognitiva

A carga cognitiva, também conhecida como esforço cognitivo, refere-se a quantidade total de esforço mental que está sendo usado em recursos psicológicos, como, memória, atenção, percepção, raciocínio, juízo, entre outros, durante a resolução de problemas. O estudo psicofisiológico da carga cognitiva em tarefas visuais necessita de uma métrica fisiológica que responda rapidamente a carga e reflita com confiança pequenas diferenças de carga [Klingner, 2010].

Visto que o aumento da atividade cortical ocasiona uma pequena e momentânea resposta nervosa autônoma, as técnicas que medem os efeitos secundários não-neurais dessa resposta também são usadas como métricas para a carga cognitiva. Tais efeitos fisiológicos incluem pequenas variações na frequência cardíaca, glicemia, atividade elétrica nos músculos faciais, detalhes dos movimentos oculares e, dilatações da pupila [Scher et al., 1984; Boxtel & Jessurun, 1993; Nakayama & Shimizu, 2004; Scholey et al., 2006].

Kahneman & Beatty [1966] em seus estudos, favoreceram o uso das dilatações pupilares nos métodos experimentais de estudo para medir o esforço cognitivo demandado durante a realização de tarefas visuais com diferentes níveis de dificuldades, pois:

- a pupilas possuem um rápido retorno em ms, o que permite estudar o tempo minucioso da carga cognitiva durante as tarefas visuais;
- o diâmetro da pupila pode ser medido usando dispositivos com câmeras infravermelhas remotas de alta velocidade, tornando-se uma das métricas psicofisiológicas para carga cognitiva menos invasiva;
- as medidas do diâmetro da pupila são registradas como uma métrica extra pelos rastreadores oculares quando se coleta dados sobre os movimentos dos olhos, ou seja, possibilitam o estudo de como a carga cognitiva está relacionada ao locus de atenção devido a sincronização da direção do olhar com a pupila.

Deste modo, pode-se inferir que as pupilas se dilatam rapidamente em resposta a cargas cognitivas. Portanto, podem ser usadas para melhorar a compreensão a respeito do desempenho humano na realização de tarefas visuais [Klingner, 2010].

2.3.1 Pupilometria

O estudo do comportamento da pupila, também conhecido como pupilometria, têm sido usado como indicativo de medição do processamento cognitivo durante a reali-

zação de tarefas desde o final do século XIX. Entretanto, somente nos anos 60 que essa técnica passou a ser reconhecida como uma ferramenta de avaliação dos processos cognitivos dos seres humanos [Beatty, 1982]. Estudos desenvolvidos a partir da avaliação da pupilometria indicam que o tamanho da pupila está associada de forma confiável ao esforço cognitivo necessário para a realização de um amplo conjunto de processos cognitivos, como, por exemplo: memória de curto prazo; atenção sustentada e seletiva; percepção; processamento de linguagem; raciocínio; entre outros [Hess, 1975; Beatty, 1982; Klingner et al., 2008; Hossain & Yeasin, 2014].

Durante a realização de uma determinada tarefa cognitiva, acontece involuntariamente uma alteração no tamanho da pupila [Klingner et al., 2008; Privitera et al., 2008]. Conforme é solicitado um maior grau de esforço cognitivo, as pupilas dilatam ligeiramente e, logo ao completar a tarefa, acontece uma diminuição dos estímulos e as pupilas se contraem progressivamente até voltar ao seu tamanho anterior. Essa reação é conhecida como Resposta Pupilar Evocada por Tarefa, do inglês, Task-Evoked Pupillary Responses (TEPR) [Kahneman & Beatty, 1966]. Entretanto, a medição da TEPR, segundo Klingner et al. [2008], não acontece de maneira segura para apenas um único caso de esforço cognitivo, sendo necessário, a utilização de muitos dados pupilares combinados durante um período de tempo para se obter uma medição de TEPR confiável para analisar.

A Figura 2.5 a seguir, demonstra um exemplo de TEPR. A Subfigura 2.5a ilustra a reação da pupila durante a execução de tarefas mentais de baixa e alta dificuldade. Nota-se a brusca dilatação da pupila durante a mudança da do exemplo de multiplicação. A Subfigura 2.5b, por meio de um gráfico, ilustra o efeito de dificuldade na dilatação pupilar evocada por questões de multiplicação mental apresentados seguidamente.

A dinâmica da alteração do diâmetro pupilar, ilustrada na Figura 2.6, é controlada pelas ações antagônicas dos dois componentes do Sistema Nervoso Autônomo que operam nos músculos lisos da pupila, o Sistema Nervoso Simpático (SNS) e o Sistema Nervoso Parassimpático (SNP). O SNS localizado no hipotálamo é responsável por estimular ações que permitem ao organismo reagir em situações de estresse. Por meio da liberação do neurotransmissor noradrenalina pelos neurônios noradrenérgicos, o SNS gera uma excitação simpática autônoma que estimula o músculo dilatador radial íris, sendo este o músculo responsável pelo aumento do diâmetro pupilar [Privitera et al., 2008; Pereira et al., 2012].

Já o SNP possui a função contrária ao SNS, sendo o responsável por excitar ações que inibem a estimulação simpática, permitindo que o organismo responda em situações de repouso. O SNP é mediado pelo complexo oculomotor de Edinger

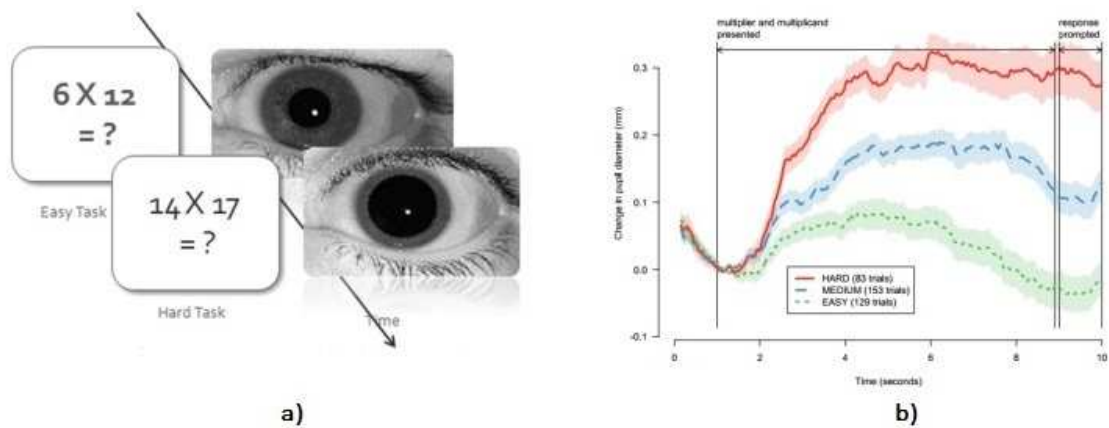


Figura 2.5. Ilustração de respostas pupilares na execução de tarefas cognitivas. Fonte: Hossain & Yeasin [2014].

Westphal no mesencéfalo e por meio da secreção do neurotransmissor acetilcolina pelos neurônios colinérgicos, inerva a esfíncter da pupila, músculo responsável pela redução do diâmetro pupilar [Smith et al., 1970; Krenz et al., 1985; Loewenfeld & Lowenstein, 1993].

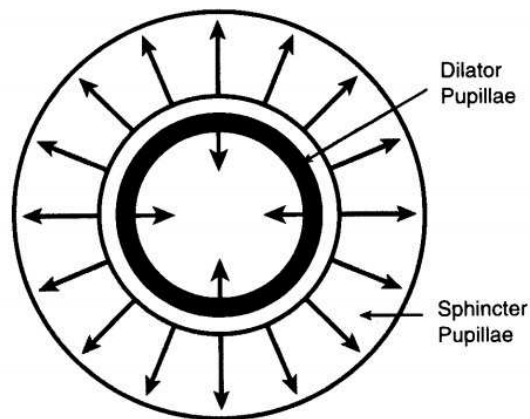


Figura 2.6. Músculos lisos responsáveis pela dilatação e constrição da pupila. Fonte: Beatty & Lucero-Wagoner [2000].

Neste capítulo, foi realizada uma revisão geral de importantes conceitos e trabalhos da literatura. Os conceitos aqui apresentados serão utilizados no decorrer desta dissertação para o estabelecimento de relações entre os aspectos visuais e cognitivos com a tomada de decisão. O próximo capítulo apresentará os algoritmos, métodos computacionais e estatísticos que serão utilizados e comparados nos experimentos das abordagens propostas.

Capítulo 3

Métodos Computacionais e Matemáticos

Para extrair informações implícitas e potencialmente úteis em bases de dados e validá-las, é necessário a utilização de técnicas sofisticadas derivadas das áreas da ciências exatas, capazes de lidar com diferentes tipos de dados para que sejam encontrados importantes padrões que possibilitam a realização de precisas observações. Neste capítulo, serão explicados os principais métodos computacionais e matemáticos utilizados, respectivamente, nas abordagens de Mineração de Dados (Capítulo 4) e Estatística (Capítulo 5) dessa pesquisa, para a descoberta de conhecimento e uma compreensão dos mesmo relacionados aos problemas abordados.

3.1 Mineração de dados

O uso generalizado de sistemas de informações e a extraordinária expansão no volume de dados das organizações decorrente da aquisição e do armazenamento de dados, principalmente nos negócios e na pesquisa científica, marcaram historicamente a era das informações [Lavrač, 1999]. Entretanto, o incrível aumento na quantidade de dados brutos geraram algumas questões desafiadoras para a análise dos mesmos, como, por exemplo: escalabilidade, dimensionalidade, qualidade, heterogeneidade, complexidade, propriedade e distribuição dos dados [Tan et al., 2006].

Tais desafios impulsionaram o desenvolvimento de técnicas e ferramentas computacionais eficientes que potencializam a análise e transformação dos dados brutos com distintas características em informações úteis [Rodrigues Júnior, 2003; Carvalho et al., 2012]. Nesse sentido, surgiu no final dos anos 80, o processo denominado como Mineração de Dados, ou do inglês, Data Mining. Este processo foi criado

como uma forma alternativa aos métodos computacionais tradicionais de análise que tornaram-se ineficazes para a manipulação e tratamento de dados por causa da sobrecarga deles em grande parte dos repositórios de dados [Tan et al., 2006; Camilo & Silva, 2009].

A Mineração de Dados é um subcampo de Ciência da Computação originado por meio da união de técnicas de Estatística, Inteligência Artificial, Aprendizagem de Máquina e Sistemas de Banco de Dados [Tan et al., 2006]. Na literatura, Mineração de Dados é comumente definido como um processo de caráter multidisciplinar, cujo o objetivo é a extração automática ou semiautomática de informações ocultas e importantes em grandes bases de dados por meio de algoritmos aprimorados que permitem a exploração, análise e modelagem dos dados [Cerqueira, 2010; Witten et al., 2016]. Na prática, geralmente a Mineração de Dados é categorizada em duas formas de análise, e a escolha da categoria em aplicações pode variar consideravelmente conforme o objetivo da pesquisa, o problema a ser estudado e suas respectivas bases de dados [Inza et al., 2010; Crockett & Eliason, 2014].

A categoria de predição tem como finalidade a construção de modelos de dados preditivos capazes de preverem, por meio de informações disponíveis, valores desconhecidos ou futuros e as classes as quais as instâncias preditas pertencem [Fayyad et al., 1996; Hollander et al., 2013; Witten et al., 2016]. A categoria descritiva, por sua vez, busca encontrar um conjunto de padrões finitos de interpretação humana, como, grupos ou clusters, correlações e anomalias semelhantes, que descrevem propriedades importantes dos dados [Tan et al., 2006; Cerqueira, 2010]. A principal diferença entre as categorias está na forma de atribuição de um objeto a classe. Enquanto a categoria preditiva utiliza classes predefinidas para verificar qual grupo um determinado objeto pertence, a categoria descritiva visa formar grupos de objetos mais homogêneos entre si sem usar classes previamente definidas [Fayyad et al., 1996; Witten et al., 2016].

Pelo fato de existirem diferentes tipos de padrões nas bases de dados, é necessário variados tipos de métodos e técnicas para encontrá-los. Logo, há distintas tarefas de Mineração de Dados nas categorias preditivas e descritivas que permitem conseguir os padrões com base nos objetivos que se almeja [Fayyad et al., 1996]. De acordo com Tan et al. [2006], as principais tarefas da categoria preditiva para prever resultados e tendências futuras são: classificação e regressão. E na categoria descritiva, as principais tarefas utilizadas para encontrar padrões e identificar características entre variáveis são: agrupamento ou clusterização, sumarização e associação.

A chegada da Computação de Alto Desempenho tem beneficiado várias áreas

de pesquisas na procura de soluções práticas para seus problemas, e a área da saúde não é uma exceção disso [Lavrač, 1999]. Segundo Acharya & Yu [2010], Hollander et al. [2013] e Crockett & Eliason [2014], a Mineração de Dados tem apresentado um grande potencial para o setor de saúde por possibilitar que os sistemas de saúde utilizem de forma sistemática a análise e exploração efetiva de dados, de modo a ajudar:

- na compreensão de forma mais clara dos dados médicos;
- na distinção de dados patológicos de dados normais;
- os clínicos a fazerem um melhor diagnóstico para fins de tratamento;
- no prognóstico de pacientes com base nas informações disponíveis em seus históricos ou sobre casos anteriores com os mesmos sintomas;
- na identificação de ineficiências e de melhores práticas que melhorem os cuidados da saúde dos pacientes;
- no apoio a tomadas de decisões;
- na redução de custos; entre outros.

A Mineração de Dados tornou-se uma metodologia fundamental para aplicações de computação em Informática Biomédica. O progresso de suas implicações na extração automática de informações por meio de vários algoritmos de aprendizagem baseados em: árvores de decisão; regras; redes neurais; clusters e séries temporais, auxiliaram significativamente diversos pesquisadores a alcançar resultados promissores e publicar pesquisas inovadoras na área [Acharya & Yu, 2010; Crockett & Eliason, 2014].

Recentemente, a IBM desenvolveu o Watson¹, uma ferramenta computacional de Big Data, baseada em tecnologia cognitiva e processamento de linguagem natural, capaz de processar dados e de aprender com eles de forma rápida e muito semelhante ao cérebro humano [Chen et al., 2016]. Na medicina, o Watson Health pretende aprimorar a qualidade e a eficiência no tratamento de doenças, funcionando como uma ferramenta de suporte à decisão médica que colabora tanto para a qualidade de vida do paciente quanto para a diminuição de custos hospitalares. Por meio do rápido cruzamento de informações, como: histórico de exames médicos reais; dados

¹É um sistema otimizado para carga de trabalho fundamentado na arquitetura de software DeepQA (arquitetura probabilística paralela maciça baseada em evidências) e executado em um cluster de servidores com processadores POWER7.

clínicos e genéticos dos pacientes; literatura científica; entre outros, o Watson analisa todas as informações e, com base em evidências, auxilia os médicos retornando todos os possíveis tratamentos recomendados para cada paciente, juntamente com as chances de cura, grau de risco e efeitos colaterais [High, 2012]. O Watson está sendo principalmente aplicado na Oncologia, na análise do genoma de tumores e no desenvolvimento de novas drogas.

Além da aprendizagem supervisionada, existem outros tipos de técnicas que possibilitam que os computadores aprendam com base na entrada fornecida, que não serão explicados neste trabalho. No entanto, para maiores detalhes a respeito das técnicas de aprendizagem existentes e Mineração de Dados, recomenda-se a leitura dos trabalhos de Fayyad et al. [1996], Tan et al. [2006], Han et al. [2011] e Witten et al. [2016].

Na subseção a seguir, o foco estará nos algoritmos de aprendizagem supervisionada da categoria de predição para a tarefa de classificação, utilizados na primeira abordagem deste trabalho. Esses algoritmos foram empregados para construir modelos preditivos usando dados oculares, com o intuito de fornecer resultados que possam apoiar o estudo da investigação médica no processo de tomada de decisão em bioética clínica.

3.1.1 Algoritmos de aprendizagem máquina

A Aprendizagem de Máquina é uma subárea da Inteligência Artificial que tem como finalidade o estudo e desenvolvimento de algoritmos que concedam a sistemas computacionais a melhoria de sua performance em determinada tarefa, por meio da habilidade de aprendizagem automatizada a partir de dados de experiências passadas [Michalski et al., 2013].

A aprendizagem supervisionada constitui-se de uma abordagem que aplica algoritmos sobre um conjunto de treinamento, formado por registros ou instâncias, já classificados, ou seja, cujos valores discretos e o rótulo da classe são conhecidos a priori, para construir um modelo que seja capaz de classificar novas instâncias em um conjunto de teste de forma precisa baseada na aprendizagem obtida com o conjunto de treinamento [Tan et al., 2006; Cerqueira, 2010]. Portanto, este modelo de classificação, também conhecido como classificador, é utilizado para prever o rótulo da classe de registros não conhecidos. A Figura 3.1 ilustra um exemplo genérico do processo de construção de um classificador.

O principal objetivo de um algoritmo de aprendizagem supervisionada é discernir um modelo com boa capacidade de generalização, isto é, que melhor se adapte

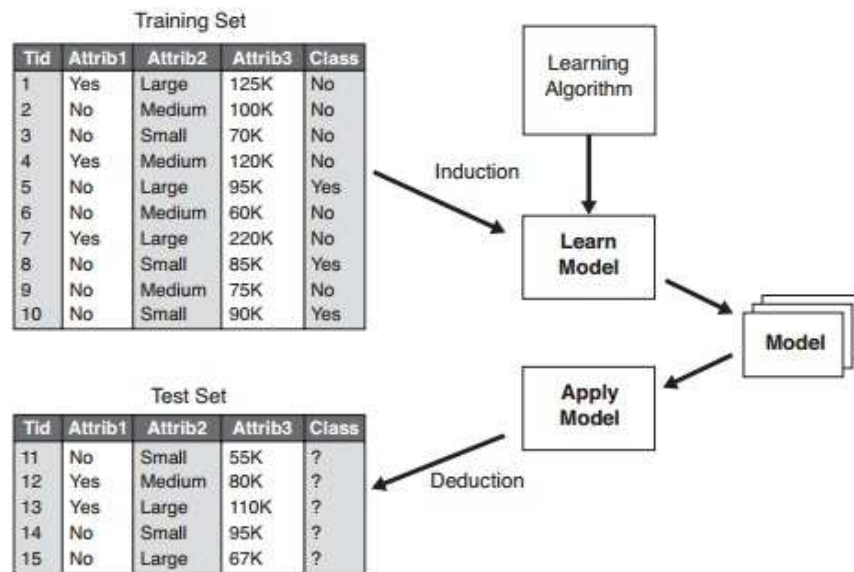


Figura 3.1. Construção de um modelo de classificação. Fonte: Tan et al. [2006].

a relação entre o vetor de atributos e o rótulo de classe de entrada, predizendo corretamente, os rótulos de classe para os novas instâncias [Tan et al., 2006].

Diversos algoritmos funcionam de formas diferentes e eficientes em certas tarefas de aprendizado, como, por exemplo, no reconhecimento de caracteres e no processamento de linguagem natural; na análise de imagens para diagnóstico médico e tomada de decisão clínica; na anotação do genoma e previsão de função ou descoberta de biomarcadores; entre outros [Geurts et al., 2009; Zhang & Rajapakse, 2009]. A seguir, será explicados os algoritmos escolhidos para realizar a classificação juntamente com o motivo pela escolha.

3.1.1.1 Multilayer perceptron

As redes neurais artificiais (RNAs) são constituídas por modelos matemáticos e computacionais inspirados na estrutura de funcionamento do sistema neural biológico, no que diz respeito a habilidade de aprender a partir de dados e melhorar o seu desempenho segundo o que foi aprendido [Tan et al., 2006]. Uma RNA é composta por um conjunto de unidades de processamento, também conhecidos como neurônios ou nós, conectados por conexões ponderadas que são responsáveis pela propagação dos valores como sinais entre os nós [Michalski et al., 2013; Cerqueira et al., 2014].

A primeira arquitetura de rede neural proposta foi chamada de perceptron, em 1953. O perceptron tinha como objetivo simular a capacidade de reconhecimento

de padrões dos seres humanos [Rosenblatt, 1957]. Esse modelo de RNA simples, ilustrado na Figura 3.2, é composto por apenas dois tipos de nós, entrada e saída. Os nós de entrada são utilizados para representar os valores dos atributos, o nó de saída representa o resultado obtido pelo modelo. Os pesos presentes nas ligações indicam a força das conexões sinápticas entre os neurônios [Tan et al., 2006; Cerqueira et al., 2014].

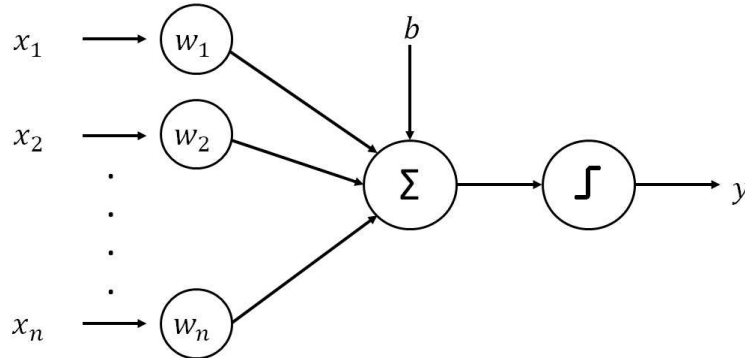


Figura 3.2. Modelo de perceptron.

O modelo de perceptron calcula o valor de saída, por meio da soma ponderada de suas entradas seguida pela subtração do termo bias (b na Figura 3.2). Deste modo, é obtida a combinação linear dos valores de entrada, cujo resultado será aplicado à função ativação, que realiza a seguinte verificação: se o valor da soma for positivo, a saída será 1. Se for negativo, a saída será -1 (a saída 0 é frequentemente utilizada em vez de -1). Para uma soma que resulta em zero, a decisão sobre a classe final depende do critério estabelecido para esses casos [Tan et al., 2006; Michalski et al., 2013; Cerqueira et al., 2014]. O modelo de perceptron representa a equação de um hiperplano dado por:

$$\mathbf{w} \cdot \mathbf{x} + b = 0, \quad (3.1)$$

onde \mathbf{w} representa o vetor de pesos, \mathbf{x} corresponde ao vetor de valores de atributo e b se refere ao termo bias [Tan et al., 2006; Cerqueira et al., 2014]. O termo bias, na forma de um peso extra constante, fornece mais flexibilidade ao hiperplano, ou seja, a linha de decisão pode se mover da origem do espaço definido pelos atributos [Braga et al., 2007; Cerqueira et al., 2014].

Os pesos são inicializados com valores aleatórios. Em seguida, eles são ajustados durante a fase de treinamento para que a saída se torne consistente com os resultados do conjunto de treinamento [Tan et al., 2006; Braga et al., 2007; Cer-

queira et al., 2014]. O processo iterativo de ajustes dos pesos sinápticos representa o processo de aprendizagem em si [Haykin, 1999]. Esse processo é expresso pela seguinte equação:

$$w_j^{(k+1)} = w_j^{(k)} + \lambda(y_i - \hat{y}_i^{(k)})x_{ij}, \quad (3.2)$$

onde $w_j^{(k)}$ representa o peso associado ao atributo j^{th} na iteração k , y_i é a classe de instância de treinamento, \hat{y}_i a classe predita, x_{ij} é o valor do atributo j^{th} da instância de treinamento x_i , e λ a taxa de aprendizado utilizada para controlar o ajuste do peso em cada iteração k [Tan et al., 2006].

Multilayer perceptron (MLP) é uma evolução do modelo de perceptron usado para resolver problemas de classificação mais complexos [Tan et al., 2006]. Neste modelo, os nós só enviam informações para camadas posteriores, não permitindo a ocorrência de ciclos ou de loops dentro da rede, realizando, portanto, uma alimentação direta [Braga et al., 2007].

As redes MLPs produzem bordas de decisão mais flexíveis e poderosas, uma vez que podem ter várias camadas e nós intermediários, também chamados de nós ocultos, entre as camadas de entrada e de saída, melhoram a eficácia da rede para extrair estatísticas mais elaboradas [Haykin, 1999]. Além das camadas e dos nós ocultos, as redes MLPs também utilizam funções de ativação aprimoradas para produzir saídas não-lineares, como, a função sigmóide e as funções tangentes hiperbólicas. A Figura 3.3 demonstra alguns exemplos das funções de ativação mais utilizadas segundo a literatura.

Para ajustar os pesos após a ocorrência do processamento, as redes MLPs aplicam um algoritmo de aprendizagem supervisionada conhecido como retropropagação ou backpropagation durante o processo de aprendizado. Este algoritmo usa a diferença observada entre o valor produzido pela rede e o valor esperado em uma determinada camada para gerar um valor de erro e utilizá-lo na atualização dos pesos da camada anterior. Portanto, os pesos são atualizados no sentido saída-entrada. Este processo se repete até o momento em que o resultado aproxima-se da saída desejada por meio da redução do erro calculado, inferindo assim, que a rede aprendeu uma determinada função de destino [Tan et al., 2006; Braga et al., 2007]. Essas características adicionais permitem que a abordagem MLP seja capaz de modelar relações mais complexas entre as variáveis de entrada e a saída [Tan et al., 2006].

A Figura 3.4 ilustra de forma simplificada, um exemplo de funcionamento da rede MLP. Os círculos presentes na camada de entrada (input layer) correspondem

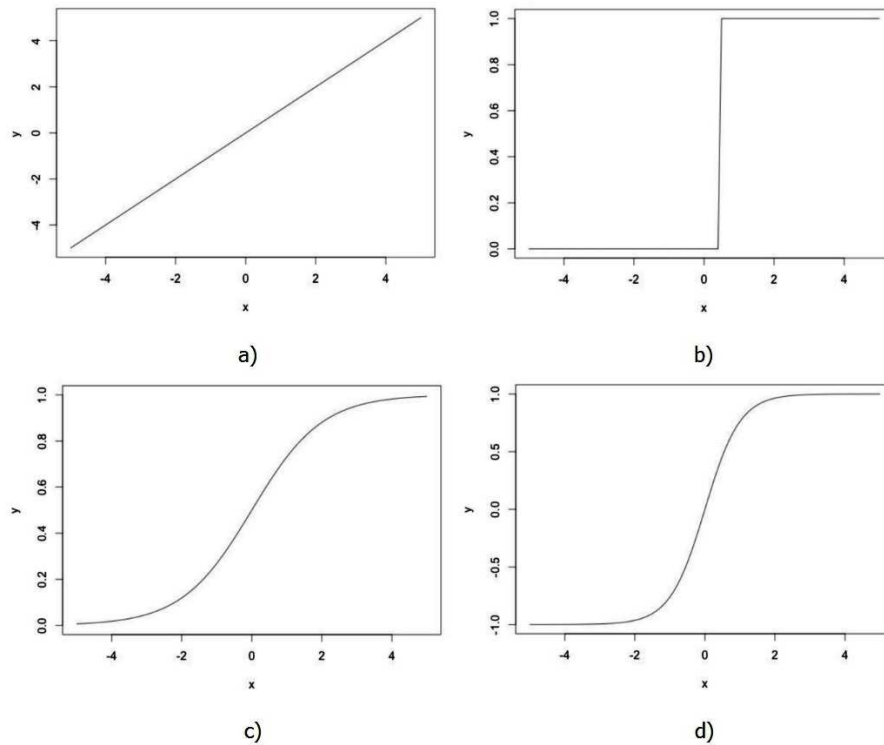


Figura 3.3. Tipos de função de ativação: (a) linear, (b) limiar, (c) sigmóide e (d) tangente hiperbólica.

aos cinco atributos utilizados no conjunto de dados. Após a camada de entrada, há uma camada oculta (hidden layer) contendo três nós ocultos, sendo estes os responsáveis por processar as informações por meio dos pesos de suas conexões. Por fim, resta a camada de saída (output layer) com apenas um nó, responsável por apresentar o resultado final alcançado pela rede. Pode assumir-se, neste caso, que o resultado final será uma combinação das saídas de três perceptrons.

Alguns exemplos recentes encontrados na literatura sobre a aplicação de MLPs em situações médicas são: na detecção de tumores do intestino delgado e doenças cerebrais [Vieira et al., 2015; Zhang et al., 2016], bem como para estimar o risco associado à cirurgia cardíaca para corrigir doenças congênitas [Ruiz-Fernández et al., 2016].

Este algoritmo de aprendizagem supervisionada foi escolhido para realizar a classificação da primeira abordagem deste estudo pelo fato de ser fortemente recomendado por pesquisadores devido o seu poder de generalização e por ter sido o algoritmo entre todos utilizados que resultou no modelo com melhor desempenho.

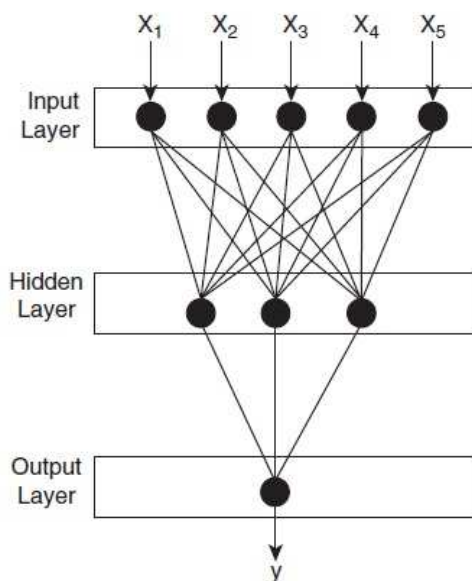


Figura 3.4. Representação simplificada do funcionamento de uma rede MLP.
Fonte: Tan et al. [2006].

3.1.1.2 Árvores de decisão C4.5

As árvores de decisão compõem um método bastante popular de classificação para a construção de modelos de predição [Rajput et al., 2011]. Neste modelo, os nós internos representam atributos a serem testados, as arestas que ligam os nós a seus nós filhos representam alternativas de valores para o atributo representado pelo nó pai, e os nós folha definem a classe (Figura 3.5) [Tan et al., 2006; Han et al., 2011].

Árvores de decisão como a C4.5, cujos nós internos representam somente um atributo, possuem limitações no tipo de borda de decisão que produzem, pois, só são capazes de realizar separação linear [Wu et al., 2008; Han et al., 2011]. Por outro lado, apresentam uma série de vantagens, dentre elas, a rapidez na construção do modelo, já que utilizam uma heurística gulosa para construção da árvore, escolhendo atributos em cada nível desta por meio de medidas de impureza, como a entropia [Tan et al., 2006; Hssina et al., 2014]. Além do mais, esses tipos de árvores de decisão possuem a vantagem de serem facilmente interpretáveis, podendo dar origem a regras que relacionam os atributos à classe [Rajput et al., 2011].

A árvore C4.5 é uma evolução de um modelo mais simples denominado ID3, onde se acrescentou a possibilidade de lidar com atributos numéricos, de tratar valores faltantes e de realizar uma poda da árvore após construí-la, de modo a diminuir sua complexidade, minimizando a possibilidade de overfitting [Quinlan, 1993; Tan et al., 2006; Rajput et al., 2011].

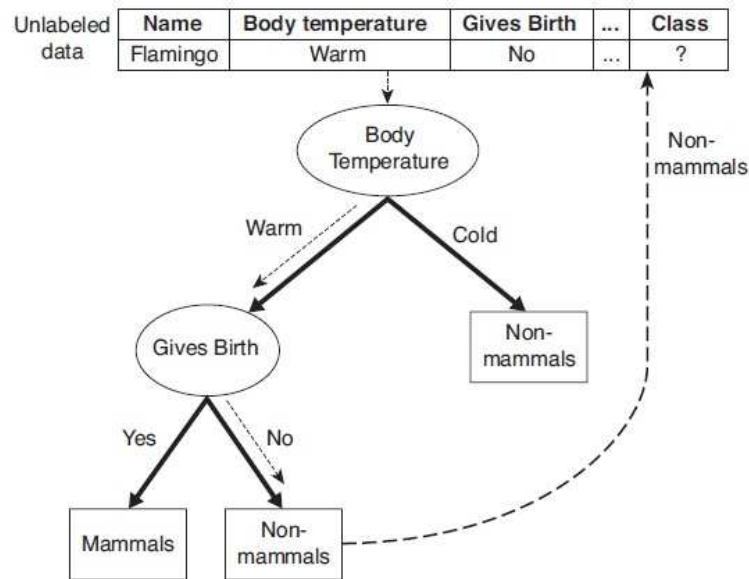


Figura 3.5. Processo de classificação usando árvore de decisão. Fonte: Tan et al. [2006].

A abordagem normalmente utilizada para a construção da árvore é a top-down, ou seja, da raiz para as folhas [Quinlan, 1993]. Neste caso, cada nó criado pelo método guloso realiza uma partição dos dados de acordo com os possíveis valores do atributo representado [Han et al., 2011]. Aplica-se então a mesma ideia, recursivamente, aos nós filhos resultantes para definir os atributos que serão testados nesses nós. Um nó para de ser expandido quando sua partição dos dados possui instâncias com a mesma classe ou quando realiza mais partições sem melhorar o poder preditivo do modelo. Esse nó, torna-se, então um nó folha e a classe atribuída ao mesmo é aquela que se encontra em maioria naquela partição [Quinlan, 1993]. Após a criação da árvore, classificar um novo exemplo torna-se trivial, pois, basta avaliar cada um dos seus atributos, percorrendo um caminho na árvore que vai da sua raiz a uma das folhas, como, mostrado na Figura 3.5.

A investigação na primeira abordagem desta dissertação utilizou a árvore de decisão para a extração de regras que pudessem trazer à luz alguns processos cognitivos relacionados à visão que fossem acionados no processo de tomada de decisão bioética do presente estudo.

Pode-se encontrar na PubMed algumas publicações de trabalhos recentes que usam a árvore de decisão C4.5 em sistemas médicos para auxílio em diagnósticos, como, por exemplo, de arritmia cardíaca [Alickovic & Subasi, 2016], da doença de Parkinson [Scherfler et al., 2016] e da doença arterial coronariana [Verma et al.,

2016].

3.2 Processo de extração de conhecimento

Graças a importância da Mineração de Dados no processo de descoberta de conhecimento, alguns autores da literatura consideram o seu conceito sinônimo do processo de Extração de Conhecimento em Bases de Dados, do inglês, Knowledge-Discovery in Databases (KDD). No entanto, esses processos, embora tenham relações, são divergentes [Fayyad et al., 1996; Witten et al., 2016]. O KDD é definido como um processo complexo, interativo, iterativo e não trivial orientado à informação, cujo objetivo é a identificação de novos padrões, modelos ou estruturas válidas, úteis e compreensíveis a partir de dados. Já a Mineração de Dados, é uma etapa específica deste processo maior de descoberta de conhecimento, responsável pela análise de dados e extração de padrões, como pode ser mostrado pela Figura 3.6 abaixo.

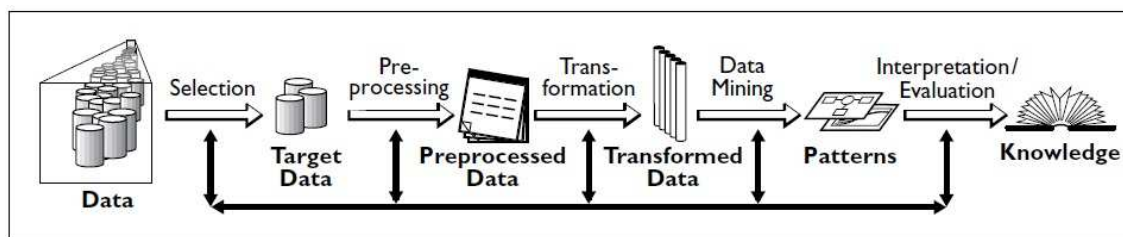


Figura 3.6. Visão geral das etapas que constituem o processo KDD. Fonte: Fayyad et al. [1996].

O processo de KDD é segmentado em uma série de etapas que podem ter seus resultados consideravelmente melhorados por meio da inspeção e das tomadas de decisões a serem feitas pelo analista. As etapas, conforme a Figura 3.6, são: seleção dos dados, pré-processamento, transformação, mineração dos dados, avaliação e interpretação das informações obtidas. É importante constatar que todo o processo de KDD apoia e incentiva o feedback. Desta forma, a qualquer momento o analista poderá voltar em uma ou mais etapas, com a finalidade de tentar alcançar melhores resultados [Witten et al., 2016].

Neste estudo, o processo de KDD foi dividido em três fases que englobam as etapas do processo, sendo elas: pré-processamento, processamento e pós-processamento dos dados. Na fase de pré-processamento estão contidas as etapas de seleção, pré-processamento e transformação de dados. Essas etapas são responsáveis por preparar por meio de operações, o conjunto de treinamento para a principal

atividade de extração de conhecimento, a mineração. Nestas etapas ocorrem as seguintes tarefas: entrada dos dados brutos; seleção dos dados relevantes; manipulação dos dados faltantes, remoção de possíveis ruídos e correção de valores inconsistentes; seleção do conjunto de atributos e redução da complexidade; transformação dos dados em formatos adequados para as análises posteriores.

A fase de pré-processamento é importante para o processo de descoberta de conhecimento uma vez que as características redundantes, ruídos ou qualquer outra inconsistência nos dados podem influenciar a detecção de informações relevantes. Portanto, a execução das etapas anteriores corretamente pode melhorar significativamente o desempenho da mineração de dados [Witten et al., 2016].

A próxima fase é a de processamento de dados, isto é, a mineração em si. Como foi mencionado anteriormente na Seção 3.1, nesta etapa do processo de KDD, ocorre a escolha dos algoritmos de aprendizagem que serão executados sobre o conjunto de treinamento para analisar os dados com o objetivo de encontrar padrões e informações úteis [Tan et al., 2006; Han et al., 2011]. Esta é a etapa mais importante e, geralmente, duradoura de todo o processo de KDD.

Por fim, a fase de pós-processamento, é composta pelas etapas de Avaliação e Interpretação, e tem como objetivo garantir que apenas resultados válidos e interessantes derivados da etapa de mineração de dados sejam ditos como conhecimentos e usados para a tomada de decisão no problema estudado [Fayyad et al., 1996; Tan et al., 2006]. Na etapa de avaliação, a performance dos modelos obtidos são validados por meio de métodos e métricas estatísticas. A etapa de interpretação já é mais dependente da interação do analista do que as demais etapas, por isso, nesta etapa, o analista verifica os padrões encontrados e decide se os resultados almejados foram alcançados e se deve realizar análises mais aprofundadas ou não [Witten et al., 2016]. Após todo o processo de KDD, é possível tomar ações com base no conhecimento adquirido.

Neste trabalho, as etapas de seleção, pré-processamento e interpretação serão explicados em suas respectivas seções no capítulo de ambas as abordagens investigadas, pois é algo comum entre as duas. Os métodos utilizado na etapa de mineração de dados já foram descritas na seção anterior. Assim, nas subseções a seguir, serão descritos os métodos utilizados nas as etapas de transformação e avaliação dos dados da primeira abordagem.

3.2.1 Transformação de dados

A transformação de dados é a última etapa da fase de pré-processamento que antecede a mineração de dados em si no processo de KDD. A transformação é responsável pela realização de algumas modificações nos dados, porém, essas alterações podem afetar bruscamente a performance da aprendizagem. Portanto, deve-se escolher os métodos mais adequados que preservem a qualidade da informação e otimizem a precisão dos algoritmos de aprendizagem que serão usados no conjunto de treinamento. Os métodos apresentados a seguir foram utilizados com intuito de obter melhores exatidões dos algoritmos.

3.2.1.1 Seleção de atributos

A seleção desses atributos é uma parte importante na fase de pré-processamento do KDD, feito para simplificar o modelo sem perder informações importantes [Blum & Langley, 1997; Yang & Pedersen, 1997]. A seleção reduz a complexidade da base de dados, o que torna a construção do modelo mais rápida e, em muitos casos, produz melhores resultados, uma vez que evita a chamada maldição de dimensionalidade que torna difícil distinguir as instâncias entre si [Santoro, 2005].

A seleção computacional dos atributos pode ser dividida em dois tipos: filtro e wrapper [Cerqueira, 2010]. Os métodos do tipo filtro são aplicados antes do processo de aprendizagem e definem a relevância de cada atributo ou de um conjunto deles para classificação de acordo com algum critério predefinido, como, por exemplo, a entropia [Santoro, 2005]. O método de filtro é recomendado para lidar com grandes conjuntos de dados, porque, como os filtros não dependem do classificador, suas análises levam pouco tempo [Cerqueira, 2010; Villacampa, 2015].

Os filtros se dividem em duas categorias: rankers e aqueles que avaliam subconjuntos. Os rankers calculam uma pontuação para cada atributo, individualmente, gerando uma classificação de atributos no final como uma lista em ordem decrescente sobre o valor da pontuação [Cerqueira, 2010]. Essa abordagem é eficiente em termos do tempo de execução. Entretanto, ignora a existência de correlações importantes entre os atributos [Villacampa, 2015].

A técnica InfoGain (IG) é um exemplo bem conhecido deste método de filtragem. O IG avalia o valor de um atributo medindo a quantidade de informação que fornece para a classe [Hall et al., 2009]. O IG de um atributo a_i em relação a uma classe c é definido pela função:

$$IG(a_i) = H(c) - H(c|a_i), \quad (3.3)$$

onde $H(c)$ representa a entropia da classe c , e $H(c|a_i)$ a entropia da classe condicionada ao atributo a_i . A entropia da classe é definida como:

$$H(c) = - \sum_{i=1}^n p_i \log_2 p_i, \quad (3.4)$$

onde n representa o número de classes distintas e p_i a proporção de instâncias que pertencem à classe i [Tan et al., 2006].

Ao contrário da categoria ranker mencionada acima, a categoria que avalia subconjuntos leva em consideração as relações entre os atributos em vez de uma avaliação individual [Cerqueira, 2010]. Um exemplo desta abordagem é o Correlation-based Feature Selection (CFS).

O CFS seleciona atributos com base na correlação entre si e com a classe. O CFS visa selecionar subconjuntos cujos atributos estão altamente correlacionados com a classe e minimamente correlacionados entre eles [Hall, 1998; Tan et al., 2006; Villacampa, 2015]. A pontuação de um subconjunto s contendo k atributos é dada pela seguinte equação:

$$Merit_s = \frac{k \overline{r_{cf}}}{\sqrt{k + k(k-1) \overline{r_{ff}}}}, \quad (3.5)$$

onde $\overline{r_{cf}}$ representa a correlação média dos atributos com a classe e $\overline{r_{ff}}$ a correlação média dos atributos entre eles [Tan et al., 2006].

Ao contrário dos métodos do tipo filtro, o método do tipo wrapper usa um algoritmo de classificação, o mesmo que será mais tarde empregado no processo de aprendizagem, para avaliar a qualidade de cada subconjunto de atributos [Kohavi & John, 1997].

Este método consiste em uma abordagem computacionalmente exigente usada para explorar, através de heurísticas, o espaço de busca dos subconjuntos de atributos que serão avaliados posteriormente por algoritmos de classificação. A dependência desses algoritmos geralmente produz maior precisão de classificação durante o processo de aprendizagem em comparação com os métodos do tipo filtro. O método wrapper é limitado pelo custo computacional mais alto à medida que o número de atributos aumenta, uma vez que o algoritmo de classificação para avaliar vários subconjuntos pode ser inviável [Santoro, 2005; Cerqueira, 2010; Villacampa, 2015].

Os três métodos de seleção de atributos mencionados foram utilizados na primeira abordagem, entretanto, o método do tipo wrapper foi o que retornou um subconjunto de atributos que possibilitou a geração do modelo com o melhor poder preditivo encontrado.

3.2.1.2 Balanceamento de dados

No mundo real, é comum nos problemas de classificação, a desigualdade entre o número de instâncias de classes nos conjuntos de treinamento. Todavia, as instâncias desequilibradas, frequentemente, fazem com que os algoritmos de aprendizagem de máquina tenham uma performance ruim na classe minoritária [Cerqueira, 2010]. Neste contexto, existem algumas técnicas que permitem lidar com essa situação.

A Classificação Sensível ao Custo (CSC) é amplamente utilizada como forma de lidar com conjuntos desbalanceados em que o modelo gerado pode ter uma tendência para a classe majoritária [Tan et al., 2006; Han et al., 2011; Cerqueira et al., 2014]. Uma vez que a classe positiva é muitas vezes a minoria, as instâncias dessa classe serão muitas das vezes avaliadas erroneamente. No entanto, essas são geralmente as instâncias mais relevantes.

Na CSC, diferentes custos podem ser determinados para a classificação errada das instâncias de cada classe [Ling & Sheng, 2008]. Por exemplo, se houver duas vezes mais instâncias da classe A em comparação ao número de instâncias da classe B , pode-se definir que o custo de classificar incorretamente as instâncias da classe B será o dobro do custo de classificar incorretamente as instâncias da classe A , de modo que um modelo equilibrado seja estabelecido para ambas as classes, mesmo com a diferença numérica [Tan et al., 2006].

3.2.2 Avaliação dos modelos de dados

Uma vez que os modelos de predição são gerados por meio dos algoritmos de aprendizagem sobre o conjunto de treinamento, torna-se necessário a utilização de métodos e métricas de avaliação para medir o desempenho dos modelos, ou seja, busca-se estimar o quão precisos e robustos são os modelos na prática [Kohavi & John, 1997; Tan et al., 2006]. A seguir, serão mencionados os métodos de avaliação utilizados na fase de pós-processamento para analisar a performance dos modelos criados.

3.2.2.1 Validação cruzada

Neste estudo, o método estatístico de validação cruzada em k -partes, do inglês, k -folds cross-validation, foi utilizado para avaliar os modelos propostos. A validação cruzada é amplamente utilizada para a avaliação ou comparação da capacidade de generalização de um modelo [Kohavi & John, 1997].

Neste método, um conjunto de treinamento original é particionado em k em partes diferentes de tamanhos iguais, onde uma partição é retida como subconjunto

de teste as outras $k - 1$ partições são utilizadas para treinamento do modelo. Este processo é repetido k vezes, de modo que, em cada iteração k , uma partição diferente seja escolhida para avaliação. Ao final do processo, o desempenho final (acurácia) relatado será o valor médio das métricas utilizadas nas k execuções, o que nos dá uma ideia de como o modelo generaliza [Tan et al., 2006; Cerqueira et al., 2014].

Foi utilizado k igual a dez ($k = 10$), que é um valor largamente utilizado na literatura, devido: experimentos terem confirmado que esta é a melhor escolha para se obter uma estimativa precisa; por ser considerado suficiente para uma avaliação imparcial e por não ser muito grande, o que poderia retardar os experimentos [Kohavi & John, 1997; Cerqueira et al., 2014]. A Figura 3.7 ilustra o funcionamento da validação cruzada 10-partes.

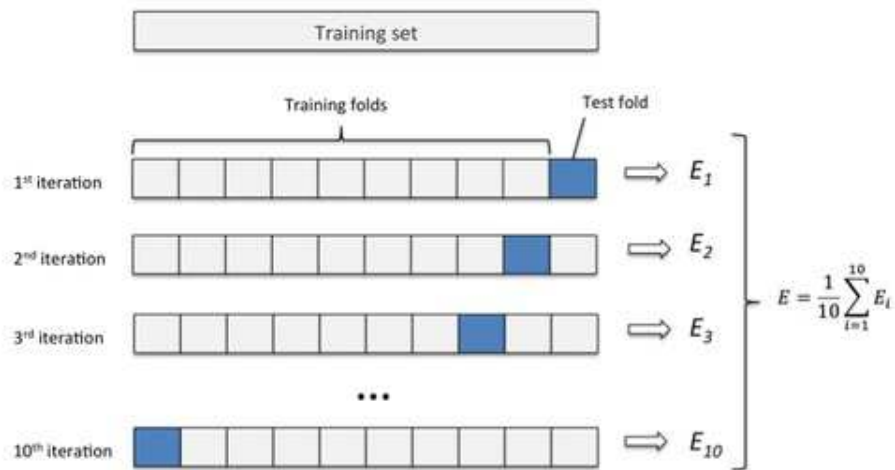


Figura 3.7. Visão geral da validação cruzada 10-partes. Fonte: Raschka [2015].

3.2.2.2 Matriz de confusão e métricas de desempenho

Uma forma adequada de avaliar o desempenho de um modelo é por meio do método visual de matriz de confusão. Essa matriz é uma tabela composta por linhas e colunas que representam, respectivamente, as instâncias da classe real e as instâncias da classe predita (Figura 3.8). Na matriz, são tabuladas as contagens de instâncias de teste preditas correta e incorretamente pelo modelo [Tan et al., 2006]. Seu nome é devido ao fato de que a matriz simplifica o processo de verificar se o algoritmo de aprendizagem está confundindo as classes [Raschka, 2015].

Na matriz de confusão, os registros previstos corretamente, denominados como verdadeiros positivos (TP - True Positives) e verdadeiros negativos (TN - True Ne-

		Predicted Class	
		Class = 1	Class = 0
Actual Class	Class = 1	f_{11}	f_{10}
	Class = 0	f_{01}	f_{00}

Figura 3.8. Matriz de confusão para um problema de classificação binária. Fonte: Tan et al. [2006].

gativos), estão localizados na diagonal principal da tabela e indicam o número de instâncias previstas corretamente nas classes 1 e 0, respectivamente. Em contrapartida, os registros previstos erroneamente, conhecidos como falsos positivos (FP - False Positives) e falsos negativos (FN - False Negatives), são representados por valores que estão fora da diagonal principal e denotam o número de instâncias previstas equivocadamente como 0 e 1 [Tan et al., 2006].

No exemplo da Figura 3.8, cada célula f_{ij} na tabela indica a quantidade de instâncias da classe i que se prevê para ser da classe j . O f_{10} , tendo como exemplo, é a quantidade de instâncias da classe 1 preditas incorretamente como classe 0. Deste modo, conforme as células na matriz de confusão, a quantidade máxima de predições exatas feitas pelo modelo é $f_{11} + f_{00}$, enquanto a quantidade total de erros é $f_{01} + f_{10}$.

Ainda que uma matriz de confusão proveja as informações necessárias para determinar a performance de um modelo de classificação, sumarizar esta informação com um único número tornaria mais conveniente para a comparação do desempenho entre diferentes modelos. Portanto, para relatar o desempenho da classificação, pode-se usar algumas das métricas clássicas que expressam o poder preditivo de um modelo [Tan et al., 2006; Raschka, 2015]. As métricas comumente utilizadas em pesquisas, segundo Kohavi & Provost [1998], Tan et al. [2006] e Raschka [2015], são:

- Acurácia: é a proporção das predições corretas pelo modelo, ou seja, o total de verdadeiros positivos e negativos, em relação ao total de instâncias do conjunto. A acurácia pode ser definida pela seguinte função:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{f_{11} + f_{00}}{f_{11} + f_{00} + f_{10} + f_{01}}; \quad (3.6)$$

- Taxa de Erro: é a métrica antagônica a acurácia, que representa a razão entre as instâncias preditas erroneamente sobre o conjunto utilizado. A taxa de erro é dada pela equação a seguir:

$$error\ rate = \frac{FP + FN}{TP + TN + FP + FN} = \frac{f_{10} + f_{01}}{f_{11} + f_{00} + f_{10} + f_{01}}; \quad (3.7)$$

- Precisão: indica a proporção de instâncias positivas previstas que são corretamente positivas reais. Desta forma, quanto maior a precisão, menor será o número de erros falsos positivos cometidos pelo classificador. A precisão é representada como:

$$p = \frac{TP}{TP + FP} = \frac{f_{11}}{f_{11} + f_{01}}; \quad (3.8)$$

- Sensibilidade ou Taxa de Verdadeiros Positivos (TPR - True Positive Rate): é a fração de instâncias positivas classificadas corretamente sobre o total de instâncias positivas. A sensibilidade é calculada da seguinte maneira:

$$TRP = \frac{TP}{TP + FN} = \frac{f_{11}}{f_{11} + f_{10}}; \quad (3.9)$$

- e Especificidade ou Taxa de Verdadeiros Negativos (TNR - True Negative Rate): de forma similar a sensibilidade, é a fração de instâncias negativas classificadas corretamente pelo modelo. A especificidade pode ser obtida pela fórmula abaixo:

$$TNR = \frac{TN}{TN + FP} = \frac{f_{00}}{f_{00} + f_{01}}. \quad (3.10)$$

3.2.2.3 Área sob a curva ROC

A curva ROC, do inglês, Receiver Operating Characteristic curve, consiste em uma técnica utilizada em vários ramos de pesquisas biomédicas e de aprendizagem de máquina para medir, analisar e comparar o desempenho entre diferentes classificadores [Metz, 1978; Fawcett, 2006]. A curva ROC representa a relação entre a sensibilidade (eixo y) e a especificidade (eixo x) de um modelo (Figura 3.9) [Tan et al., 2006].

A área sob o curva ROC, do inglês, Area Under the Curve (AUC), fornece um resumo do desempenho dos classificadores com base na distribuição dos resultados [Metz, 1978]. Um bom modelo de classificação é o que possui uma curva mais aberta, isto é, mais distante da linha diagonal e mais próximo de 1 [Tan et al., 2006]. Desta forma, quanto maior for a área sob a curva, mais alto será o poder de previsão do modelo.

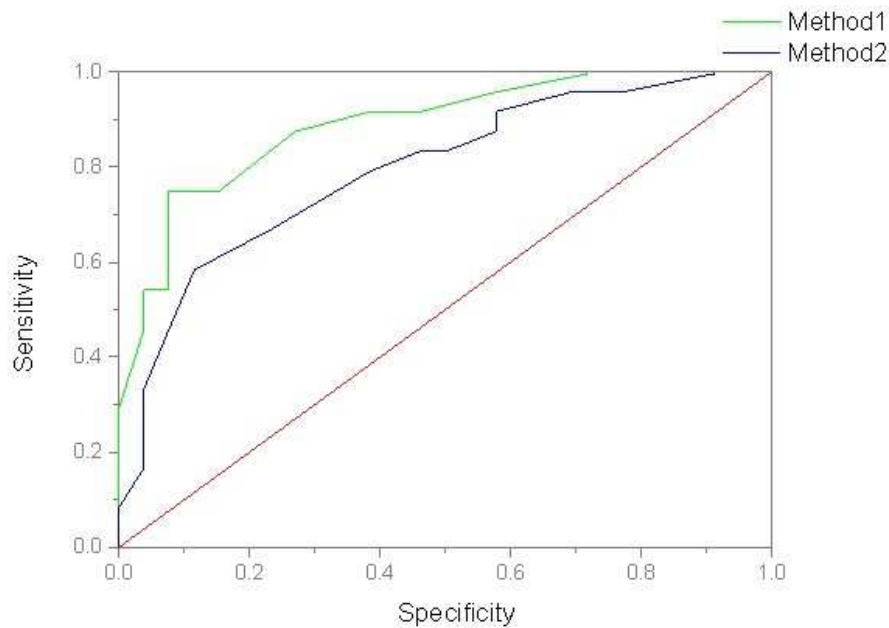


Figura 3.9. Comparação do desempenho de dois métodos por meio da área sob o curva ROC.

No exemplo da Figura 3.9 é possível notar que a curva do método 1 (linha verde) possui uma maior área em relação ao método 2 (linha azul), sendo, portanto, o modelo que apresenta melhor performance de previsão entre os dois modelos comparados. A linha diagonal (linha bordô) corresponde a uma classificação aleatória com a saída positiva ou negativa.

3.3 Análise estatística

A segunda abordagem investigada nesta dissertação difere da primeira, principalmente, na questão de análise e validação dos dados observados. Enquanto a primeira abordagem utiliza técnicas voltadas para mineração de dados, a segunda foca no uso de testes de estatísticos. Portanto, as subseções a seguir, descrevem os principais métodos utilizados na segunda abordagem deste estudo.

3.3.1 Testes de hipótese estatística

A forma científica mais comum para realizar a inferência estatística é por meio do teste de hipóteses, também conhecido como teste de significância [Marusteri & Bacarea, 2010; Nayak & Hazra, 2011]. O teste de hipóteses constitui-se de um procedimento de decisão, no qual possibilita decidir se uma determinada afirmação,

também conhecida como hipótese nula (H_0), sobre um determinado parâmetro é, ou não, sustentada pela evidência adquirida por meio dos dados amostrais [Box et al., 2005].

Portanto, para lidar com os níveis de incerteza, é utilizado nos testes estatísticos o nível de significância (α), sendo este, o responsável por julgar se os resultados dos testes são estatisticamente significativos ou não [Box et al., 2005; Marusteri & Bacarea, 2010]. Geralmente, o nível de significância utilizado é 0,05 (5%), mas, há registros na literatura de trabalhos que também utilizam 0,01 (1%). Tais níveis de significância correspondem a um nível de confiança de 95% e 99%, respectivamente.

De acordo com Marusteri & Bacarea [2010] e Nayak & Hazra [2011], a seleção do teste de hipótese estatística mais apropriado para um experimento é uma tarefa árdua e desafiadora que leva em conta o tipo de pesquisa além de algumas considerações fundamentais, como, por exemplo:

- tipo de dados (quantitativos/numéricos ou qualitativos/ordinais);
- número de grupos/amostras utilizados (um, dois, três ou mais);
- se os grupos são pareados/dependentes ou não pareados/independentes;
- caso os dados sejam numéricos, deve-se verificar se eles seguem os parâmetros da curva de distribuição normal (gaussiana);
- caso a distribuição dos dados seja normal, deve-se verificar se as amostras possuem a mesma variância.

Após a análise das considerações acima e suas definições, torna-se possível a escolha do melhor teste estatístico para um experimento em questão. A Figura 3.10 a seguir, baseada nos esquemas descritos no trabalho de Nayak & Hazra [2011], demonstra, por meio de um fluxograma, o processo de seleção dos testes estatísticos conforme a análise das considerações fundamentais.

Por causa da interdisciplinaridade deste estudo, busca-se uma melhor compreensão a respeito das considerações fundamentais dos testes. Visto isso, as próximas subseções descrevem, resumidamente, como descobrir se os grupos utilizados são ou não pareados e a importância dos testes de normalidade, homogeneidade, paramétrico e não-paramétrico.

3.3.1.1 Grupos pareados e não pareados

Ter ciência se os grupos são pareados ou não é fundamental para se definir qual versão do teste estatístico usar, pois, existem muitos testes que possuem versões

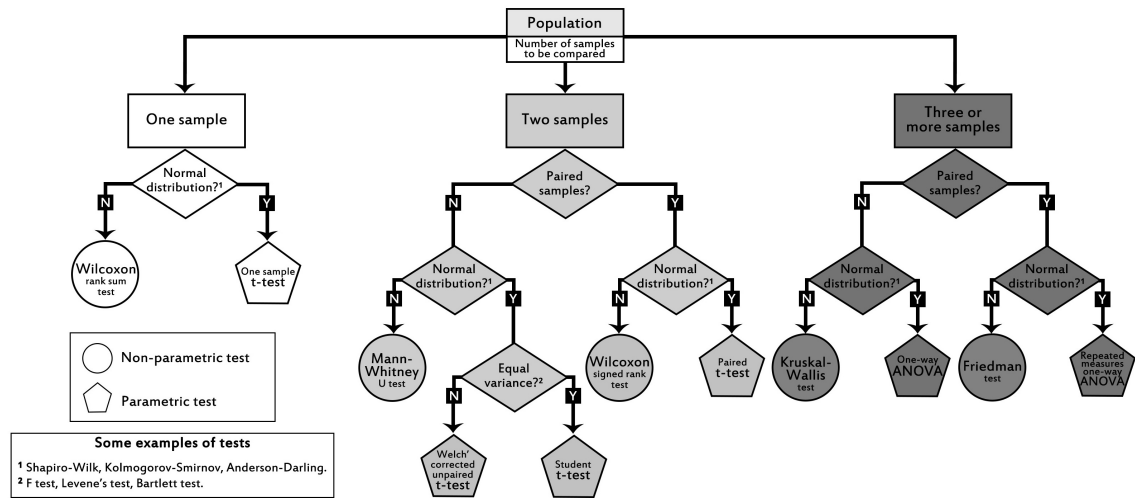


Figura 3.10. Fluxograma para escolher o teste estatístico.

diferentes para grupos pareados e não pareados. Portanto, a escolha errada do teste pode acarretar em falsas conclusões para o processo de inferência estatística [Marusteri & Bacarea, 2010]. Os grupos são ditos como pareados se houver possibilidade de:

- os dados serem derivados por medições repetidas (antes e depois), no caso de ser um grupo;
- os valores presentes no conjunto de dados utilizado estarem influenciados ou relacionados de alguma forma pelos valores do outro conjunto, no caso de serem dois ou mais grupos.

Caso contrário, os grupos são considerados como não pareados ou independentes [Marusteri & Bacarea, 2010; Nayak & Hazra, 2011].

3.3.1.2 Testes de normalidade

Os testes de normalidade são cruciais para se determinar qual tipo de teste de inferência estatística usar em uma pesquisa, devido a suposição de normalidade ser requerida em várias situações [Öztuna et al., 2006]. Esses testes, também conhecidos como testes aderência, são responsáveis pela verificação do formato da distribuição de uma variável aleatória contínua, por meio da tentativa de representação do comportamento de determinado intervalo de valores em função da frequência de sua ocorrência [Torman et al., 2012]. Em outras palavras, o teste de normalidade

calcula a probabilidade de um conjunto de dados observados ser extraído de uma população normal.

Existem diversos tipos de métodos que permitem verificar se a distribuição dos dados que estão sendo analisados se ajusta a uma distribuição normal ou não. Entre os principais tipos, há os gráficos e os testes estatísticos [Öztuna et al., 2006; Ghasemi & Zahediasl, 2012; Torman et al., 2012]. Como exemplo dos métodos gráficos, pode-se citar os: histogramas; Q-Q Plots; P-P Plots; stem-and-leaf plot; etc. A Figura 3.11 exemplifica dois exemplos de verificação da normalidade por meio de métodos gráficos.

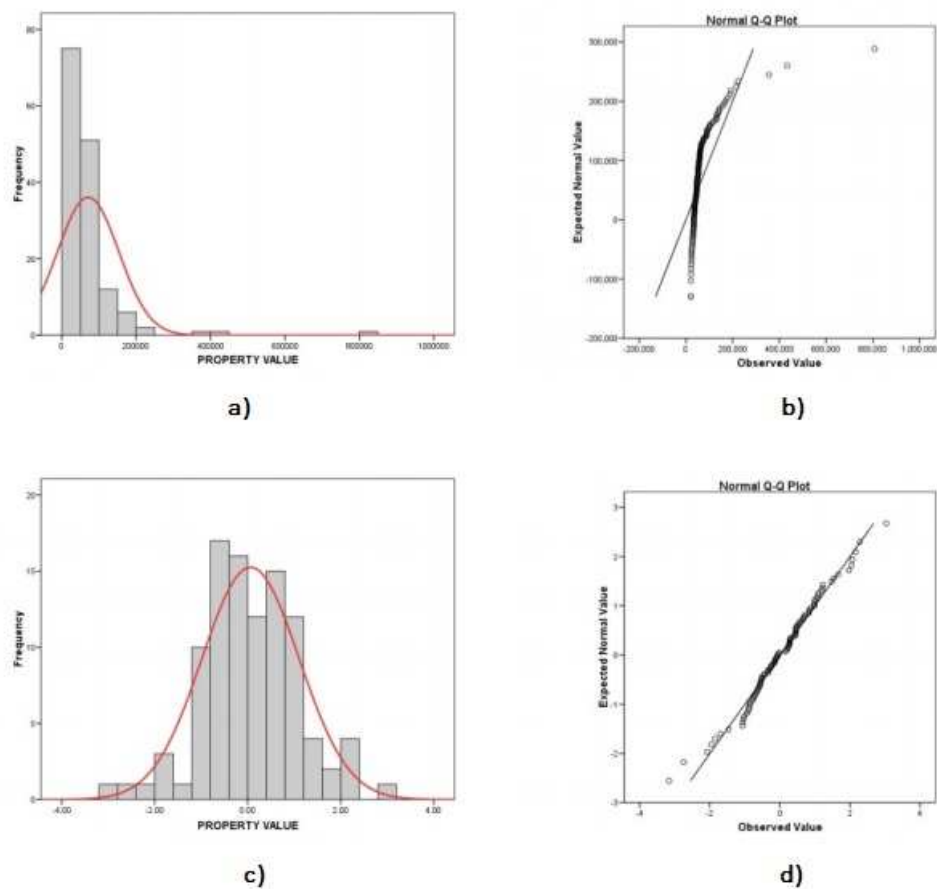


Figura 3.11. Verificação da normalidade por meio do histograma e Q-Q plots. As subfiguras (a) e (b) demonstram o comportamento de dados que não seguem uma distribuição normal, enquanto as subfiguras (c) e (d) apresentam o formato de sino e a linearidade dos pontos, sugerindo, portanto, que os dados são normalmente distribuídos.

No entanto, mesmo que os métodos gráficos sejam intuitivos, também são subjetivos, por exemplo, o que para um analista pode ser um sino ou uma linearidade, para outro pode ser algo diferente. Portanto, existe a necessidade da utilização dos tes-

tes estatísticos, como, os testes de: Kolmogorov-Smirnov, Lilliefors, Shapiro-Wilk, Anderson-Darling, D'Agostino, entre outros, para ratificar ou negar o resultado dos métodos gráficos.

As hipóteses utilizadas neste tipo de teste são:

- H_0 : os dados do conjunto provêm de uma distribuição normal;
- H_1 : os dados do conjunto derivam de uma distribuição não normal.

Após a execução do teste de normalidade, se o valor retornado pelo teste for maior ou igual à 0,05 (p-valor $\geq 0,05$), a H_0 é dita como verdadeira, caso contrário, é rejeitada.

A escolha de um teste deve ser feita examinando as características que o analista julgar adequadas para a pesquisa, tais como: poder do teste; controle da taxa do erro I e o tamanho do conjunto de dados utilizado [Nayak & Hazra, 2011]. Em virtude disso, de acordo com a literatura, o teste de Shapiro-Wilk tem sido o teste de normalidade mais recomendado por apresentar ser mais poderoso na resolução de problemas de normalidade para diferentes tamanhos de amostras e situações [Ghasemi & Zahediasl, 2012].

3.3.1.3 Testes de homogeneidade

Outro pré-requisito que necessita ser verificado quando está interessado na realização de inferências estatísticas, principalmente na paramétrica, é a análise de igualdade de variâncias. Quando esta pressuposição não é analisada, pode comprometer a validade dos resultados dos testes, impactando diretamente nas conclusões alcançadas pela pesquisa [Nogueira & Pereira, 2013; Riboldi et al., 2014].

Os testes de homogeneidade ou homocedasticidade são responsáveis por verificar se a variância das variáveis em estudo são iguais nos dois grupos de dados comparados [Nayak & Hazra, 2011]. As hipóteses utilizadas neste tipo de teste são:

- H_0 : os conjuntos de dados apresentam a mesma variância (homogêneas);
- H_1 : os conjuntos de dados apresentam diferentes variâncias (heterogêneas).

Após a execução do teste de homogeneidade, se o valor retornado pelo teste for maior ou igual à 0,05 (p-valor $\geq 0,05$), a H_0 é dita como verdadeira, caso contrário, é rejeitada.

Os principais testes encontrados na literatura para verificar a hipótese de homogeneidade de variâncias são, os testes de: Bartlett; Levene; Brown-Forsythe;

Fligner-Killeen; O'Brien e o teste F. Entretanto, a escolha do teste varia, assim como no teste de normalidade, conforme as características do conjunto utilizado. Segundo Milliken & Johnson [2009] e Riboldi et al. [2014], é recomendável usar o teste de de Bartlett se os dados são aproximadamente normais, caso contrário, sugere-se o uso do teste de Levene.

3.3.1.4 Testes paramétricos e não-paramétricos

Os testes de hipóteses são divididos em duas abordagens estatísticas, a paramétrica e a não-paramétrica. A distinção entre as duas abordagens é necessária, pois, sem um conhecimento adequado sobre ambos procedimentos, resultados enganosos podem ser gerados [Marusteri & Bacarea, 2010].

A normalidade das distribuições é a principal diferença entre as abordagens [Öztuna et al., 2006; Marusteri & Bacarea, 2010]. Os testes paramétricos são aqueles que se baseiam nas suposições de normalidade dos dados e também, às vezes, na de homogeneidade das variâncias. A abordagem não-paramétrica, por sua vez, é livre de distribuição, ou seja, não é necessário que se façam suposições sobre a distribuição de probabilidades das variáveis em estudo, como, a normalidade dos dados e a homogeneidade das variâncias [Öztuna et al., 2006; Marusteri & Bacarea, 2010; Nayak & Hazra, 2011; Hoskin, 2012]. Pelo fato dos testes não-paramétricos possuírem um leque de maior abrangência do que os testes paramétricos, muitas vezes, os pesquisadores possuem preferência de seu emprego por não dependerem da verificação das pressuposições.

Para cada teste estatístico paramétrico, conforme a Figura 3.10, existe um teste não-paramétrico correspondente. Os exemplos de testes paramétricos mais comuns encontrados na literatura são: o t de Student e o ANOVA. Da mesma forma, os testes não-paramétricos mais conhecidos são: o de Mann-Whitney U; de Wilcoxon; de Friedman e de Kruskal-Wallis [Marusteri & Bacarea, 2010; Nayak & Hazra, 2011; Hoskin, 2012; Torman et al., 2012].

Por causa da falta de conhecimento sobre os conceitos de robustez e poder de um teste estatístico, o desempenho dos testes utilizados pode ser comprometido [Marusteri & Bacarea, 2010; Nogueira & Pereira, 2013]. Um teste é classificado como robusto quando possui uma menor probabilidade de cometer o erro do tipo I, ou seja, de rejeitar a H_0 quando, na realidade, é verdadeira [Hoskin, 2012; Nogueira & Pereira, 2013]. Neste caso, os testes não-paramétricos tendem a ser mais robustos do que os testes paramétricos por terem uma boa performance mesmo com a violação das pressuposições [Marusteri & Bacarea, 2010]. Por outro lado, um teste

é considerado como mais poderoso se possui menor probabilidade de cometer o erro do tipo II, ou seja, não rejeitar a H_0 quando esta é falsa [Hoskin, 2012; Nogueira & Pereira, 2013]. Os testes paramétricos por não violarem as pressuposições são ditos como mais poderosos e menos questionáveis do que os testes não-paramétricos [Marusteri & Bacarea, 2010].

Para informações complementares a respeito dos testes estatísticos, além das referências bibliográficas utilizadas, recomenda-se a leitura do livro de Ribeiro Júnior [2012].

Capítulo 4

Abordagem de Aprendizagem de Máquina

Neste capítulo, é proposta uma abordagem computacional baseada em algoritmos de aprendizagem de máquina para construir modelos preditivos usando dados de rastreamento ocular móvel, do inglês, mobile eye tracking (MET). A intenção é proporcionar resultados que possam auxiliar o estudo da investigação médica no processo decisório em bioética clínica, como, na prática de eutanásia.

A eutanásia, que é claramente um tema clássico na bioética clínica, é uma questão conflitante, tendo em vista a subjetividade dos conceitos que envolvem este tema [Lepargneur, 2009]. Segundo França [2009] e Felix et al. [2013], eutanásia é definida como a boa morte, morte tranquila ou apropriada, sem dor nem sofrimento ou golpe de misericórdia. Também é dita como a ação ou omissão por parte do médico ou de outra pessoa com intenção de, por compaixão, provocar o óbito do paciente em sofrimento insuportável, a pedido voluntário e competente dessa pessoa com finalidade de acabar com a sua dor [Berghs et al., 2005; Radbruch et al., 2016]. Segundo Goldim [2006] e Lepargneur [2009], é quando uma pessoa quer livremente morrer, mas não consegue realizar seu desejo sozinha, por motivos físicos, e solicita o auxílio de um outro indivíduo (suicídio assistido).

Embora os estudos encontrados na literatura tenham demonstrado uma evolução importante na pesquisa de rastreamento ocular, eles não foram realizados em conjunto com suporte computacional avançado, como, por exemplo, empregando técnicas de aprendizagem de máquina, especialmente em bioética. Portanto, a abordagem proposta é inovadora por aplicar essas técnicas em um novo contexto ainda não explorado.

Os modelos preditivos obtidos neste estudo resultaram em padrões e regras in-

interessantes que associam os aspectos visuais relacionados à cognição com o processo de tomada de decisão. Esses resultados possuem um grande potencial para futuras pesquisas com o objetivo de desenvolver ferramentas computacionais; testar hipóteses científicas; auxiliar nas definições de equipe médicas mais aptas e no processo de tomada de decisão em situações críticas onde não há consenso moral.

4.1 Framework de aprendizagem de máquina

Embora este estudo se centre no caso específico de correlação das características visuais com a prática da eutanásia, vale ressaltar que a ideia geral do procedimento aqui apresentado pode ser aplicada a qualquer campo que use recursos MET para associar dados de visão com resultados relativos ao estudo em questão. A Figura 4.1, a seguir, ilustra, resumidamente, os passos utilizados na abordagem proposta que serão detalhados nas seções seguintes.

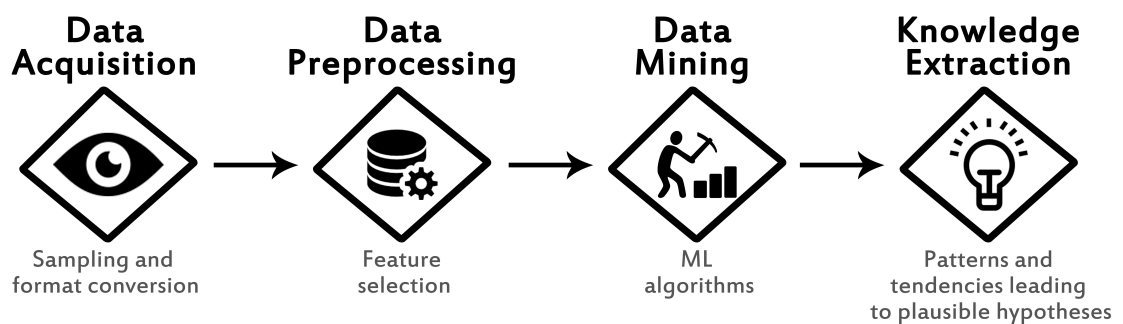


Figura 4.1. Pipeline de extração de conhecimento.

O pipeline apresentado é semelhante ao processo de KDD. O pipeline é dividido em quatro passos, sendo eles: aquisição, pré-processamento, mineração de dados e extração de conhecimento. O primeiro passo consiste em coletar dados oculares através do MET, realizar a amostragem adequada e converter os dados amostrados resultantes em um formato adequado. O segundo passo compreende o pré-processamento de dados, ou seja, melhorar a qualidade dos dados e escolher as variáveis adequadas por seleção de atributos. O terceiro passo inclui experimentos com algoritmos de aprendizagem de máquina no conjunto de treinamento resultante e a avaliação do desempenho estatístico dos modelos gerados. O quarto e último passo consiste em extrair conhecimento por meio da obtenção de padrões e tendências que podem levar a hipóteses plausíveis.

Nas próximas seções, cada passo será apresentado de forma mais detalhada, informando os métodos e técnicas utilizadas, a sua importância para os experimentos realizados, os resultados encontrados e o conhecimento inferido no final para resolução da problemática.

4.2 Materiais e métodos

4.2.1 Conjunto de dados

O conjunto de treinamento usado neste estudo pelos algoritmos de aprendizagem de máquina foi criado com base em dados coletados de 75 alunos voluntários (17 a 47 anos, média de 21,36, 14 do sexo masculino) do primeiro ao nono período do curso de graduação de Enfermagem de uma universidade pública brasileira durante o ano de 2015¹. Os dados foram coletados usando o rastreador ocular móvel conhecido como Mobile Eye-XG da Applied Science Laboratories (ASL) com uma cadência de 30 Hz².

O Mobile Eye-XG é um dispositivo capaz de monitorar o comportamento dos olhos ao analisar a visão central dos participantes usando câmeras especializadas conectadas aos óculos. Para esse fim, o equipamento usa uma fonte de luz infravermelha para determinar os locais de fixação da visão central mínima e exata a cada 40 ms, além de identificar a imagem da pupila e da córnea [Granka et al., 2004]. Antes da aquisição de dados, o ambiente de coleta foi devidamente preparado e o dispositivo MET foi calibrado.

Em seguida, um trecho de dois minutos e 24 segundos do filme “Million Dollar Baby”, dirigido por Clint Eastwood e lançado em 2005, conhecido no Brasil como “Menina de Ouro”, foi mostrado na mesma tela usada para calibração do Mobile Eye-XG para que os movimentos oculares dos participantes pudessem ser identificados. O filme conta a história de Maggie Fitzgerald, uma pugilista que sofreu uma lesão grave e se tornou tetraplégica após uma luta injusta. Com o tempo, ocorreram várias complicações, o que levou Maggie a tentar suicídio. Depois de falhar, ela se virou para seu treinador (Frankie) e pediu-lhe a eutanásia.

Depois que a boxeadora expressou sua vontade e fez o pedido ao treinador, o trecho do filme foi interrompido em um momento crucial quando a decisão seria tomada. Os participantes da pesquisa foram então questionados sobre sua decisão

¹Dados brutos disponíveis em: <https://drive.google.com/drive/folders/OB8aRarbjcs87VGDLa3FvYnZRV1k?usp=sharing>

²O equipamento utilizado nesta pesquisa: <https://www.youtube.com/watch?v=fvJPVInvsNk>

(seja para cumprir ou não o pedido da boxeadora) e foram convidados a escrever os motivos. Essas respostas permaneceram confidenciais. No final do experimento, os seguintes dados foram obtidos de cada participante: (i) o registro computacional do rastreamento ocular e (ii) a resposta na decisão juntamente com sua justificativa. O trecho do filme é composto por cenas simples que retratam o diálogo dos personagens no quarto de um hospital. As cenas do trecho podem ser visualizadas na Figura 4.2 abaixo.



Figura 4.2. Capturas de tela do filme que os participantes assistiram, sendo: (a) cena da Maggie e (b) cena do Frankie. Fonte: Eastwood & Hagglund [2005].

Durante a composição do conjunto de treinamento, as justificativas das respostas na tomada de decisão dos participantes foram analisadas e observou-se que a decisão de alguns participantes foi influenciada por questões como a sacralidade da vida, a religião ou o código de ética dos profissionais de enfermagem. Portanto, uma vez que um dos objetivos deste estudo é investigar a correlação dos dados oculares com o processo de tomada de decisão, os dados provenientes desses participantes foram removidos do nosso conjunto, evitando assim distorções indesejadas nos modelos preditivos.

No final, o conjunto de treinamento continha 27 casos positivos (iriam praticar eutanásia) e 16 casos negativos, ou seja, 43 participantes (17 a 26 anos, média de 20,93, dez do sexo masculino) que, de acordo com suas justificativas, eram neutros e tinham nenhuma das influências mencionadas acima. Depois de concluir o conjunto de treinamento final, ele foi convertido no formato ARFF, sendo este o formato mais apropriado para ambos os softwares de mineração de dados utilizados e que serão mencionados mais a diante.

4.2.2 Atributos

O próximo passo, depois de gerar o conjunto de treinamento, é a seleção de atributos, ou seja, definindo quais atributos são relevantes e não redundantes entre os 47 iniciais (incluindo dados socioeconômicos e oculares), dado o escopo da pesquisa pela equipe de pesquisadores. Observe que o atributo da classe, ou seja, a decisão ética (prática ou não a eutanásia) não faz parte do procedimento de seleção de atributos.

A seleção reduz a complexidade do conjunto de dados, o que torna a construção do modelo mais rápida e, em muitos casos, produz melhores resultados [Kohavi & John, 1997; Cerqueira et al., 2014]. A seleção pode ser realizada através de métodos computacionais, por intervenção humana, ou por ambos. Neste trabalho, a seleção foi feita em três etapas, primeiramente pela intervenção humana e posteriormente por meio de métodos computacionais.

Primeiramente, 12 atributos foram identificados como socioeconômicos (por exemplo, ID, idade, sexo, religião, escolaridade dos pais, etc.), enquanto os restantes 35 eram atributos oculares (por exemplo, fixação em uma AOI, dilatação da pupila, duração das fixações nas AOI, etc). À medida que este estudo se concentra em buscar correlações entre dados oculares e tomada de decisão na bioética, os atributos socioeconômicos foram removidos do conjunto de dados.

Em seguida, com base em uma análise anterior realizada por meio da experiência da equipe especializada de pesquisadores, 14 atributos oculares foram removidos manualmente por serem claramente redundantes (por exemplo, a mesma medida é englobada por duas variáveis, uma como valor percentual e outra como valor absoluto). Assim, a dimensionalidade do conjunto inicial foi reduzida para 21 atributos numéricos, juntamente com o atributo da classe de decisão binária, como mostrado na Tabela 4.1.

Em seguida, realizamos a seleção computacional de atributos usando as técnicas de IG e CFS, mencionadas na Subseção 3.2.1.1. Ambas as técnicas de seleção foram aplicadas aos 21 atributos do conjunto de treinamento para verificar quais atributos seriam os mais apropriados. No caso do CFS, a seleção depende de uma heurística de pesquisa efetiva que permita diminuir o número de subconjuntos a serem avaliados [Cerqueira, 2010]. Neste contexto, utilizamos o CFS com cinco estratégias de busca diferentes, sendo elas: Best First Search [Rich & Knight, 1991], Greedy Forward or Backward Search [Hall et al., 2009], Evolutionary Search [Kim et al., 2000], Scatter Search [López et al., 2006] e, Tabu Search [Hedar et al., 2008].

Após a aplicação de ambas as técnicas do tipo filtro, o único atributo retornado entre os 21 presentes no conjunto de treinamento foi o `vertical_pupil_dilation`, com

Atributo	Descrição
<code>maggies_face_fixation</code>	Número de fixações no rosto da Maggie
<code>tube_fixation</code>	Número de fixações no tubo traqueal da Maggie
<code>maggies_chest_fixation</code>	Número de fixações no peito da Maggie
<code>maggies_scene_space_fixation</code>	Número de fixações no espaço da cena da Maggie
<code>frankies_face_fixation</code>	Número de fixações no rosto do Frankie
<code>frankies_chest_fixation</code>	Número de fixações no peito do Frankie
<code>frankies_scene_space_fixation</code>	Número de fixações no espaço da cena do Frankie
<code>without_classification_fixation</code>	Número de fixações em áreas sem classificação
<code>total_fixation</code>	Número total de fixações durante o experimento
<code>maggies_face_duration</code>	Tempo das fixações no rosto da Maggie (ms)
<code>tube_duration</code>	Tempo das fixações no tubo traqueal da Maggie (ms)
<code>maggies_chest_duration</code>	Tempo das fixações no peito da Maggie (ms)
<code>maggies_scene_space_duration</code>	Tempo das fixações no espaço da cena da Maggie (ms)
<code>frankies_face_fixation_duration</code>	Tempo das fixações no rosto do Frankie (ms)
<code>frankies_chest_duration</code>	Tempo das fixações no peito do Frankie (ms)
<code>frankies_scene_space_duration</code>	Tempo das fixações no espaço da cena do Frankie (ms)
<code>without_classification_duration</code>	Tempo das fixações em áreas sem classificação
<code>total_duration</code>	Tempo total de fixações durante o experimento (ms)
<code>overall_pupil_dilation</code>	Valor médio do diâmetro geral da pupila (pixels)
<code>horizontal_pupil_dilation</code>	Valor médio do diâmetro horizontal da pupila (pixels)
<code>vertical_pupil_dilation</code>	Valor médio do diâmetro vertical da pupila (pixels)
<code>decision</code>	Classe

Tabela 4.1. Atributos restantes após o passo de seleção manual.

IG de 0,225 e *Merit* de 0,265. Este experimento mostrou que a permanência de um único atributo não produz melhores resultados do que quando são usados mais atributos. Isso corresponde à observação de Santoro [2005], que salienta que a permanência de apenas um atributo, embora recomendada por uma abordagem de filtro, pode causar perda de poder preditivo pelos algoritmos de classificação em certos casos.

Considerando que a aplicação de filtros não funcionou bem, isto é, não conseguiram capturar o impacto real dos atributos na classificação, realizou-se um último passo que fez uso de outro método de seleção do tipo wrapper. Ao contrário dos métodos do filtro, a abordagem do wrapper usa um algoritmo de classificação, o mesmo que será mais tarde empregado no processo de aprendizagem, para avaliar a qualidade de cada subconjunto de atributo [Kohavi & John, 1997].

O método wrapper consiste em uma abordagem computacionalmente exigente usada para explorar, por meio de heurísticas, o espaço de busca dos subconjuntos de atributos que serão posteriormente avaliados por algoritmos de classificação [Santoro, 2005; Cerqueira, 2010]. A dependência desses algoritmos geralmente produz maior precisão de classificação durante o processo de aprendizagem em compara-

ção com os métodos de filtragem [Cerqueira, 2010; Villacampa, 2015]. Este método de seleção de atributos é limitado pelo aumento do custo computacional à medida que o número de atributos aumenta. Desta forma, usar algoritmos de classificação para avaliar vários subconjuntos pode ser inviável [Santoro, 2005; Cerqueira, 2010; Villacampa, 2015].

No entanto, o custo computacional não é um problema neste caso, uma vez que os 21 atributos das 43 instâncias contidas no conjunto de treinamento representam um valor que pode ser processado em tempo hábil, de modo que mesmo as possibilidades de $2^{21} - 1$ possam ser analisados. Assim, usando o método wrapper juntamente com a estratégia de busca exaustiva para analisar sistematicamente todas as possibilidades do subconjunto, a combinação de atributos que resultaram no modelo com maior poder preditivo inclui cinco atributos, nomeadamente: `maggies_chest_fixation`, `tube_duration`, `total_duration`, `overall_pupil_dilation`, and `vertical_pupil_dilation`.

Com o objetivo de saber o quão importante são os cinco atributos selecionados, eles foram classificados com base na medida F, do inglês, F-measure. Esta medida representa a média harmônica entre precisão e sensibilidade e que atinge seu melhor valor em 1 e o pior em 0 [Powers, 2011]. O ranking foi realizado removendo um atributo do subconjunto por vez e depois executando a classificação. Após a classificação, o valor da medida F do modelo com o conjunto de treinamento original (que inclui os cinco atributos) foi subtraído do valor da medida F do modelo com o subconjunto contendo o número reduzido de atributos (quatro atributos). O valor da subtração representa o impacto de cada atributo na classificação. Portanto, quando um atributo é removido e sua ausência faz com que o valor da medida F original caia, pode-se inferir que essa queda está associada à sua importância. Assim, quanto maior a queda, mais importante o atributo no modelo. A Tabela 4.2 mostra o ranking dos atributos retornados pelo método envolvido com base na medida F.

Atributo	Pontuação (Decadência da medida F)
<code>overall_pupil_dilation</code>	28,2
<code>vertical_pupil_dilation</code>	24,2
<code>tube_duration</code>	19,3
<code>maggies_chest_fixation</code>	14,0
<code>total_duration</code>	6,7

Tabela 4.2. Nível de importância dos atributos com base na medida F.

Após essas etapas de seleção de atributos, o foco tornou-se testar os algoritmos de aprendizagem para criar modelos preditivos com o objetivo de extrair regras

que relacionem os cinco atributos mencionados com a classe, de modo que o conhecimento sobre aspectos de busca visual durante a tomada de decisão possa ser extraído.

4.2.3 Ferramentas de aprendizagem de máquina

O presente estudo usou dois programas de mineração de dados, ambos de código aberto e multiplataforma. Os vários algoritmos de aprendizagem de máquina que serão mencionados na subseção a seguir foram aplicados usando o software Waikato Environment for Knowledge Analysis (WEKA³), que é amplamente utilizado para tarefas de classificação em vários trabalhos científicos [Hall et al., 2009].

O outro software utilizado foi NICEsim, um simulador que visa compreender a correlação entre os atributos e a variável de classe [Cerqueira et al., 2014]. O ambiente de simulação inclui uma interface gráfica para o pesquisador alterar os valores dos atributos para criar cenários distintos (instâncias simuladas). Para cada intervenção do usuário, o sistema mostra a probabilidade da instância simulada ser positiva. Como resultado, as tendências podem ser extraídas de acordo com os cenários simulados, que facilitam o estudo de hipóteses científicas. As simulações realizadas no NICEsim baseiam-se em modelos preditivos criados com o Support Vector Machine (SVM) e MLP.

4.2.4 Algoritmos de aprendizagem

Vários algoritmos podem ser usados na tarefa de classificação para prever a qual classe pertence um determinado registro. A seleção de algoritmos de classificação deve ser baseada em alguns critérios, como: precisão; robustez; escalabilidade; velocidade e interoperabilidade do modelo [Fayyad et al., 1996; Tan et al., 2006].

No presente estudo, alguns algoritmos de classificação supervisionada foram testados, todos amplamente utilizados em pesquisas de informática biomédica [Chen et al., 2005]. Devido à característica interdisciplinar do presente trabalho, este texto também é dirigido a outros profissionais, além de cientistas computacionais ou matemáticos. Portanto, uma breve explicação de cada um desses algoritmos é dada abaixo.

Naïve Bayes é um algoritmo de classificação que fornece uma abordagem simples e rápida baseada no teorema de Bayes [John & Langley, 1995]. Para estimar a

³Todos os resultados gerados pelos experimentos computacionais foram obtidos pelo WEKA versão de desenvolvimento 3.7.12 64 bits em um notebook Intel Core i7-4500 @1.80GHz com 8 GB de RAM rodando o Windows 8.1 Pro 64 bits.

probabilidade do vetor de atributos condicionados à classe, assume-se que os atributos são independentes, o que simplifica o cálculo. Enquanto não houver dependência, ou esta seja fraca, o algoritmo resulta em um modelo preciso com um tempo de aprendizagem bastante curto em comparação com vários outros algoritmos [John & Langley, 1995; Tan et al., 2006].

A abordagem SVM, para dados linearmente separáveis, estabelece um hiperplano de margem máxima como limite de decisão para evitar overfitting. Para esse fim, o objetivo do SVM é formulado como um problema de programação quadrática. Quando os dados não são linearmente separáveis, é utilizada uma função kernel para mapear os dados para um espaço de dimensão superior para torná-los linearmente separáveis e, portanto, aplicar o método de separação linear [Tan et al., 2006]. O Sequential Minimal Optimization (SMO) é um algoritmo amplamente utilizado para resolver o problema de programação quadrática estabelecido na abordagem do SVM. Diferente das implementações tradicionais de SVM, o SMO usa um otimizador analítico e memória linear para lidar com grandes conjuntos de dados [Bottou & Lin, 2007].

O algoritmo de aprendizagem k -Nearest Neighbour (k NN) classifica uma instância ao analisar seus k vizinhos mais próximos [Michalski et al., 2013]. A classe é atribuída de acordo com a classe majoritária dos vizinhos selecionados. Neste cálculo, os votos geralmente têm pesos diferentes de acordo com a distância do vizinho, ou seja, quanto mais perto o vizinho, maior será o peso que sua classe terá na definição da classe da instância em análise. Portanto, este algoritmo é altamente impactado pela métrica de proximidade escolhida para a definição da vizinhança [Tan et al., 2006; Michalski et al., 2013].

Random Forest é um método de classificação em conjunto que usa o resultado de várias árvores de decisão. A decisão final da classe é adquirida por um esquema de votação em que a classe majoritária obtida das diferentes árvores prevalece. Para construir cada árvore, é utilizada uma amostra do conjunto de treinamento. Cada amostra é construída extraindo um subconjunto aleatório dos atributos originais [Quinlan, 1986; Tan et al., 2006].

Além desses quatro algoritmos citados, também foram utilizados a rede MLP e a árvore de decisão C4.5, descritos anteriormente nas Subseções 3.1.1.1 e 3.1.1.2, respectivamente.

4.3 Resultados e discussão

4.3.1 Avaliação do desempenho do modelo

Utilizou-se a validação cruzada 10-partes para avaliar os modelos de aprendizagem de máquina resultantes. A Tabela 4.3 resume o desempenho de cada algoritmo de aprendizagem mencionado na seção anterior, com e sem classificação sensível ao custo (CSC), em relação a acurácia média (Ac), precisão (Pr), sensibilidade (Sn), área média sob a curva ROC (AUC) e, intervalo de confiança da acurácia (IC). É importante mencionar que SMO foi usado com (RBFkernel) e sem (classificação linear) kernel. Além disso, o número de vizinhos foi definido como 3 para o k NN. Exceto por essas modificações, os parâmetros padrão do WEKA foram mantidos.

Classificador	Tipo	Ac	Pr	Sn	AUC	IC (95%)
Naïve Bayes	Normal	65.1%	64.7%	65.1%	0.660	[50.16%, 80.04%]
MLP	Normal	90.7%	91.9%	90.7%	0.903	[75.76%, 100.00%]
SMO (sem kernel)	Normal	62.8%	58.4%	62.8%	0.513	[47.86%, 77.74%]
SMO (RBFKernel)	Normal	62.8%	39.4%	62.8%	0.500	[47.86%, 77.74%]
k NN	Normal	72.1%	75.1%	72.1%	0.667	[57.16%, 87.04%]
C4.5	Normal	65.1%	64.7%	65.1%	0.668	[50.16%, 80.04%]
Random Forest	Normal	62.8%	61.1%	62.8%	0.686	[47.86%, 77.74%]
Naïve Bayes	CSC	58.1%	60.4%	58.1%	0.660	[43.16%, 73.04%]
MLP	CSC	81.4%	81.4%	81.4%	0.852	[66.46%, 96.34%]
SMO (sem kernel)	CSC	67.4%	71.1%	67.4%	0.690	[52.46%, 82.34%]
SMO (RBFKernel)	CSC	44.2%	46.7%	44.2%	0.428	[29.26%, 59.14%]
k NN	CSC	72.1%	75.1%	72.1%	0.635	[57.16%, 87.04%]
C4.5	CSC	67.4%	71.1%	67.4%	0.667	[52.46%, 82.34%]
Random Forest	CSC	60.5%	59.0%	60.5%	0.699	[45.56%, 75.44%]

Tabela 4.3. Avaliação estatística (validação cruzada 10-partes) dos modelos gerados a partir de vários algoritmos de aprendizagem com e sem CSC.

No nosso caso, a classe positiva é a maioria, mas o desequilíbrio entre ela e a classe negativa não é grave. Ainda assim, considerou-se prudente incluir a CSC nos experimentos a serem comparados com as versões de aprendizado com custos homogêneos para averiguar se a diferença numérica entre as classes causaria qualquer tipo de viés nos modelos.

Como pode ser visto, o modelo gerado pela rede MLP com custos homogêneos se destaca consideravelmente em comparação com os outros modelos, proporcionando melhor desempenho em todas as métricas. O valor elevado da AUC (Figura 4.3) demonstra um alto poder preditivo e significa que as probabilidades geradas pelo modelo são robustas. Essas probabilidades são fundamentais para encontrar os

padrões relatados nas próximas subseções.

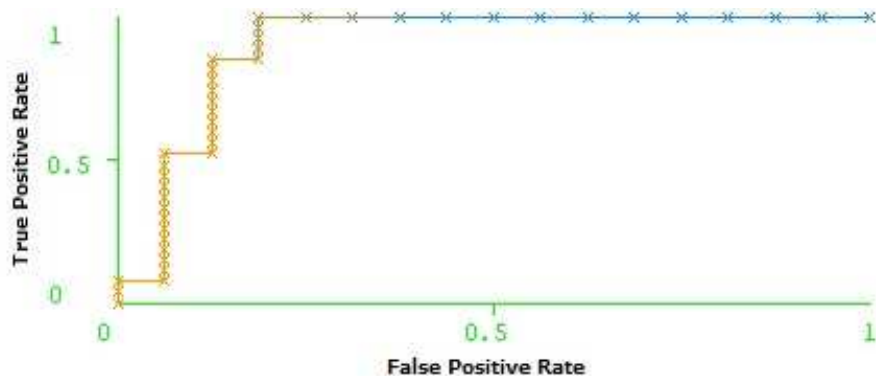


Figura 4.3. Curva ROC do modelo de MLP gerado. A AUC média é equivalente a 0,903.

4.3.2 Extraindo padrões por meio de simulações com modelos MLP

No contexto da extração de conhecimento, o NICEsim foi utilizado para simular vários cenários, combinando os valores dos atributos (instâncias simuladas), a fim de proporcionar uma melhor compreensão da relação entre os atributos e a classe. Após realizar as simulações, foram gerados cinco gráficos que mostram a relação entre os atributos e a probabilidade de aceitar participar da prática da eutanásia.

As Figuras 4.4 e 4.5 ilustram a relação entre a duração das fixações e a probabilidade de aceitar a participar da prática da eutanásia. A duração da fixação do olhar é uma medida que examina a relação entre a carga cognitiva e a cena de visualização [Granka et al., 2004]. Esta medida corresponde a duração do processamento cognitivo de uma pessoa em relação a uma determinada AOI [Irwin, 2004]. Como pode ser visto na Figura 4.4, para fixações com duração de 1.400 ms, os participantes apresentam um pouco mais de 60% de chance de aceitar a participar da prática da eutanásia. Esse valor é próximo de 100% quando a duração das fixações ultrapassa 32.768 ms. A Figura 4.5, por outro lado, mostra que as fixações com duração inferior a 7.768 ms no tubo traqueal (traqueostomia) levam praticamente a nenhuma chance (0%) de aceitar a participar da prática da eutanásia, enquanto valores acima de 11.652 ms mostram que a chance é próxima de 100%. Ambos os gráficos demonstram que, quanto maior for a fixação geral, maior a chance dos participantes em aceitar a participar da prática da eutanásia. Acredita-se que os

participantes que tiveram mais atenção as AOIs - como a traqueostomia - ficaram mais consternados com a situação da paciente e, conseqüentemente, optaram por interromper o sofrimento.

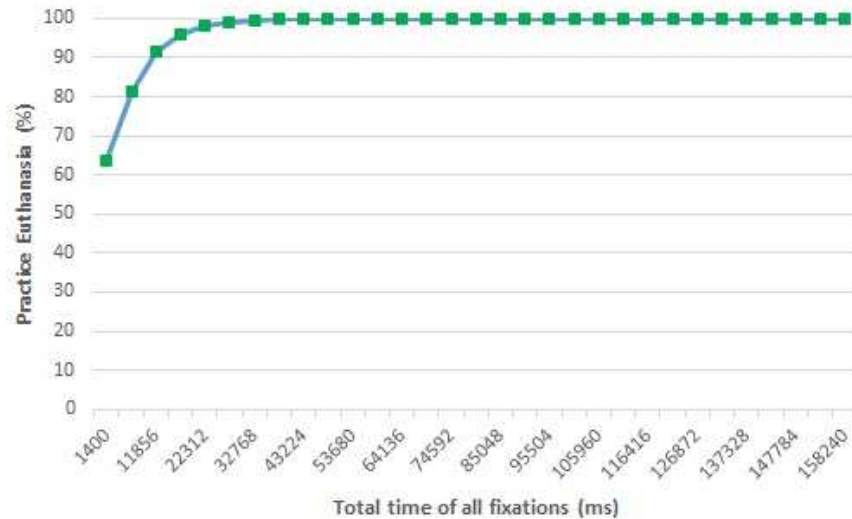


Figura 4.4. Probabilidade prevista de aceitar a participar da prática da eutanásia versus tempo total de duração das fixações gerais.

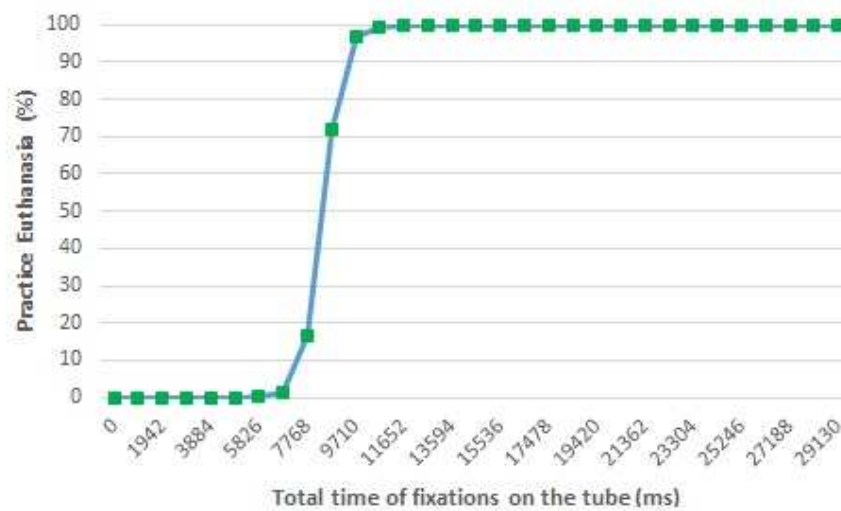


Figura 4.5. Probabilidade prevista de aceitar a participar da prática da eutanásia versus tempo de duração das fixações no tubo traqueal.

A Figura 4.6 indica que a atenção no peito de Maggie é um fator determinante, mostrando que este é definitivamente um AOI, porque elementos importantes recebem fixações mais frequentemente [Fitts et al., 2005]. Nesta análise, um número

baixo de fixações (abaixo de 32) representa uma alta probabilidade de aceitar a participar da prática da eutanásia (aproximadamente 100%), enquanto valores acima de 44 reduzem a chance de aproximadamente 0%. À primeira vista, a tendência mostrada na Figura 4.6 parece antagônica às conclusões tiradas das Figuras 4.4 e 4.5. Mas, é importante notar que a área do tórax da paciente tem um cordão com um crucifixo. Uma explicação plausível é que, quanto mais os participantes fixaram seus olhos na região desse objeto religioso, procurando informações, mais seu viés religioso foi levado em consideração. Mesmo que os participantes que declararam ser religiosos tenham sido removidos do conjunto de dados, quando se trata de processamento cognitivo, nada impede que o objeto na cena talvez tenha influenciado a decisão dos participantes até certo ponto.

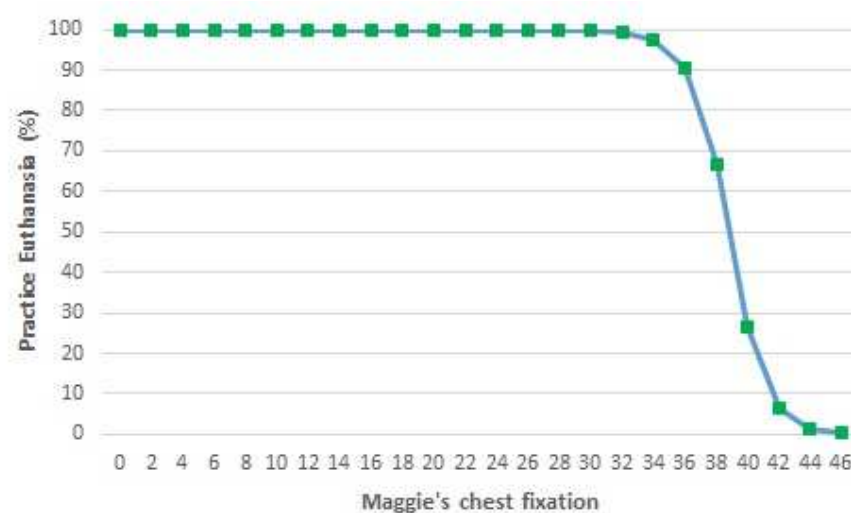


Figura 4.6. Probabilidade prevista de aceitar a participar da prática da eutanásia versus quantidade de fixações no peito de Maggie (que contém um cordão com um crucifixo).

Para o comportamento da pupila, tanto vertical (Figura 4.7) quanto geral (Figura 4.8), os resultados mostram que quanto maior o diâmetro da pupila, menor a probabilidade de aceitar a participar da prática da eutanásia. Os valores acima de 312 pixels indicam uma probabilidade muito baixa de aceitar a participar da prática da eutanásia, enquanto valores abaixo de 268 pixels indicam uma chance próxima de 100%. A pupila representa com precisão o esforço cognitivo, bem como o julgamento do interesse de alguém na realização de uma determinada tarefa e, é considerado um importante elemento cognitivo [Granholm & Steinhauer, 2004; Cardoso & Costa, 2013]. Então, uma dilatação elevada da pupila indica um alto nível de esforço cognitivo. Quando o estímulo diminui, a pupila se contrai até retornar ao

seu tamanho de repouso [Cardoso & Costa, 2013] - um processo também conhecido como miose [Hess & Polt, 1964]. Neste contexto, acredita-se que os participantes que tiveram uma demanda energética muito alta optaram por uma decisão mais racional do que emocional, respeitando o juramento e o código de ética dos profissionais de enfermagem. Além disso, a eutanásia é considerada homicídio de acordo com os artigos 121 e 122 do Código Penal Brasileiro vigente [Dodge, 2009].

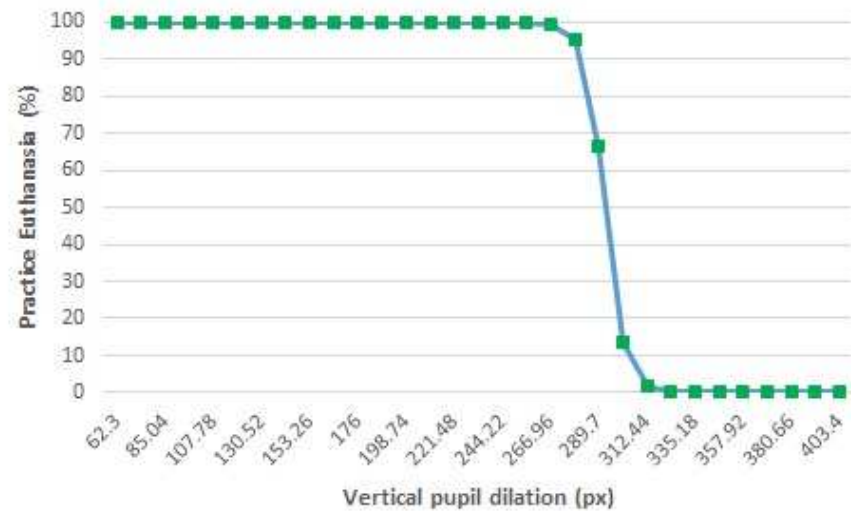


Figura 4.7. Probabilidade prevista de aceitar a participar da prática da eutanásia versus tamanho da dilatação vertical média da pupila.

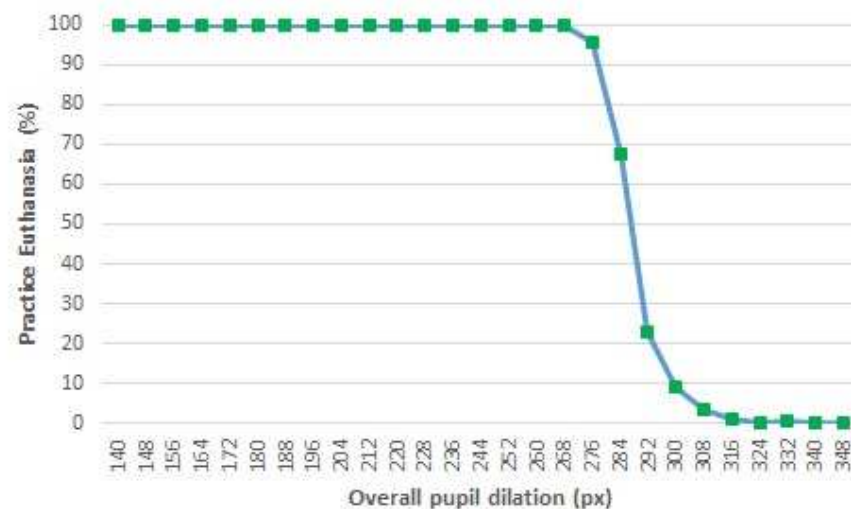


Figura 4.8. Probabilidade prevista de aceitar a participar da prática da eutanásia versus tamanho da dilatação geral média da pupila.

Os experimentos demonstram que as variáveis escolhidas são adequadas para construir modelos preditivos para a tomada de decisão ética estudada. Era esperado, uma vez que os componentes avaliados pelo MET estão associados às capacidades cognitivas e emocionais do participante [Granholm & Steinhauer, 2004]. Portanto, a relação dessas variáveis com a tomada de decisão ética mostra o papel significativo que a visão desempenha no caso estudado.

4.3.3 Extraíndo padrões por meio de regras de árvores de decisão

Embora a Tabela 4.3 mostre que C4.5 é muito menos preciso quando comparado ao MLP, um experimento foi realizado para verificar se as regras extraídas de um modelo C4.5 corroborariam com as tendências encontradas a partir do modelo MLP. A Figura 4.9 ilustra as regras junto com a árvore de decisão de onde as regras podem ser extraídas e demonstram que os padrões encontrados correspondiam aos resultados do modelo MLP. Observe que as regras foram tiradas de uma abordagem C4.5 sensível ao custo, porque o modelo C4.5, na sua forma original, apresentou desempenho pior (Tabela 4.3).

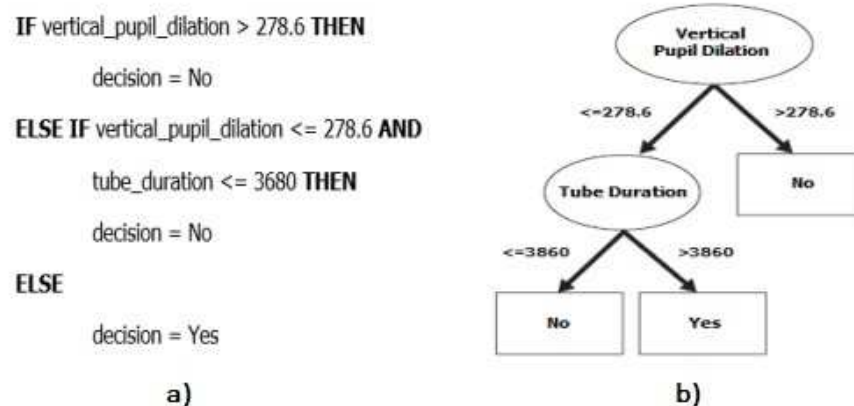


Figura 4.9. Padrões gerados por um modelo C4.5 sensível ao custo. As subfiguras (a) e (b) demonstram as regras extraídas e a árvore de decisão pelo modelo C4.5 sensível ao custo.

As regras obtidas mostram que os atributos referentes à dilatação vertical da pupila e duração da fixação na traqueostomia da paciente são fatores que melhor predizem a decisão a ser tomada.

A dilatação do pupila pode ser indicativa da carga cognitiva no momento da tomada de decisão [Granholm & Steinhauer, 2004]. O aumento da ativação cortical

no sistema límbico, devido aos níveis de recursos psicológicos utilizados para resolver problemas (como atenção, raciocínio e julgamento), promove um aumento no diâmetro pupilar - relacionado à ativação do SNS [Critchley et al., 2013] - um processo também conhecido como midríase [Hess & Polt, 1964]. O resultado mostrado na Figura 4.9 evidencia mais uma vez uma correlação entre a dilatação da pupila e o processo decisório do participante no presente protocolo experimental, o que mostra que a carga cognitiva estimulada pelo nível de dificuldade da tarefa influencia principalmente a tomada de decisão do participante. Ou seja, de acordo com a árvore de decisão, para alcançar a decisão “não”, a dilatação vertical da pupila deve ser necessariamente alta (significando alta carga cognitiva - neste caso, provavelmente relacionada a uma decisão racional).

Os resultados apresentados nas Figuras 4.4, 4.5, 4.6, 4.7, 4.8 e na Figura 4.9 indicam que também é muito importante analisar a região onde o participante procura informações para tomar decisões, para que possamos entender se as atividades cognitivas estão relacionadas a emoção ou razão. Da mesma forma que as figuras da Subseção 4.3.2, as regras mostradas na Figura 4.9 fornecem evidências de que uma alta carga cognitiva, sinalizada pelo alto número de fixações no tubo traqueal, relaciona-se com a decisão “sim”. Neste caso, é provável uma carga emocional devido ao fato de que o tubo demonstra que a paciente está realmente sofrendo, que é um dos elementos centrais para a tomada de decisões éticas em casos de eutanásia [Berghs et al., 2005].

4.4 Conclusão

A presente investigação propõe uma abordagem computacional, baseada em técnicas de aprendizagem de máquina e rastreamento ocular, que visa investigar dados oculares gerados a partir da apreciação de um trabalho cinematográfico para auxiliar estudos que investigam a influência da busca visual no processo de tomada de decisão bioética, em uma questão muito relevante e controversa de prática clínica.

Para esse fim, os algoritmos de aprendizagem de máquina foram aplicados - nomeadamente, técnicas de classificação supervisionada para um conjunto de dados construído a partir de dados oculares de 75 (embora seja aparentemente baixo, este número é expressivo em relação a estudos anteriores), a fim de entender como a atenção visual afeta a decisão ética dos futuros profissionais. Para o conjunto de dados utilizado, o MLP foi o mais preciso entre todos os algoritmos de classificação empregados, apresentando 90,7% de acurácia e AUC média de 0,903.

A abordagem computacional aqui apresentada provou ser útil para pesquisas sobre cognição, – mais especificamente sobre atenção visual –, uma vez que ele constrói e valida um modelo preditivo altamente preciso para simular os efeitos sobre as variáveis relativas ao número de fixações, a duração da fixação em AOIs e a dilatação da pupila sobre a probabilidade dos profissionais da área da saúde em aceitar a participar da prática da eutanásia. Portanto, os resultados deste estudo mostram claramente uma conexão entre o comportamento cognitivo e a tomada de decisão ética.

Capítulo 5

Abordagem de Estatística e Visualização de Dados

Nos últimos anos, a análise dos movimentos dos olhos e da atenção visual têm ganhando popularidade como importantes temas de investigação em diversos campos de aplicações quando deseja-se estudar o comportamento cognitivo de indivíduos durante a realização de tarefas.

Para tirar conclusões sobre os comportamentos dos sujeitos, é necessário comparar um determinado conjunto de dados e organizá-los de acordo com padrões comuns encontrados entre os sujeitos [Wu & Munzner, 2015]. Contudo, os métodos de visualização de dados existentes para analisar os movimentos oculares são limitados e sozinhos não permitem interpretar facilmente os dados [Räihä et al., 2005; Bednarik, 2007]. As técnicas de visualização, geralmente, são capazes de analisar somente uma ou outra métrica obtida pelo rastreador ocular (por exemplo, o MET), sendo, então, restritas a uma análise mais exploratória sem base estatística. Desta forma, a dificuldade de extrair conhecimento a partir dos dados gerados, torna-se um grande desafio [Andrienko et al., 2012; Blascheck et al., 2014].

Em virtude disso, para determinar qual técnica utilizar, deve-se selecionar os métodos adequados que suportem a análise desejada. Entretanto, é necessário realizar um levantamento exaustivo na literatura a fim de encontrar as técnicas mais eficazes para o estudo a ser investigado. Pesquisas sugerem a utilização de abordagens visuais aliadas a métodos estatísticos e de aprendizagem de máquina com intuito de facilitar a compreensão dos movimentos oculares humanos [Andrienko et al., 2012; Blascheck et al., 2014].

Neste capítulo, a abordagem apresentada é continuação da abordagem proposta anteriormente, na qual foi construído e validado um modelo preditivo alta-

mente preciso por meio de algoritmos de aprendizagem de máquina para simular os efeitos das variáveis oculares sobre a tomada de decisão em bioética clínica. Visto que a abordagem anterior não fez uso de recursos de visualização de dados e somente da estatística para validação do modelo preditivo, este capítulo consiste em apresentar uma abordagem baseada em técnicas de estatística e visualização de dados, como forma de realizar uma análise mais frutífera dos dados de rastreamento ocular, mais especificamente, dados sobre o tamanho da pupila durante intervalos de tempo ao longo do experimento.

O interesse em relação aos dados referentes a variabilidade do diâmetro pupilar é devido ser uma variável fisiológica mensurável associada às capacidades cognitivas e emocionais de sujeitos durante uma tomada de decisão e graças aos interessantes padrões encontrados no estudo anterior sobre o comportamento pupilar no caso bioético estudado [Beatty, 1982; Granholm & Steinhauer, 2004; Granka et al., 2004; Querino et al., 2015].

Os resultados alcançados pela presente investigação demonstram que, ao comparar os dados do diâmetro pupilar dos participantes na tomada de decisão, há diferenças significativas principalmente associadas ao aspecto da variabilidade da pupila que reflete diretamente no esforço cognitivo necessário para o processamento de informações. Além disso, o estudo também proporcionou uma maneira capaz de identificar características e padrões para formulação de possíveis hipóteses sobre como a tomada de decisão clínica pode ser influenciada, principalmente, pelas capacidades emocionais e racionais do participante.

5.1 Framework de análise de séries temporais

Da mesma forma que o Capítulo 4, nesta abordagem, também foi proposto um pipeline para extração de conhecimento que pode ser aplicado em qualquer estudo que requeira a utilização dos recursos dos dispositivos de rastreamento ocular para encontrar padrões úteis, válidos e não-triviais em uma grande quantidade de dados oculares. No entanto, o foco do pipeline está na análise do comportamento da variabilidade do diâmetro pupilar ao longo do tempo. A Figura 5.1 a seguir expõe os passos utilizados.

O pipeline consiste em quatro etapas consecutivas: aquisição, pré-processamento de dados, análise de séries temporais e extração de conhecimento. As etapas um, dois e quatro são comuns em processos de KDD e são praticamente as mesmas do pipeline do capítulo anterior, sendo, apenas adaptadas conforme o cu-

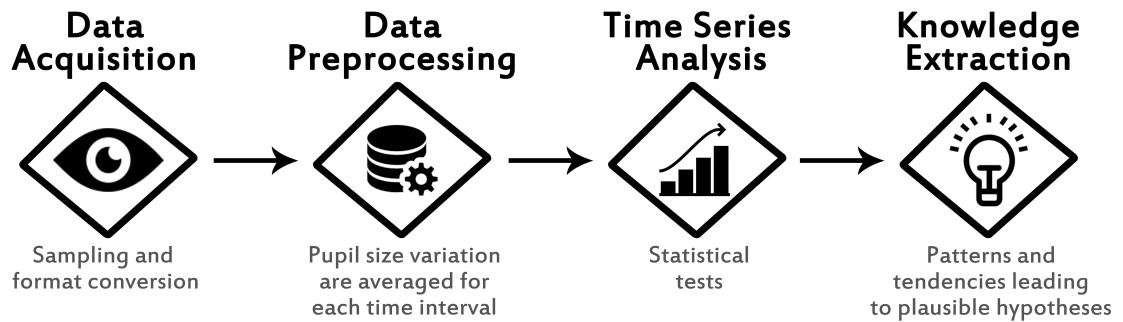


Figura 5.1. Pipeline de extração de conhecimento.

nho científico da pesquisa a ser realizada. Nossa principal contribuição está na etapa 3, onde apresentamos uma abordagem que, por meio da combinação de metodologias (testes estatísticos e de visualização de dados), auxilia equipes de pesquisadores na investigação e análise dos movimentos dos olhos ao longo do tempo, mais precisamente, na verificação da existência de diferenças significativas entre os movimentos pupilares dos sujeitos avaliados. Os detalhes sobre o estudo de caso serão fornecidos nas seções a seguir.

5.2 Materiais e métodos

5.2.1 Conjunto de dados

Foi utilizado nesta abordagem o conjunto de dados resultante do experimento da abordagem anterior, isto é, composto pelos dados de 43 estudantes voluntários do curso de Enfermagem de uma universidade pública do Brasil. Assim, como mencionado na Subseção 4.2.1, utilizou-se o dispositivo Mobile Eye-XG da ASL com cadência de 30 Hz para coletar os dados oculares durante o trecho do filme “Million Dollar Baby” ou “Menina de Ouro”, com cerca de dois minutos e 24 segundos de duração. Ao final do experimento, foram obtidos de cada participante além da tomada de decisão, os registros computacionais do rastreamento ocular, mais especificamente, os registros sobre a variabilidade do dilatação pupilar.

É válido ressaltar que antes de iniciar a aquisição dos dados: o ambiente onde seria realizado o experimento foi devidamente preparado para a aquisição dos dados em condições de luz constante e em resposta a fluxos informativos comuns e não manipulados; foi realizada a calibração do dispositivo em todos os sujeitos participantes para evitar erros e ruídos nas medidas providas pelo dispositivo rastreador

móvel por causa da liberdade de movimentar a cabeça; e foram excluídos os indivíduos que apresentavam alguma disfunção visual que impossibilitava a utilização do rastreador ocular ou que pudessem interferir na integridade dos dados, mantendo somente os participantes que tinham uma visão normal.

Nesta abordagem, exclusivamente, o trecho do filme foi dividido em 13 cenas (bins) ao todo, cada uma com um tempo de duração diferente. A definição e transição entre os bins dá-se pela mudança de protagonista, ou seja, se um trecho inicia-se exibindo a Maggie, então é definido o início de um bin. Instantes depois, se o foco visual é alterado para o Frankie, é dito o fim do bin anterior e o início de um outro bin, e assim sucessivamente. A Tabela 5.1 sumariza os dados de cada um dos bins (b), quanto: ao protagonista (p); ao tempo de duração; ao número de dados coletados (n) e ao conteúdo apresentado visualmente e sonoramente.

b	p	Tempo (ms)	n	Conteúdo
1	Maggie	0 - 2.330	71	Maggie pede um favor ao Frankie.
2	Frankie	2.370 - 9.610	218	Frankie diz que faz o que Maggie quiser.
3	Maggie	9.650 - 14.320	141	Maggie pergunta se Frankie lembra o que o pai dela fez com Axel.
4	Frankie	14.350 - 18.960	139	Frankie observa Maggie com atenção.
5	Maggie	19.000 - 21.770	84	Maggie respira forte.
6	Frankie	21.800 - 29.380	228	Frankie pede para Maggie não pensar nessa opção (eutanásia).
7	Maggie	29.420 - 55.700	788	Maggie diz que não pode viver assim.
8	Frankie	55.730 - 61.510	174	Frankie observa Maggie pensativo.
9	Maggie	61.540 - 106.990	1362	Maggie fala sobre a sua vida e enfatiza seu pedido.
10	Frankie	107.020 - 111.300	129	Frankie observa Maggie com atenção.
11	Maggie	111.330 - 124.390	392	Maggie pede a Frankie que não deixem que tirem as conquistas dela.
12	Frankie	124.420 - 126.990	78	Frankie observa Maggie com atenção.
13	Maggie	127.030 - 134.410	222	Maggie pede a Frankie para que não a deixe no hospital naquelas condições.

Tabela 5.1. Resumo das 13 cenas do trecho filme.

Após o experimento, a decisão dos participantes foram verificadas e, por fim, o conjunto de dados foi dividido em dois, um com 27 instâncias, referentes aos participantes que aceitaram o pedido da paciente e praticariam a eutanásia, e outro com as 16 instâncias restantes, referentes aos que eram contra o pedido. Por fim, foi realizada a conversão dos dados brutos gerados em planilhas no formato CSV.

5.2.2 Extração de séries temporais

A extração de conhecimento a partir da análise de séries temporais é um grande desafio para os pesquisadores, devido às limitações das abordagens existentes e pela carência de métodos capazes de induzir conhecimento compreensível a partir de dados sequenciais [Maletzke et al., 2009].

Neste trabalho, os dados do comportamento ocular foram registados sequencialmente por meio do software ASL Plus. A cada 40 ms, durante todo o tempo de exibição do trecho do filme, o diâmetro pupilar foi capturado e armazenado. Desta forma, sucessivamente, foram obtidos os dados referentes ao tamanho da pupila de todos os participantes ao longo de determinado número de instantes distintos durante todo o experimento.

A ordem, ao se trabalhar com dados temporais, é de suma importância, devido a compor uma valiosa informação para o entendimento desses dados [Maletzke et al., 2009]. Para cada bin, utilizou algumas medidas típicas da estatística descritiva para sumarizar as dilatações das pupilas registradas. Foram utilizadas medidas, como, média, mediana, variância, valores mínimos e máximos. Entretanto, por causa do piscar, do tamanho dos olhos e do próprio dispositivo de rastreamento, é comum durante a coleta de dados oculares a obtenção de ruídos e valores faltantes [Klingner et al., 2008]. Esses valores interferem diretamente na tarefa de análise e comparação entre duas séries temporais [Maletzke et al., 2009]. Assim, com intuito de evitar o comprometimento na qualidade do conhecimento gerado a partir desses dados, na etapa de pré-processamento, os valores faltantes e ruídos foram ignorados. Após a análise, notou-se que os valores das médias e das medianas entre os sujeitos, para os respectivos bins, não possuíam discrepâncias, logo, optou-se pela utilização da medida da média, assim como Klingner et al. [2008]; Privitera et al. [2008]; Hossain & Yeasin [2014] e Muthumanickam et al. [2016] em seus respectivos trabalhos.

Os registros médios das variações do diâmetro pupilar do experimento foram obtidos da seguinte forma:

- após ser assistido pelos participantes, o trecho do filme foi dividido em 13 cenas (bins) distintas;
- para cada participante, foram gerados 13 vetores de valores pupilares referentes as cenas;
- os vetores dos participantes foram separados em dois grupos de acordo com a tomada de decisão (a favor e contra o pedido de eutanásia);

- para cada cena, foi calculado o vetor médio de cada grupo. O valor do vetor corresponde à média dos dados pupilares em pixels de todas as instâncias de um mesmo grupo em um mesmo instante.

Por fim, 26 vetores (13 de cada grupo) com as médias do diâmetro pupilar dos participantes ao longo do tempo foram criados para realizar a análise estatística das séries temporais. Portanto, após este processamento, cada bin definido previamente pode ser analisado comparando um vetor do grupo a favor com um vetor do grupo contra. A Figura 5.2 abaixo demonstra este procedimento.

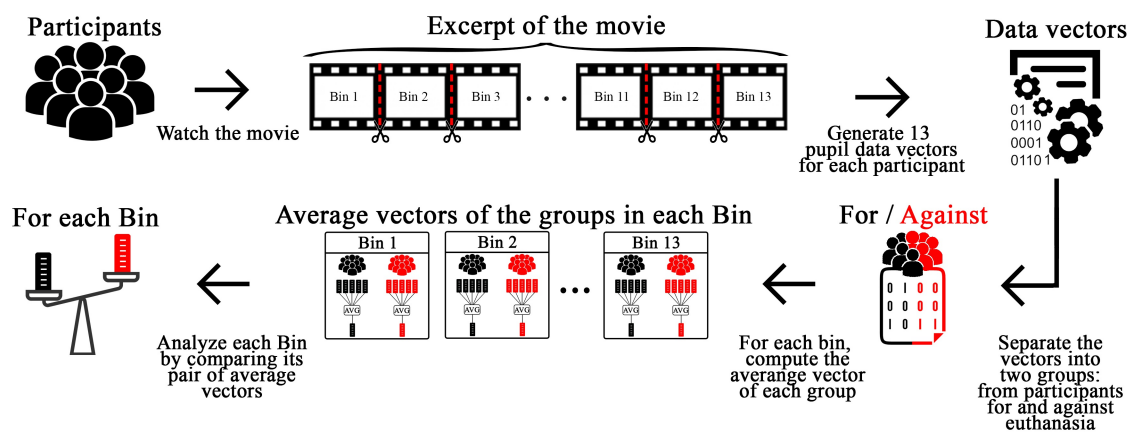


Figura 5.2. Procedimento para obtenção de dados pupilares.

5.2.3 Testes estatísticos

Em análise de dados, não se pode inferir em algo de forma confiável somente com base em experiências anteriores ou no julgamento dos olhos humanos [Marusteri & Bacarea, 2010]. Então, para saber se existem diferenças estatisticamente significativas entre os conjuntos de dados observados, é necessário a utilização de testes estatísticos, sendo eles, a base para inferência na maioria das pesquisas científicas [Nayak & Hazra, 2011].

Conforme explicado anteriormente na Seção 3.3, a inferência estatística é realizada por meio dos testes de hipóteses e a escolha do melhor teste para uma determinada investigação é baseada em alguns pontos importantes, como, tipo de dados usados, número de grupos a serem comparados, se os grupos possuem ou não relação, entre outros.

Neste trabalho, os dados utilizados são do tipo quantitativo e é verificado a possível existência de diferenças estatisticamente significativas em dois grupos não

pareados. Por meio do uso do teste de normalidade de Shapiro-Wilk¹, foi constatado que os dados não seguem uma distribuição normal (p -valor $< 0,05$). Essa informação corrobora com as seguintes observações:

- as respostas pupilares não são estacionárias por natureza, por causa de sua associação com a carga cognitiva gerada pelas atividades mentais [Murata & Iwase, 1998; Hossain & Yeasin, 2014];
- os dados biológicos frequentemente não seguem uma distribuição gaussiana com precisão [Marusteri & Bacarea, 2010].

Em razão dos dados possuírem uma distribuição não-normal, torna-se dispensável o uso do teste de homogeneidade da variância. Assim, após a verificação das pressuposições de normalidade e homogeneidade da variância, conclui-se que o teste de significância estatística mais recomendado para o experimento em questão, segundo a Figura 3.10, é o teste não-paramétrico de Mann-Whitney U [Callegari-Jacques, 2009; Nayak & Hazra, 2011; Torman et al., 2012].

5.2.4 Ferramentas de visualização e estatística

Nos últimos anos, as práticas estatísticas aliaram-se ao advento da computação. O uso de ferramentas computacionais voltadas à estatística e visualização de dados tornaram-se indispensáveis para a interpretação das informações obtidas por meio de grandes volumes de dados. Atualmente, existem diversos softwares e pacotes estatísticos genéricos que executam uma gama de funções. Os softwares mais comuns na literatura são: MATLAB, SPSS, SAS, PaSt e R.

Neste estudo, assim como Torman et al. [2012] e Nogueira & Pereira [2013] em seus respectivos trabalhos, recorreremos ao uso do software R. O R é uma ferramenta gratuita, multiplataforma, distribuída por meio da licença de software livre da GNU, que funciona como uma linguagem de programação e um ambiente para computação estatística e gráfica. Por meio do uso de bibliotecas com funções preestabelecidas, o R permite agregar ferramentas de exibição gráfica, de cálculos, de análise estatística e manipulação de dados em um ambiente de desenvolvimento [Ferreira & Cymrot, 2012].

Nos últimos anos, o software começou a ser aceito e bem requisitado por diversos pesquisadores de várias áreas por ser bem documentado, reproduzível e adaptável. Essas características fizeram com que se torna-se uma poderosa ferramenta

¹O teste de normalidade de Shapiro-Wilk usado nesta pesquisa foi executado por meio da função `shapiro.test` no R.

para análise de dados, capaz de realizar uma ampla variedade de funções [Ferreira & Cymrot, 2012; Henning et al., 2016]. Por esses motivos, o software R foi escolhido para realizar a análise dos dados por meio dos testes estatísticos e a exibição dos mesmos através das técnicas de visualização.

5.2.4.1 Gráficos exploratórios

Embora a análise estatística seja adequada para validação e geração de conclusões sobre os dados, as técnicas de visualização exploratória permitem a extração de plausíveis hipóteses, por meio da revelação e distinção de padrões, regularidades, anomalias e informações ocultas de interesse não conhecidos previamente nos dados [Kurzahls et al., 2015].

Mesmo não sendo desenvolvidos especificamente para esse fim, existem diversos gráficos que permitem a representação de conjunto de dados oculares e o ganho de discernimentos sobre eles [Blascheck et al., 2014]. Neste trabalho, optou-se pela escolha de dois tipos de gráficos que permitem a visualização dos dados ao longo do tempo e sua distribuição estatística.

O primeiro modelo gráfico escolhido foi o boxplot. O boxplot é uma técnica de visualização amplamente utilizada com intuito de aprimorar a compreensão sobre os dados quantitativos da amostra por meio de sua avaliação e síntese descritiva [Blascheck et al., 2014]. É um esquema alternativo ao histograma que provê um bom resumo visual dos principais aspectos de uma distribuição estatística, como, por exemplo, amplitude, simetria, variabilidade, valores típicos e, por fim, os valores discrepantes [Öztuna et al., 2006; Steen, 2012]. Os boxplots são extremamente úteis para a realização da comparação visual entre dois ou mais grupos por causa do uso das estatísticas mais robustas, como, a mediana e a distância interquartílica, ao invés das estatísticas mais sensíveis, como, a média e o desvio padrão, tendo, portanto, um alto grau de confiabilidade [Reis & Reis, 2002; Kurzahls et al., 2015].

O segundo modelo é conhecido como gráfico de linhas. Esta técnica gráfica permite estudar as características dos movimentos oculares representadas em grandes quantidades de dados ao longo de um período de tempo contínuo [Blascheck et al., 2014]. Os gráficos de linhas são ideais para observar visualmente o comportamento da pupila quando estímulos são apresentados, possibilitando a identificação de tendências, irregularidades e picos de amplitude da variável num instante de alteração brusca.

5.3 Resultados e discussão

Para investigar a relação entre o comportamento cognitivo refletido pela variabilidade do diâmetro pupilar com a tomada de decisão, o presente estudo propôs uma abordagem para análise que faz uso de técnicas de estatística e de visualização de dados que permitem a realização de comparações entre os dados e o apoio a hipóteses de estudo.

Inicialmente, serão apresentados os resultados² das análises estatísticas e logo em seguida, as conclusões referentes à exploração visual. As explicações de ambas as análises serão dadas tanto de forma geral quanto segmenta entre os bins. Por fim, avaliaremos o comportamento do diâmetro pupilar a fim de proporcionar uma melhor compreensão da relação entre essa variável e a tomada de decisão bioética.

5.3.1 Análise estatística

5.3.1.1 Análise geral

Como mencionado na Subseção 5.2.3, a verificação e validação do esforço cognitivo refletido por meio da pupila entre os dois grupos de participantes durante o experimento foram realizadas por meio do teste de hipótese estatística não-paramétrica de Mann Whitney U. As hipóteses definidas para o teste foram as seguintes:

- H_0 : as médias do diâmetro pupilar dos dois grupos são iguais;
- H_1 : as médias do diâmetro pupilar dos dois grupos são diferentes.

O teste³ mostrou diferenças significativas ao comparar as médias do diâmetro pupilar entre cada um dos grupos de participantes (p-valor < 0,05). Logo, isto significa que podemos rejeitar a H_0 .

5.3.1.2 Analisando as 13 cenas

Com o intuito de obter um detalhamento maior das diferenças dos dados oculares nas cenas individualmente, também optou-se em realizar uma análise estatística dos grupos por bins. A Tabela 5.2 exhibe os valores do teste de Mann Whitney U usado para os dois grupos nos bins separadamente.

²Todos os resultados gerados pelos experimentos computacionais foram obtidos pelo software R versão 3.2.5 32 bits em um notebook Intel Core i7-4500 @1.80GHz com 8 GB de RAM rodando o Windows 8.1 Pro 64 bits.

³O teste de hipótese estatística de Mann Whitney U usado nesta pesquisa foi executado por meio da função `wilcox.test` no R e com nível de confiança de 95%.

Bin	p-valor
1	0,000
2	0,000
3	0,004
4	0,000
5	0,006
6	0,080
7	0,000
8	0,561
9	0,000
10	0,000
11	0,000
12	0,000
13	0,231

Tabela 5.2. Resultados do teste de Mann Whitney U para os 13 bins.

Os resultados do teste apresentam diferenças significativas ao comparar 10 dos 13 bins dos grupos (p-valor $< 0,05$), rejeitando-se a H_0 de igualdade das médias. Contudo, não foram encontradas diferenças significativas quando se comparou os dados dos bins 6, 8 e 10 (p-valor $\geq 0,05$), aceitando, portanto, a H_0 . Logo, conclui-se que 76,92% dos valores médios presentes nos grupos durante os bins possuem diferenças estatisticamente significativas.

Essas informações não foram obtidas por meio da análise anterior, por causa da abstração que ocorre quando analisamos todos os dados de uma só vez, diferente de quando são analisados detalhadamente, pode-se encontrar informações ocultas que não foram percebidas antes. Esses resultados são fundamentais para encontrar os padrões relatados na próxima subseção.

5.3.2 Análise exploratória

5.3.2.1 Análise geral

No contexto da extração do conhecimento, após ser constatado pelas análises estatísticas a existência de diferenças significativas entre os dados, foi realizado um estudo exploratório por meio do uso dos boxplots e gráficos de linhas para proporcionar uma melhor compreensão a respeito desse comportamento. A Figura 5.3 ilustra o comportamento médio do diâmetro pupilar geral dos grupos de participantes ao longo do experimento.

A Subfigura 5.3a mostra um gráfico no formato boxplot, que possibilita evidenciar a mediana, os outliers e as respectivas variâncias de cada grupo. Primeiramente,

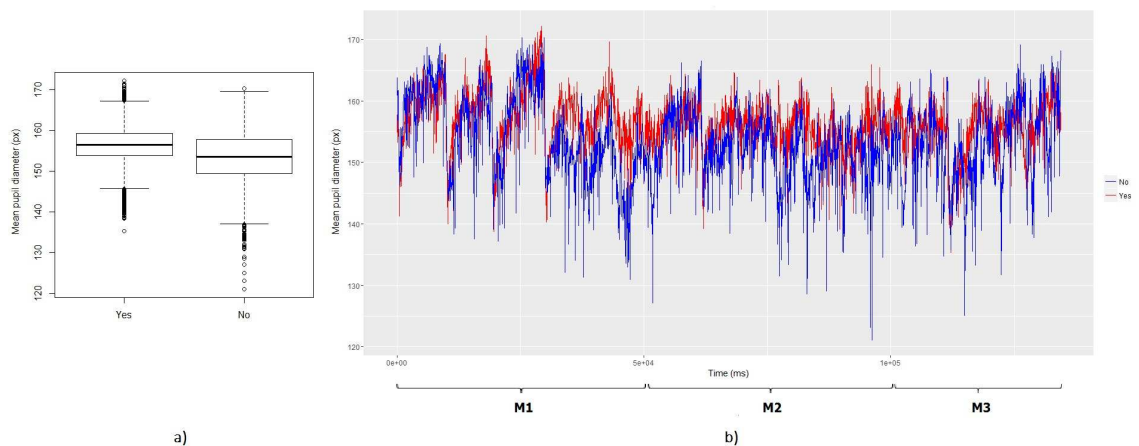


Figura 5.3. Comportamento médio do diâmetro pupilar geral dos grupos de participantes ao longo do experimento por meio de (a) boxplot e (b) gráfico de linhas.

observa-se que as linhas das medianas estão posicionadas praticamente no centro das caixas, dividindo-as em duas partes aproximadamente iguais. Isto significa que os dois grupos podem ser considerados simétricos [Reis & Reis, 2002; Pintilie, 2013]. A simetria e assimetria dos dados pupilares estão associadas ao desempenho de funções cognitivas [Cicco et al., 2016]. Assim, quanto menos assimétricos forem os dados, melhor a performance cognitiva dos indivíduos durante uma ação. Visto isso, alega-se então que não houve déficits cognitivos para ambos os grupos e que o aspecto de simetria não foi o responsável por apresentar as diferenças significativas.

Outro aspecto exibido pelos boxplots são os outliers. A interpretação desses valores torna-se um pouco mais complicada em relação a outros aspectos, pois a sua definição varia de acordo com a finalidade e a situação da pesquisa. Outliers são definidos como quaisquer pontos acima do limite superior ou abaixo do limite inferior e, nem sempre indicam erro de leitura, mas sim, valores incomuns [Steen, 2012]. Como pode ser visto na Subfigura 5.3a, os valores ditos como outliers não estão tão distantes dos limites definidos nos grupos, além disso, grande parte dos outliers presentes no grupo Sim (Yes) estão dentro dos intervalos de valores do grupo Não (No). As mudanças sutis no tamanho da pupila em curto prazo estão relacionadas a uma maior necessidade de processamento de informação [Klingner et al., 2008]. Em razão disso e de ter realizado o pré-processamento dos dados (Subseção 5.2.2), esses valores foram considerados como aceitáveis e mantidos nos grupos.

O aspecto de dispersão pode ser facilmente verificado por meio da altura das caixas, uma vez que os boxplots exibem o intervalo completo de variação dos dados

[Steen, 2012]. O grupo Sim além de apresentar maiores médias do diâmetro pupilar, também exibe uma menor variabilidade desses valores. O grupo Não em compensação, apresenta menores médias e uma maior variabilidade. Durante a realização de uma tarefa cognitiva, acontece involuntariamente uma alteração no tamanho da pupila [Kahneman & Beatty, 1966; Beatty, 1982]. Deste modo, conforme é solicitado um maior grau de esforço cognitivo, as pupilas dilatam ligeiramente (ativação do SNS) e, logo ao completar a tarefa, acontece a diminuição dos estímulos e as pupilas se contraem progressivamente até voltar ao seu tamanho anterior (inibição da estimulação simpática por meio da ativação do SNP) [Klingner et al., 2008; Privitera et al., 2008; Hossain & Yeasin, 2014].

Uma dilatação elevada da pupila indica um alto nível de esforço cognitivo [Kahneman & Beatty, 1966; Beatty, 1982]. Mesmo que o grupo Sim apresente valores mais altos do diâmetro pupilar, isso não significa que os indivíduos desse grupo tiveram um maior esforço cognitivo. Porque, há um maior grau de esforço cognitivo quando há fortes dilatações seguidas por contrações e não ao contrário ou uma permanência em um determinado valor alto. Segundo a Subfigura 5.3b, no decorrer do experimento, as alterações dos valores do diâmetro da pupila desse grupo não foram tão significativas. Pelo gráfico de linhas, verificou-se que a partir do M2 até um pouco depois do início de M3, uma variabilidade não muito abrupta em comparação ao grupo Não, acreditando assim que pela ocorrência de menores dilatações seguidas de contrações, houve um menor esforço cognitivo. Na Figura 5.3, está evidente que o grupo Não teve variabilidades mais longas, isto é, dilatações e contrações mais fortes durante vários momentos do experimento. Em virtude disso, inferi-se que esse grupo de indivíduos requereram uma demanda cognitiva mais alta para tomar a decisão.

Embora a caixa do boxplot do grupo Sim apresenta-se acima da caixa do grupo Não, existe uma interseção entre elas, ou seja, valores comuns. Na Subfigura 5.3b, diferente da Subfigura 5.3a, percebe-se nitidamente os momentos que apresentam as diferenças e similaridades entre as distribuições dos dois grupos. Os testes estatísticos da Subseção 5.3.1 comprovam a discrepância do comportamento do diâmetro pupilar dos grupos investigados. Neste contexto, demonstra-se que os grupos de participantes tiveram comportamentos pupilares distintos, evidenciando assim, uma correlação entre o comportamento do diâmetro da pupila e o processo decisório. Portanto, a relação dessa variável com a tomada de decisão bioética mostra um papel significativo no caso estudado.

5.3.2.2 Analisando as 13 cenas

Após a discussão e interpretação dos resultados da análise geral, foi realizado novamente uma análise segmentada dos 13 bins para possibilitar um melhor entendimento dos padrões encontrados. Pelo fato dos aspectos de simetria e outliers não terem demonstrado diferenças significativas nas observações anteriores, na discussão a seguir, serão explorados apenas os aspectos de dispersão e interseção. As Figuras 5.4 e 5.5, diferente da Figura 5.3, ilustram de maneira detalhada, o comportamento do diâmetro pupilar médio dos grupos em cada um dos 13 bins individualmente.

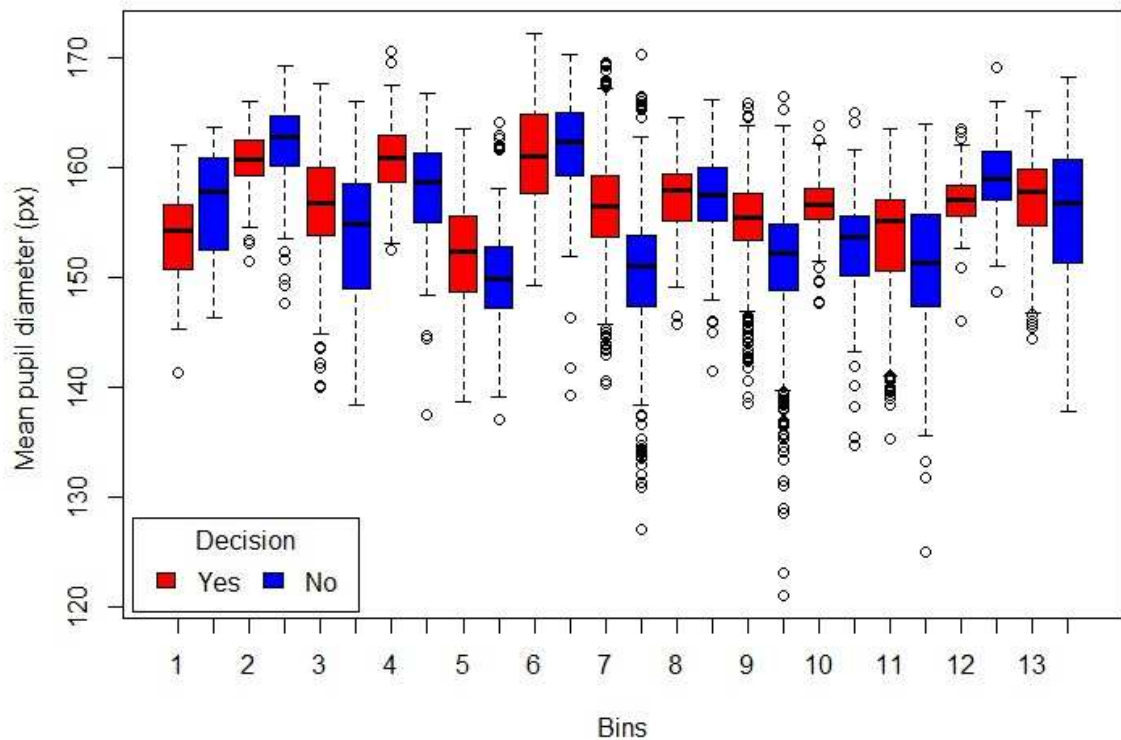


Figura 5.4. Múltiplos boxplots dos dados do diâmetro pupilar dos grupos de participantes durante os 13 bins.

O profissional da saúde, ao deparar-se em uma situação clínica que envolve a vida e o viver, deve levar em conta para sua tomada de decisão tanto os preceitos da ética, da legalidade e da moral impostos a ele por meio do juramento feito ao código de ética dos profissionais de enfermagem (decisão mais racional), quanto o senso de humanidade e o respeito à vontade do próprio paciente difundidos pela reflexão da bioética (decisão mais emocional) [Siqueira-Batista et al., 2015; Goldim, 2006]. Entretanto, em casos delicados que possuem um alto nível de complexidade e de

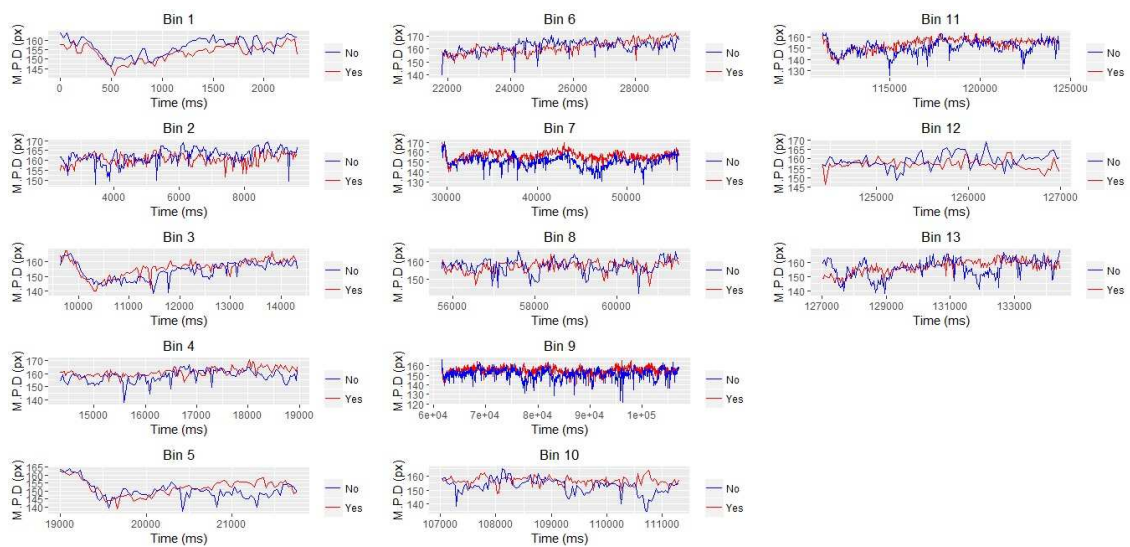


Figura 5.5. Múltiplos gráficos de linhas de séries temporais do comportamento do diâmetro pupilar dos grupos participantes durante os 13 bins.

juízo, como, a eutanásia, tomar uma decisão que concilie de modo colaborativo as vias racionais e emocionais é extremamente difícil.

Observa-se que 11 dos 13 bins do grupo Não se sobressaíram em relação ao grupo Sim, com exceção dos bins 5 e 6, pois apresentam uma maior variabilidade dos valores. No contexto da tomada de decisão clínica, uma explicação plausível para o alto esforço cognitivo gerado pelas constantes variações do diâmetro pupilar do grupo Não seja devido o nível de dificuldade apresentado pela tarefa, o que acarreta na existência de uma situação psicologicamente desconfortável, conhecida como dissonância cognitiva [Hossain & Yeasin, 2014]. A dissonância cognitiva ocorre quando uma pessoa encontra-se em inconsistência entre duas opiniões, esforçando-se para manter a coerência entre seus pensamentos [Festinger, 1962]. A constante variação do diâmetro pupilar durante todo o experimento evidencia o processamento de conflito na escolha da decisão [Querino et al., 2015]. Em virtude disso, suponha-se que a decisão final desses participantes tenha sido “não” por terem tido uma ponderação muito intensa para se contrapor a vontade do paciente, isto é, por terem pensado mais nas consequências que sua decisão acarretaria para ele e para sua profissão pelo desrespeito ao código de ética (neste caso, uma decisão mais racional).

Em relação a dispersão do grupo Sim, nota-se também uma alta variabilidade entre os bins 1 e 7 (Figura 5.4), onde pressupôs-se ser o momento do experimento em que os participantes desse grupo também tiveram dúvidas sobre qual decisão escolher. Todavia, após o bin 7, constatou-se que os dados variavam em uma pro-

porção bem menor, mantendo, portanto, um diâmetro pupilar mais concentrado, principalmente entre os intervalos de 150 e 160 pixels (Figura 5.5). Segundo a visão de Querino et al. [2015], o tamanho das pupilas visam aumentar nos estágios iniciais de uma tarefa devido essas etapas serem consideradas mais sensíveis às dificuldades apresentadas a um indivíduo, e tendem a diminuir consideravelmente nos estágios finais por causa da redução da exigência da tarefa, pelo fato do indivíduo ter se contentado com a decisão a ser tomada. O comportamento do grupo Sim se encaixa nessa perspectiva, visto isso, acredita-se que no bin 7, os participantes do grupo já tenham escolhido sua decisão final, o que justifica a diminuição do processamento de conflito juntamente com o do esforço cognitivo, representados por meio da baixa variabilidade dos valores pupilares médios a partir desse bin.

Sobre a questão clínica do grupo Sim, pressupõe-se que a tomada de decisão tenha sido influenciada mais pela emoção do que pela razão, visto que o conteúdo emocional presente em cenas agradáveis ou repulsivas, neste caso, a segunda opção, evoca estímulos associados ao esforço cognitivo [Privitera et al., 2008]. Esse esforço refletido pela TEPR é nítido nas primeiras sete cenas do experimento (Tabela 5.1), principalmente entre as cenas de número 5 e 6, onde se nota a maior dilatação antes da tomada de decisão. Portanto, implica-se que os participantes do grupo Sim sensibilizaram-se, por meio das cenas visualizadas, do sofrimento da paciente e por impulso (demonstrado pela redução da variabilidade do diâmetro pupilar após a maior dilatação do bin 5 para o 6), resolveram atender ao pedido, abrindo mão do juramento feito ao código de ética em troca do senso de humanidade e do respeito à vontade da própria paciente. Além disso, quando uma situação como a da Maggie, ou seja, uma boxeadora que estava indo bem na carreira e por causa de uma luta injusta teve sua vida arruinada, é vista com um movimento afetivo, faz sentido a pessoa na qual foi submetida ao pedido de eutanásia, querer ajudá-la nesse processo de encerrar a própria vida, pois, dependendo da visão de mundo e existência, algumas pessoas não gostariam de ser submetidas ao mesmo sofrimento.

Por fim, na Subseção 5.3.1.2, foi verificado por meio do teste de hipótese que três dos 13 bins, não possuem diferenças significativas. Os boxplots da Figura 5.4 ratificam esses resultados, ilustrando que o boxplot de um dos grupos dos bins 6, 8 e 13 realmente possui todos os valores em comum ao boxplot do grupo oposto. Também é possível observar que os bins 7 e 10 quase não se intercedem, ou seja, são os bins que mais apresentam diferenças. Os gráficos de linhas desses dois bins na Figura 5.5 corroboram com essa informação, demonstrando claramente as diferenças de comportamento.

Os resultados apresentados nesta investigação além de complementar, corrobo-

raram com algumas das conclusões alcançadas na abordagem anterior (Capítulo 4) que mostram a conexão entre os aspectos cognitivos refletidos pelo comportamento do diâmetro pupilar e a tomada de decisão.

5.4 Conclusão

Esta investigação consiste na análise da compreensão dos movimentos oculares humanos, mais especificamente, dos dados pupilares, durante tomadas de decisões bioéticas na esfera clínica por meio do uso de métodos de estatística e visualização. O estudo visa facilitar as necessidades dos pesquisadores em obter possíveis insights que envolvam a atenção visual e os processos cognitivos durante a realização de tarefas, visto que as ferramentas e métodos já existentes são limitados para satisfazer plenamente tais necessidades.

A abordagem proposta foi composta pelos métodos visuais conhecidos como boxplot e gráfico de linhas, além do teste de hipóteses estatística de Mann-Whitney U. Um estudo de caso com a abordagem foi realizado sobre um experimento de neurociências composto por dados pupilares de 43 estudantes voluntários do curso de Enfermagem de uma universidade brasileira, obtidos por meio de um dispositivo de rastreamento ocular móvel, a fim de analisar o comportamento da pupila sobre efeitos da carga cognitiva e entender sua relação com as tomadas de decisões.

Os resultados alcançados revelaram diferenças de comportamento do olhar dos indivíduos que optariam a favor e contra a participação da prática da eutanásia ao ver a cena cinematográfica apresentada no experimento. A análise estatística comprovou diferenças significativas, enquanto a análise exploratória demonstrou, de forma mais detalhada, resultados mais perspicazes a respeito das características do comportamento pupilar, como, por exemplo: simetria; dispersão; valores típicos e atípicos, possibilitando assim, a identificação de padrões nos dados analisados. Com os padrões encontrados, foi possível associar os aspectos cognitivos e pressupor motivos pela tomada de decisão.

Capítulo 6

Conclusões

Este trabalho propôs um framework de mineração de dados oculares, composto por duas abordagens baseadas em metodologias recomendadas no estado da arte, sendo em algoritmos de aprendizagem de máquina e em técnicas de estatística e visualização de dados, para tentar inferir em padrões que possibilitem associar processos cognitivos refletidos pela atenção visual com a tomada de decisão em bioética clínica.

Os dados oculares do conjunto de dados foram gerados a partir da apreciação de um trabalho cinematográfico com o MET sobre uma questão ética muito relevante e controversa da prática clínica, a eutanásia. A metodologia de investigação da participação da atenção visual na bioética não havia sido até então encontrada na literatura, sendo, portanto, desenvolvida pelos pesquisadores da UFV.

Na primeira abordagem, após a seleção dos dados e dos atributos relevantes, foram aplicados algoritmos de aprendizagem supervisionada, para construir modelos preditivos usando dados de rastreamento ocular. Os experimentos estatísticos mostraram que o modelo preditivo resultante do algoritmo MLP levou ao melhor desempenho comparado aos outros algoritmos testados, apresentando uma acurácia de 90,7% e uma AUC média de 0,903. O conhecimento interessante (padrões e regras) para a tomada de decisão bioética estudada foi extraído usando simulações com modelos MLP e inspecionando as regras de decisão tomadas.

Na segunda abordagem, após também a fase de seleção e pré-processamento dos dados pupilares, foram extraídas as séries temporais dos grupos de participantes que seriam analisadas e comparadas. Os experimentos estatísticos realizados sobre o conjunto de dados por meio do teste de hipótese de Mann-Whitney U demonstraram que há diferenças estatisticamente significativas entre os dois grupos de participantes. Deste modo, pode-se confirmar que as decisões dos grupos são influenciadas por fatores distintos. Posteriormente, os métodos visuais de boxplot e do gráfico

de linhas forneceram um detalhamento maior sobre o comportamento do diâmetro pupilar, o que possibilitou identificar hipóteses plausíveis baseadas nos padrões de desempenho do processamento cognitivo durante tarefas decisórias.

Os resultados forneceram novos insights e podem ser úteis para pesquisadores que desejam investigar a relação dos processos cognitivos com tomadas de decisões e atenção visual por terem demonstrado uma forma simples que facilita a extração de informações, o surgimento de hipóteses e a interpretações dos resultados provenientes de dados oculares.

Este estudo envolveu várias áreas do conhecimento, o que contribuiu para revelar o papel da atenção visual no processo de tomada de decisão, verificando até que ponto a visão é um fator determinante, particularmente, em bioética clínica, quando se trata de questões sobre o fim de vida.

A ideia abordada nesta dissertação é nova e relevante. Mesmo utilizando métodos clássicos de aprendizagem de máquina, visualização e estatística, eles são usados em um novo contexto (bioética) que ainda não havia sido explorado na literatura. Por esse motivo, não foi possível fazer comparações com qualquer outro trabalho conhecido até o momento.

Como trabalhos futuros, pretende-se: (i) utilizar além dos dados oculares, outros dados fisiológicos, como, dados cardiovasculares ou de eletroencefalograma para melhorar a abrangência dos resultados; (ii) integrar ambas as abordagens em um sistema computacional genérico que possibilite a investigação de outras bases de dados do complexo sistema de processamento de informações humanas na tomada de decisões em diferentes cenários.

Por fim, a primeira abordagem deste trabalho resultou em uma publicação no International Conference on Computational Science 2017 (ICCS 2017). A segunda abordagem gerou um artigo que está em sua fase final e que futuramente será submetido a uma conferência ou revista. O artigo publicado está disponível no Apêndice A.

Referências Bibliográficas

- Acharya, U. R. & Yu, W. (2010). Data mining techniques in medical informatics. *The Open Medical Informatics Journal*, 4:21--22.
- Alickovic, E. & Subasi, A. (2016). Medical decision support system for diagnosis of heart arrhythmia using dwt and random forests classifier. *Journal of Medical Systems*, 40(4):1--12.
- Andrá, C.; Lindström, P.; Arzarello, F.; Holmqvist, K.; Robutti, O. & Sabena, C. (2015). Reading mathematics representations: An eye-tracking study. *International Journal of Science and Mathematics Education*, 13(2):237--259.
- Andrienko, G.; Andrienko, N.; Burch, M. & Weiskopf, D. (2012). Visual analytics methodology for eye movement studies. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2889--2898.
- Araujo, R. R. d. & Carreiro, L. R. R. (2009). Orientação voluntária e automática da atenção e indicadores de desatenção e hiperatividade em adultos. *Avaliação Psicológica*, 8(3):325--336.
- Barreto, A. M. (2012). Eye tracking como método de investigação aplicado às ciências da comunicação. *Revista Comunicando*, 1(1):168--186.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2):276--292.
- Beatty, J. & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of Psychophysiology*, 2:142--162.
- Bednarik, R. (2007). *Methods to analyze visual attention strategies: Applications in the studies of programming*. University of Joensuu.

- Berghs, M.; De Casterle, B. D. & Gastmans, C. (2005). The complexity of nurses' attitudes toward euthanasia: A review of the literature. *Journal of Medical Ethics*, 31(8):441--446.
- Bernstam, E. V.; Smith, J. W. & Johnson, T. R. (2010). What is biomedical informatics?. *Journal of Biomedical Informatics*, 43(1):104--110.
- Blascheck, T.; Kurzhals, K.; Raschke, M.; Burch, M.; Weiskopf, D. & Ertl, T. (2014). State-of-the-art of visualization for eye tracking data. In *Proceedings of EuroVis*, volume 2014, pp. 1--20.
- Blum, A. L. & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1):245--271.
- Bojko, A. A. (2009). Informative or misleading? heatmaps deconstructed. In *International Conference on Human-Computer Interaction*, pp. 30--39. Springer.
- Bottou, L. & Lin, C.-J. (2007). Support vector machine solvers. *Large Scale Kernel Machines*, pp. 301--320.
- Box, G. E.; Hunter, J. S. & Hunter, W. G. (2005). *Statistics for experimenters: Design, innovation, and discovery*, volume 2. Wiley-Interscience New York.
- Boxtel, A. & Jessurun, M. (1993). Amplitude and bilateral coherency of facial and jaw-elevator emg activity as an index of effort during a two-choice serial reaction task. *Psychophysiology*, 30(6):589--604.
- Braga, A. d. P.; Ferreira, A. C. P. d. L. & Ludermir, T. B. (2007). *Redes neurais artificiais: Teoria e aplicações*. LTC Editora, Rio de Janeiro.
- Brosch, T.; Scherer, K. R.; Grandjean, D. M. & Sander, D. (2013). The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, 143:1--10.
- Callegari-Jacques, S. M. (2009). *Bioestatística: Princípios e aplicações*. Artmed Editora.
- Camilo, C. O. & Silva, J. C. d. (2009). Mineração de dados: Conceitos, tarefas, métodos e ferramentas. *Universidade Federal de Goiás*, pp. 1--29.
- Cardoso, F. d. S. L. & Costa, I. T. d. (2013). O comportamento pupilar como indicativo de conhecimento específico do jogador de futebol. *Revista Mineira de Educação Física*, Edição Especial(9):1087--1094.

- Carvalho, D. R.; Moser, A. D.; da Silva, V. A. & Dallagassa, M. R. (2012). Mineração de dados aplicada à fisioterapia. *Fisioterapia em Movimento*, 25(3):595--605.
- Cerqueira, F. R. (2010). *New computational strategies for tandem mass spectrometry data analysis in phosphoproteomics*. PhD thesis, University for Health Sciences, Medical Informatics and Technology, Österreich.
- Cerqueira, F. R.; Ferreira, T. G.; de Paiva Oliveira, A.; Augusto, D. A.; Krempser, E.; Barbosa, H. J. C.; Franceschini, S. d. C. C.; de Freitas, B. A. C.; Gomes, A. P. & Siqueira-Batista, R. (2014). NICeSim: An open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making. *Artificial Intelligence in Medicine*, 62(3):193--201.
- Chen, H.; Fuller, S. S.; Friedman, C. & Hersh, W. (2005). Knowledge management, data mining, and text mining in medical informatics. *Medical Informatics*, pp. 3--33.
- Chen, Y.; Argentinis, J. E. & Weber, G. (2016). IBM Watson: How cognitive computing can be applied to big data challenges in life sciences research. *Clinical therapeutics*, 38(4):688--701.
- Chiavenato, I. (2004). *Introdução à Teoria da Gestão*. Elsevier-Campos, São Paulo.
- Chun, M. M. & Wolfe, J. M. (2001). Chapter nine visual attention.
- Cicco, V. d.; Barresi, M.; Fantozzi, M. P. T.; Cataldo, E.; Parisi, V. & Manzoni, D. (2016). Oral implant-prostheses: New teeth for a brighter brain. *PloS One*, 11(2):1--21.
- Coco, M. I. (2009). The statistical challenge of scan-path analysis. In *2009 2nd Conference on Human System Interactions*, pp. 372--375. IEEE.
- Collen, M. F. (1986). Origins of medical informatics. *Western Journal of Medicine*, 145(6):778--785.
- Critchley, H. D.; Eccles, J. & Garfinkel, S. N. (2013). Interaction between cognition, emotion, and the autonomic nervous system. *Handb Clin Neurol*, 117:59--77.
- Crockett, D. & Eliason, B. (2014). What is data mining in healthcare?. *HealthCatalyst*, pp. 1--10.
- Denning, P. J. (2005). Is computer science science?. *Communications of the ACM*, 48(4):27--31.

- Dodge, R. E. F. (2009). Eutanásia-aspectos jurídicos. *Revista Bioética*, 7(1):1--7.
- Eastwood, C. & Hagglund, K. J. (2005). Million dollar baby. *PsycCritiques*, 50(36).
- Evans, D. A. & Patel, V. L. (2013). *Advanced models of cognition for medical training and practice*, volume 97. Springer Science & Business Media.
- Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861--874.
- Fayyad, U.; Piatetsky-Shapiro, G. & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37--54.
- Felix, Z. C.; Costa, S. F. G. d.; Alves, A. M. P. d. M.; Andrade, C. G. d.; Duarte, M. C. S. & Brito, F. M. d. (2013). Eutanásia, distanásia e ortotanásia: Revisão integrativa da literatura. *Ciência & Saúde Coletiva*, 18(9):2733--2746.
- Ferreira, D. S. & Cymrot, R. (2012). *Uso do software R no tratamento estatístico de dados na engenharia*. Congresso de Iniciação Científica do INATEL - INCITEL, Santa Rita do Sapucaí.
- Festinger, L. (1962). *A theory of cognitive dissonance*, volume 2. Stanford university press.
- Figueiredo, A. M. d. (2011). Bioética clínica e sua prática. *Revista Bioética*, 19(2).
- Fitts, P. M.; Jones, R. E. & Milton, J. L. (2005). Eye movements of aircraft pilots during instrument-landing approaches. *Ergonomics: Psychological Mechanisms and Models in Ergonomics*, 3:56.
- França, G. V. d. (2009). Eutanásia: Um enfoque ético-político. *Revista Bioética*, 7(1).
- Galgani, F.; Sun, Y.; Lanzi, P. L. & Leigh, J. (2009). Automatic analysis of eye tracking data for medical diagnosis. In *Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on*, pp. 195--202. IEEE.
- Geurts, P.; Irrthum, A. & Wehenkel, L. (2009). Supervised learning with decision tree-based methods in computational and systems biology. *Molecular Biosystems*, 5(12):1593--1605.
- Ghasemi, A. & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2):486--489.

- Goldim, J. R. (2006). Bioética: Origem e complexidade. *Revista Hospital de Clínicas de Porto Alegre*, 26(2):86--92.
- Görgen, K. (2010). Combining eyetracking and eeg. *Publications of the Institute of Cognitive Science*, 15.
- Granholm, E. & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, 52(1):1--6.
- Granka, L. A.; Joachims, T. & Gay, G. (2004). Eye-tracking analysis of user behavior in www search. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 478--479. ACM.
- Gutnik, L. A.; Hakimzada, A. F.; Yoskowitz, N. A. & Patel, V. L. (2006). The role of emotion in decision-making: A cognitive neuroeconomic approach towards understanding sexual risk behavior. *Journal of Biomedical Informatics*, 39(6):720--736.
- Haladjian, H. H. & Montemayor, C. (2015). On the evolution of conscious attention. *Psychonomic Bulletin & Review*, 22(3):595--613.
- Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P. & Witten, I. H. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11(1):10--18.
- Hall, M. A. (1998). *Correlation-based feature subset selection for machine learning*. PhD thesis, University of Waikato, Hamilton, New Zealand.
- Han, J.; Pei, J. & Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier.
- Haykin, S. (1999). *Neural network: A comprehensive foundation*. Prentice-Hall, New Jersey.
- Hedar, A.-R.; Wang, J. & Fukushima, M. (2008). Tabu search for attribute reduction in rough set theory. *Soft Computing*, 12(9):909--918.
- Henning, E.; Ramos, M. S.; de Aguiar, R.; Siple, I. Z. & dos Santos, L. M. (2016). Para além da computação estatística: O uso do ambiente R para o ensino de métodos numéricos. *RENOTE*, 14(1):1--10.
- Hess, E. (1975). The tell-tale eye.

- Hess, E. H. & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143(3611):1190--1192.
- High, R. (2012). The era of cognitive systems: An inside look at ibm watson and how it works. *IBM Corporation, Redbooks*.
- Hollander, Z.; Chen, V.; Sidhu, K.; Lin, D.; Ng, R. T.; Balshaw, R.; Cohen-Freue, G. V.; Ignaszewski, A.; Imai, C.; Kaan, A. et al. (2013). Predicting acute cardiac rejection from donor heart and pre-transplant recipient blood gene expression. *The Journal of Heart and Lung Transplantation*, 32(2):259--265.
- Hoskin, T. (2012). Parametric and nonparametric: Demystifying the terms. In *Mayo Clinic*, pp. 1--5.
- Hossain, G. & Yeasin, M. (2014). Understanding effects of cognitive load from pupillary responses using Hilbert analytic phase. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 375--380.
- Hssina, B.; Merbouha, A.; Ezzikouri, H. & Erritali, M. (2014). A comparative study of decision tree ID3 and C4. 5. *International Journal of Advanced Computer Science and Applications*, 4(2):13--19.
- Inza, I.; Calvo, B.; Armañanzas, R.; Bengoetxea, E.; Larrañaga, P. & Lozano, J. A. (2010). Machine learning: An indispensable tool in bioinformatics. *Bioinformatics Methods in Clinical Research*, pp. 25--48.
- Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*, 217:105--133.
- John, G. H. & Langley, P. (1995). Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pp. 338--345. Morgan Kaufmann Publishers Inc.
- Junqueira, C. R. (2011). Bioética: Conceito, fundamentação e princípios. *São Paulo: Universidade Federal de São Paulo*.
- Kahneman, D. & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154(3756):1583--1585.

- Kim, Y.; Street, W. N. & Menczer, F. (2000). Feature selection in unsupervised learning via evolutionary search. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 365--369. ACM.
- Klingner, J.; Kumar, R. & Hanrahan, P. (2008). Measuring the task-evoked pupillary response with a remote eye tracker. In *Proceedings of the 2008 symposium on eye tracking research & applications*, pp. 69--72. ACM.
- Klingner, J. M. (2010). *Measuring cognitive load during visual tasks by combining pupillometry and eye tracking*. PhD thesis, Stanford University.
- Kohavi, R. & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, 97(1):273--324.
- Kohavi, R. & Provost, F. (1998). Glossary of terms. *Machine Learning*, 30(2-3):271--274.
- Krenz, W.; Robin, M.; Barez, S. & Stark, L. (1985). Neurophysiological model of the normal and abnormal human pupil. *IEEE Transactions on Biomedical Engineering*, (10):817--825.
- Kübler, T.; Eivazi, S. & Kasneci, E. (2015). Automated visual scanpath analysis reveals the expertise level of micro-neurosurgeons. In *MICCAI Workshop on Interventional Microscopy*, pp. 1--8.
- KurzHALS, K.; Burch, M.; Blascheck, T.; Andrienko, G.; Andrienko, N. & Weiskopf, D. (2015). A task-based view on the visual analysis of eye-tracking data. In *Workshop on Eye Tracking and Visualization*, pp. 3--22. Springer.
- Lavrač, N. (1999). Selected techniques for data mining in medicine. *Artificial Intelligence in Medicine*, 16(1):3--23.
- Lepargneur, H. (2009). Bioética da eutanásia argumentos éticos em rono da eutanásia. *Revista Bioética*, 7(1).
- Ling, C. X. & Sheng, V. S. (2008). Cost-sensitive learning and the class imbalance problem. *Encyclopedia of Machine Learning*, pp. 1--8.
- Loewenfeld, I. E. & Lowenstein, O. (1993). *The pupil: Anatomy, physiology, and clinical applications*, volume 2. Wiley-Blackwell.

- López, F. G.; Torres, M. G.; Batista, B. M.; Pérez, J. A. M. & Moreno-Vega, J. M. (2006). Solving feature subset selection problem by a parallel scatter search. *European Journal of Operational Research*, 169(2):477--489.
- Lupu, R. G. & Ungureanu, F. (2013). A survey of eye tracking methods and applications. *Bul Inst Polit Iasi*, pp. 71--86.
- Maletzke, A. G.; Batista, G. E.; Lee, H. D.; Wu, F. C. & Itaipu-PTI, P. T. (2009). Mineração de séries temporais por meio da extração de características e da identificação de motivos. In *Anais do VII Encontro Nacional de Inteligência Artificial, XXIX Congresso da Sociedade Brasileira de Computação*, pp. 637--646.
- Mansour, J. K. & Flowe, H. D. (2010). Eye tracking and eyewitness memory. In *Forensic Update*, pp. 11--15. British Psychological Society.
- Marusteri, M. & Bacarea, V. (2010). Comparing groups for statistical differences: How to choose the right statistical test?. *Biochemia Medica*, 20(1):15--32.
- McLaren, B. M. (2006). Computational models of ethical reasoning: Challenges, initial steps, and future directions. *IEEE Intelligent Systems*, 21(4):29--37.
- Metz, C. E. (1978). Basic principles of roc analysis. In *Seminars in nuclear medicine*, volume 8, pp. 283--298. Elsevier.
- Michalski, R. S.; Carbonell, J. G. & Mitchell, T. M. (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.
- Milliken, G. A. & Johnson, D. E. (2009). *Analysis of messy data volume 1: Designed experiments*, volume 1. CRC Press.
- Murata, A. & Iwase, H. (1998). Evaluation of mental workload by fluctuation analysis of pupil area. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE*, volume 6, pp. 3094--3097. IEEE.
- Musen, M. A. et al. (2002). Medical informatics: Searching for underlying components. *Methods of Information in Medicine*, 41(1):12--19.
- Muthumanickam, P. K.; Forsell, C.; Vrotsou, K.; Johansson, J. & Cooper, M. (2016). Supporting exploration of eye tracking data: Identifying changing behaviour over long durations. In *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pp. 70--77. ACM.

- Nakayama, M. & Shimizu, Y. (2004). Frequency analysis of task evoked pupillary response and eye-movement. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, pp. 71--76. ACM.
- Nayak, B. K. & Hazra, A. (2011). How to choose the right statistical test?. *Indian Journal of Ophthalmology*, 59(2):85--86.
- Nogueira, D. A. & Pereira, G. M. (2013). Desempenho de testes para homogeneidade de variâncias em delineamentos inteiramente casualizados. *Sigmae*, 2(1):7--22.
- Öztuna, D.; Elhan, A. H. & Tüccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences*, 36(3):171--176.
- Pärnamets, P.; Johansson, P.; Hall, L.; Balkenius, C.; Spivey, M. J. & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences*, 112(13):4170--4175.
- Patel, V. L. & Kannampallil, T. G. (2015). Cognitive informatics in biomedicine and healthcare. *Journal of Biomedical Informatics*, 53:3--14.
- Patel, V. L. & Kaufman, D. R. (1998). Medical informatics and the science of cognition. *Journal of the American Medical Informatics Association*, 5(6):493--502.
- Pereira, C. U.; Campos, G. J. d. L.; Carvalho, R. W. F. d.; Gusmão, L. C. B. d. et al. (2012). Nervo oculomotor: Anatomia, fisiologia e clínica. *Revista de Cirurgia e Traumatologia Buco-maxilo-facial*, 12(2):93--104.
- Pintilie, A. C. (2013). Statistical analysis and comparative visualization of cellular particle-based simulations. Master's thesis, University of Stuttgart, Stuttgart.
- Plöchl, M.; Ossandón, J. P. & König, P. (2012). Combining eeg and eye tracking: Identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Frontiers in Human Neuroscience*, 6:278.
- Powers, D. M. (2011). Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2:37--63.
- Primak, F. V. (2008). *Decisões com BI (Business Intelligence)*. Ciência Moderna Ltda, Rio de Janeiro.

- Privitera, C. M.; Renninger, L. W.; Carney, T.; Klein, S. & Aguilar, M. (2008). The pupil dilation response to visual detection. In *Electronic imaging 2008*, pp. 1--11. International Society for Optics and Photonics.
- Querino, E.; dos Santos, L.; Ginani, G.; Nicolau, E.; Miranda, D.; Romano-Silva, M. & Malloy-Diniz, L. (2015). Cognitive effort and pupil dilation in controlled and automatic processes. *Translational Neuroscience*, 6(1):168--173.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1):81--106.
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- Radbruch, L.; Leget, C.; Bahr, P.; Müller-Busch, C.; Ellershaw, J.; de Conno, F.; Vanden Berghe, P. & board members of the EAPC (2016). Euthanasia and physician-assisted suicide: A white paper from the european association for palliative care. *Palliative medicine*, 30(2):104--116.
- Räihä, K.-J.; Aula, A.; Majaranta, P.; Rantala, H. & Koivunen, K. (2005). Static visualization of temporal eye-tracking data. *Human-Computer Interaction-INTERACT 2005*, pp. 946--949.
- Rajput, A.; Aharwal, R. P.; Dubey, M.; Saxena, S. & Raghuvanshi, M. (2011). J48 and JRIP rules for e-governance data. *International Journal of Computer Science and Security*, 5(2):201--207.
- Raschka, S. (2015). *Python machine learning*. Packt Publishing Ltd.
- Reis, E. A. & Reis, I. A. (2002). Análise descritiva de dados: Síntese numérica. *Belo Horizonte, Universidade Federal de Minas Gerais*.
- Reyna, V. F. & Rivers, S. E. (2008). Current theories of risk and rational decision making. *Developmental Review: DR*, 28(1):1--11.
- Ribeiro Júnior, J. I. (2012). *Métodos estatísticos aplicados à melhoria da qualidade*. Editora UFV.
- Riboldi, J.; Barbian, M. H.; da Silva Kolowski, A. B.; Selau, L. P. R. & Torman, V. B. L. (2014). Precisão e poder de testes de homocedasticidade paramétricos e não-paramétricos avaliados por simulação. *Revista Brasileira de Biomedicina*, 32(3):334--344.
- Rich, E. & Knight, K. (1991). *Artificial intelligence*. McGraw-Hill, New York.

- Ristovski, G.; Hunter, M.; Olk, B. & Linsen, L. (2013). EyeC: Coordinated views for interactive visual exploration of eye-tracking data. In *Information Visualisation, 2013 17th International Conference*, pp. 239--248. IEEE.
- Rodrigues Júnior, J. F. (2003). Desenvolvimento de um framework para análise visual de informações suportando data mining. Master's thesis, Instituto de Ciências Matemáticas e de Computação, São Carlos.
- Rosenblatt, F. (1957). *The perceptron: A perceiving and recognizing automaton*. Cornell Aeronautical Laboratory, New York.
- Ruiz-Fernández, D.; Torra, A. M.; Soriano-Payá, A.; Marín-Alonso, O. & Palencia, E. T. (2016). Aid decision algorithms to estimate the risk in congenital heart surgery. *Computer Methods and Programs in Biomedicine*, 126:118--127.
- Santoro, D. M. (2005). Sobre o processo de seleção de subconjuntos de atributos-as abordagens filtro e wrapper. Master's thesis, Universidade Federal de São Carlos.
- Sasson, N. J. & Elison, J. T. (2012). Eye tracking young children with autism. *Journal of Visualized Experiments*, (61).
- Scher, H.; Furedy, J. J. & Heslegrave, R. J. (1984). Phasic t-wave amplitude and heart rate changes as indices of mental effort and task incentive. *Psychophysiology*, 21(3):326--333.
- Scherfler, C.; Göbel, G.; Müller, C.; Nocker, M.; Wenning, G. K.; Schocke, M.; Poewe, W. & Seppi, K. (2016). Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism. *Neurology*, 86(13):1242--1249.
- Scholey, A.; Jackson, P. & Kennedy, D. (2006). Mental effort, blood glucose and performance. *Appetite*, 47(2):277.
- Shic, F. & Scassellati, B. (2007). A behavioral analysis of computational models of visual attention. *International Journal of Computer Vision*, 73(2):159--177.
- Shortliffe, E. H. (1984). The science of biomedical computing. *Medical Informatics*, 9(3-4):185--193.
- Siqueira-Batista, R.; Cardoso, F.; Gomes, A. P.; Fonseca, J. d. O.; Bernini, A. F.; Barros, L. V.; Souza, C. R.; Esperidião-Antonio, V. & Costa, I. T. d. (2015). Decision-making in (bio)ethics: A preliminary study using mobile eye tracking. *Revista Brasileira de Educação Médica*, 39(4):496--501.

- Smith, J. D.; Masek, G. A.; Ichinose, L. Y.; Watanabe, T. & Stark, L. (1970). Single neuron activity in the pupillary system. *Brain Research*, 24(2):219--234.
- Steen, K. V. (2012). *Statistics - theory*. Liège, *Université de Liège*.
- Strandvall, T. (2008). Eye tracking as a tool in package and shelf testing.
- Sugarman, J. & Sulmasy, D. P. (2010). *Methods in medical ethics*. Georgetown University Press.
- Tan, P.-N.; Steinbach, M. & Kumar, V. (2006). *Introduction to data mining*. Pearson Education, São Carlos.
- Torman, V. B. L.; Coster, R. & Riboldi, J. (2012). Normalidade de variáveis: Métodos de verificação e comparação de alguns testes não-paramétricos por simulação. *Clinical & Biomedical Research*, 32(2):227--234.
- Tourassi, G.; Voisin, S.; Paquit, V. & Krupinski, E. (2013). Investigating the link between radiologists' gaze, diagnostic decision, and image content. *Journal of the American Medical Informatics Association*, 20(6):1067--1075.
- Tsang, H. Y.; Tory, M. & Swindells, C. (2010). eSeetrack: Visualizing sequential fixation patterns. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):953--962.
- Verma, L.; Srivastava, S. & Negi, P. (2016). A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data. *Journal of Medical Systems*, 40(7):1--7.
- Vieira, P. M.; Ramos, J. & Lima, C. S. (2015). Automatic detection of small bowel tumors in endoscopic capsule images by roi selection based on discarded lightness information. In *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3025--3028. IEEE.
- Villacampa, O. (2015). *Feature selection and classification methods for decision making: A comparative analysis*. PhD thesis, College of Engineering and Computing - Nova Southeastern University, Fort Lauderdale.
- West, J. M.; Haake, A. R.; Rozanski, E. P. & Karn, K. S. (2006). eyepatterns: Software for identifying patterns and similarities across fixation sequences. In *Proceedings of the 2006 symposium on Eye tracking research & applications*, pp. 149--154. ACM.

- Williams, A. & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport*, 69(2):111--128.
- Witten, I. H.; Frank, E.; Hall, M. A. & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Wooding, D. S. (2002). Fixation maps: Quantifying eye-movement traces. In *Proceedings of the 2002 symposium on eye tracking research & applications*, pp. 31--36. ACM.
- Wu, M. M. & Munzner, T. (2015). SEQIT: Visualizing sequences of interest in eye tracking data. *IEEE TVCG*, 22(1):449--458.
- Wu, X.; Kumar, V.; Quinlan, J. R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G. J.; Ng, A.; Liu, B.; Philip, S. Y. et al. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1--37.
- Yang, Y. & Pedersen, J. O. (1997). A comparative study on feature selection in text categorization. In *ICML*, volume 97, pp. 412--420.
- Yarbus, A. (1967). Eye movements and vision. 1967. *New York*.
- Yu, C.; Yurovsky, D. & Xu, T. L. (2012). Visual data mining: An exploratory approach to analyzing temporal patterns of eye movements. *Infancy*, 17(1):33--60.
- Zagermann, J.; Pfeil, U. & Reiterer, H. (2016). Measuring cognitive load using eye tracking technology in visual computing. In *BELIV'16: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pp. 78--85.
- Zhang, L.; Wade, J.; Swanson, A.; Weitlauf, A.; Warren, Z. & Sarkar, N. (2015). Cognitive state measurement from eye gaze analysis in an intelligent virtual reality driving system for autism intervention. In *Affective Computing and Intelligent Interaction, 2015 International Conference on*, pp. 532--538. IEEE.
- Zhang, Y. & Rajapakse, J. C. (2009). *Machine learning in bioinformatics*, volume 4. John Wiley & Sons.
- Zhang, Y.; Sun, Y.; Phillips, P.; Liu, G.; Zhou, X. & Wang, S. (2016). A multilayer perceptron based smart pathological brain detection system by fractional fourier entropy. *Journal of Medical Systems*, 40(7):1--11.

Apêndice A

Artigo publicado no International Conference on Computational Science 2017 (Qualis A1)

Fernandes, Daniel L.; Siqueira-Batista, Rodrigo; Gomes, Andréia P.; Souza, Camila R.; Costa, Israel T.; Cardoso, Felipe da S. L.; Assis, João Vitor; Caetano, Gustavo Henrique L.; Cerqueira, Fabio R. **“Investigation of the visual attention role in clinical bioethics decision-making using machine learning algorithms”**. In International Conference on Computational Science 2017 (ICCS 2017), Zürich. Procedia Computer Science, 2017. v. 108. p. 1165-1174.



International Conference on Computational Science, ICCS 2017, 12-14 June 2017,
Zurich, Switzerland

Investigation of the visual attention role in clinical bioethics decision-making using machine learning algorithms

Daniel L. Fernandes¹, Rodrigo Siqueira-Batista³, Andréia P. Gomes³, Camila R. Souza³, Israel T. da Costa⁴, Felipe da S. L. Cardoso⁴, João V. de Assis⁴, Gustavo H. L. Caetano⁴, and Fabio R. Cerqueira^{1,2}

¹ Graduate Program in Computer Science, Universidade Federal de Viçosa, Minas Gerais, Brazil
{[daniel.louzada](mailto:daniel.louzada@ufv.br), [fabio.cerqueira](mailto:fabio.cerqueira@ufv.br)}@ufv.br

² Department of Production Engineering, Universidade Federal Fluminense, Rio de Janeiro, Brazil
frcerqueira@id.uff.br

³ Department of Medicine and Nursing, Universidade Federal de Viçosa, Minas Gerais, Brazil
{[rsbatista](mailto:rsbatista@ufv.br), [andrea.gomes](mailto:andrea.gomes@ufv.br), [camila.r.souza](mailto:camila.r.souza@ufv.br)}@ufv.br

⁴ Department of Physical Education, Universidade Federal de Viçosa, Minas Gerais, Brazil
{[israel.teoldo](mailto:israel.teoldo@ufv.br), [felippe.cardoso](mailto:felippe.cardoso@ufv.br), [joao.assis](mailto:joao.assis@ufv.br), [gustavo.caetano](mailto:gustavo.caetano@ufv.br)}@ufv.br

Abstract

This study proposes the use of a computational approach based on machine learning (ML) algorithms to build predictive models using eye tracking data. Our intention is to provide results that may support the study of medical investigation in the decision-making process in clinical bioethics, particularly in this work, in cases of euthanasia. The data used in the approach were collected from 75 students of the nursing undergraduate course using an eye tracker. The available data were processed through feature selection methods, and were later used to create models capable of predicting the euthanasia decision through ML algorithms. Statistical experiments showed that the predictive model resultant from the multilayer perceptron (MLP) algorithm led to the best performance compared with the other tested algorithms, presenting an accuracy of 90.7% and a mean area under the ROC curve of 0.90. Interesting knowledge (patterns and rules) for the studied bioethical decision-making was extracted using simulations with MLP models and inspecting the obtained decision-tree rules. The good performance shown by the obtained MLP predictive model demonstrates that the proposed investigation approach may be used to test scientific hypotheses related to visual attention and decision-making.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: Visual attention, Decision-making in bioethics, Mobile eye tracking, Machine learning in medicine

1 Introduction

Decision-making and problem-solving capacity represent important cognitive abilities for people's daily lives. In the context of large raw datasets, data mining techniques, particularly

machine learning (ML) algorithms, have become remarkably useful in the decision-making process due to the ability to reveal hidden patterns, converting data into useful information [9].

In clinical practice, in particular in situations that involve bioethical issues, decision-making by a healthcare professional arises from a complex interaction between the work environment and neural processes parallel to cognition and emotion, leading to choices and judgments that take into account knowledge, risk, uncertainty, and regret, among other aspects [3]. Along with the use of bioethics frameworks in the decision-making process, the adoption of techniques for investigating and supporting ethical decision-making in healthcare has been discussed [11].

This is the case of mobile eye tracking (MET), a mobile technological device that enables measuring eye position and movement in order to map visual attention [27]. Recently, this tool has shown great potential for application in a wide range of fields of knowledge [21], e.g., in mental health [11], radiology and diagnostic imaging [26], human-computer interaction [14], studies of teaching-learning processes [27], and sports [4]. In addition to its broad application, the main interest in the use of MET is due to the fact that it is widely accepted as a valuable non-invasive method for studying visual attention characteristics as well as cognitive, emotional, and sensory status of a subject [11]. Therefore, MET plays an important role in advanced understanding of human perception, reasoning studies, and strategies for problem solving [26].

It is important to notice that the most challenging part of researches using MET is data analysis due to the typical small number of participants - in spite of the volume and complexity of the information produced - and due to imperfect data collection given the interferences produced in the visual exploration process. Hence, combining methodologies is recommended to enhance the analysis of the acquired data, thus optimizing the assessment of the influence of visual aspects in the cognitive process [27]. Although the aforementioned studies have shown an important evolution in eye-tracking research, they were not carried out in conjunction with advanced computational support, such as by employing ML techniques, specially in bioethics.

The present investigation was designed based on these considerations aiming to experimentally study the role of visual search - using MET - in the decision-making process in the case of euthanasia (this project was approved by the Research Ethics Committee of Federal University of Viçosa - CAAE No. 25353714.8.0-000.5153). Euthanasia, which is clearly a classic theme in clinical bioethics [12], is defined as mercy killing or dying process of an ill person, i.e, when someone other than the patient is committed to an action with the goal of alleviating ample and unbearable suffering of the patient [1]. The idea of studying MET data to relate visual features to decision-making for such a complex moral issue was first proposed by Siqueira-Batista et al. (2015) [24], and recommended by Pärnamets et al. (2015) [21]. The present work carries out the authors' proposal using advanced computational methods.

The predictive models obtained in this study resulted in interesting patterns that associate cognition-related visual aspects with the decision-making process. Those results have great potential for future research aiming to build computational tools to help in the ideal set-up of medical teams and to support the decision-making procedure in critical situations where no moral consensus has been established.

2 Materials and Methods

2.1 Machine learning framework

Although this study focuses on the specific case of correlating visual characteristics to the practice of euthanasia, it is worth pointing out that the overall idea of the procedure hereby presented may be applied to any field that uses MET resources to associate vision data with

results pertaining to the study at hand.

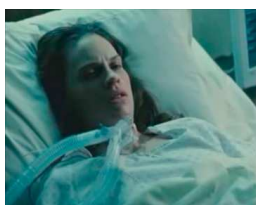
The presented pipeline is similar to that of the process of knowledge discovery in databases (KDD) [9]. The pipeline is split into four steps, namely: acquisition, preprocessing, data mining, and knowledge extraction. The first step consists of collecting eye data through MET, performing the appropriate sampling, and converting the resulting sampled data into a suitable format. The second step comprises data preprocessing, i.e., improving data quality and choosing proper variables by attribute selection. The third step includes experiments with ML algorithms on the resulting training set (TS), and the assessment of the statistical performance of the generated models. The fourth, and final step, consists of extracting knowledge by obtaining patterns and tendencies that may lead to plausible hypotheses.

2.2 Dataset

The TS used in this study for building the ML algorithms was created based on data collected from 75 voluntary students (17 to 47 years old, mean 21.36, 14 of whom males) from the first to ninth semesters of the nursing undergraduate course of a large Brazilian public university¹. Data were collected using ASL’s Mobile Eye-XG eye tracker with a frame rate of 30 Hz².

MET is a device able to monitor eye behavior by analyzing the central vision of the participants using specialized cameras attached to glasses. To that end, the equipment uses an infrared light source to determine the minimum and exact central vision fixation sites every 40 milliseconds (ms), besides identifying the image of pupil and cornea [14]. Prior to data acquisition, the collection environment was duly prepared and the MET device was calibrated.

Next, a 2 min 24 s excerpt of the movie “Million Dollar Baby” (Figure 1) was shown on the same screen used for MET calibration so that the participants’ eye movements could be identified. The film tells the story of Maggie Fitzgerald, a boxer that suffered a serious injury and became tetraplegic after an unfair fight. Over time, several complications occurred, which led Maggie to attempt suicide. After failing, she turned to her coach (Frankie) and asked him for euthanasia.



(a) Maggie’s scene.



(b) Frankie’s scene.

Figure 1: Screenshots of the movie the participants watched.

After the boxer expressed her will and made the request to the coach, the movie was interrupted at a crucial moment when a decision would be made. The participants in the research were then inquired about their decision (whether to fulfill or not the request of the boxer) and were asked to write down their reasons. Those answers remained confidential. By the end of the experiment, the following data were obtained from each participant: (i) the computational eye-tracking record and (ii) the answer on the decision along with its justification.

¹Raw data available at: <https://drive.google.com/drive/folders/0B8aRarbjcs87VGdLa3FvYnZRV1k?usp=sharing>

²The equipment used in this research: <https://www.youtube.com/watch?v=fvJPVInvsNk>

During TS composition, the justifications in the answers on the participants' decision-making were analyzed and it was observed that the decision of some participants was influenced by matters such as sanctity of life, religion, or the code of ethics of nursing professionals. Therefore, since one of the goals of this study is to investigate the correlation of eye data with the decision-making process, the data coming from those participants were removed from our TS, thus preventing undesired biases in the predictive models.

At the end, the TS contained 27 positive (would practice euthanasia) instances and 16 negative instances, i.e., 43 participants (17 to 26 years old, mean 20.93, ten of whom males) who, according to their justifications, were neutral and had none of the influences mentioned above. After concluding the final TS, it was converted into the ARFF format [15].

2.3 Attributes

The next step, after generating the TS, is attribute selection, i.e., defining which attributes are relevant and non-redundant among the 35 initial ones. The selection reduces the dataset complexity, which makes model building quicker and, in many cases, yields better results [5, 19].

First, based on a previous analysis, 14 eye attributes were manually removed for being clearly redundant (e.g., the same measure is encompassed by two variables, one as a percentage value and another as an absolute value).

Thus, the dimensionality of the initial set was reduced to 21 numerical attributes: number of fixations on: Maggie's face, Maggie's tracheal tube, Maggie's chest, Maggie's scene space, Frankie's face, Frankie's chest, Frankie's scene space, and areas without classification; total number of fixations during the experiment; duration of fixations (ms) on: Maggie's face, Maggie's tracheal tube, Maggie's chest, Maggie's scene space, Frankie's face, Frankie's chest, Frankie's scene space, and areas without classification; total duration of fixations during the experiment; mean value (pixels) of: overall pupil diameter, horizontal pupil diameter, and vertical pupil diameter; binary decision class variable (whether to practice or not euthanasia).

Next, we performed a computational attribute selection using a wrapper approach. Wrapper methods use a classification algorithm to assess the quality of each selected attribute subset [19]. We used this approach along with an exhaustive search strategy to systematically analyze all subset possibilities ($2^{21} - 1$). As a result, the combination of attributes that led to the ML model with the highest predictive power was the subset: number of fixations on Maggie's chest (`maggies_chest_fixation`), time of fixations on Maggie's tracheal tube (`tube_duration`), total time of fixations during the experiment (`total_duration`), mean value of overall pupil diameter (`overall_pupil_dilation`), and mean value of vertical pupil diameter (`vertical_pupil_dilation`).

2.4 Machine learning toolkits

The present study employed two data mining programs, both open-source and multiplatform. The several ML algorithms that will be mentioned later on in the text were applied using the software WEKA, which is widely used for classification tasks in several scientific works [15].

The other software used was NICeSim, a simulator that aims at understanding the correlation between the attributes and the class variable [5]. The simulation environment includes a graphical interface for the researcher to change the values of the attributes to create distinct scenarios (simulated instances). For each user intervention, the system shows the probability of the simulated instance to be positive. As a result, tendencies can be extracted according to the simulated scenarios, which facilitate the study of scientific hypotheses. The simulations performed in NICeSim are based on predictive models built with Support Vector Machine (SVM) and Multilayer Perceptron (MLP).

2.5 Learning algorithms

Several algorithms may be used in the classification task to predict to which class a given record belongs. The selection of classification algorithms must be based on some criteria such as precision, robustness, scalability, velocity, and interoperability of the model [9, 25]. In the present study, some supervised classification algorithms were tested, all widely used in medical informatics research [6]. A brief explanation of each of these algorithms is given below.

Naïve Bayes (NB) is a classification algorithm that provides a simple and quick approach based on the Baye's theorem [18]. In order to estimate the probability of the vector of attributes conditioned to the class, it is assumed that the attributes are independent, which simplifies the calculation. As long as there is no dependency, or it is weak, the algorithm results in an accurate model with quite shorter learning time compared with several other algorithms [18, 25].

Artificial neural networks comprise learning approaches inspired by the functional structure of the biological neural system. MLP is an example used to solve complex classification problems. An MLP has several layers and intermediate nodes (artificial neurons), also called hidden nodes, between the input and output layers, and typically uses the sigmoid function as activation function to produce non-linear outputs. In order to adjust the weights of connections between neurons, MLPs apply an algorithm called backpropagation during the learning process. This algorithm uses the difference observed between the value yielded by the network and the value expected in a certain layer to update the weights of the previous layer [20, 25].

The SVM approach, for linearly separable data, establishes a maximum-margin hyperplane as decision boundary in order to prevent overfitting. To that end, the SVM goal is formulated as a quadratic programming problem. When the data are not linearly separable, a kernel function is employed to map the data to a higher dimension space to make them linearly separable and, therefore, apply the linear separation method [25]. Sequential Minimal Optimization (SMO) is an algorithm widely used to solve the SVM quadratic programming problem. SMO uses an analytical optimizer and linear memory to deal with large datasets [2].

The learning algorithm k -nearest neighbor (k NN) classifies an instance by analyzing its closest k -neighbors [20]. The class is attributed according to the majority class of the selected neighbors. In this calculation, the votes often have different weights according to the neighbor distance, i.e., the closer the neighbor, the higher the weight its class will have in defining the class of the instance under analysis [20, 25].

Decision trees, such as C4.5, are part of a very popular and easily interpretable classification method, as the resultant rules that relate the attributes to the class are very easy to extract [23]. The C4.5 is an evolution of a simple model called ID3. C4.5 presents the following enhancements: inclusion of numerical attributes, treatment of missing values, and pruning after building the tree in order to decrease its complexity and minimize the possibility of overfitting [23, 25].

Random Forest (RF) is an ensemble classification method that uses the result of several decision trees. The final decision of the class is acquired by a voting schema in which the majority class obtained from the different trees prevails. To build each tree, a TS sample is used. Each sample is built by extracting a random subset from the original attributes [22, 25].

3 Results and Discussion

3.1 Model performance evaluation

We used a 10-fold cross validation to assess the resultant ML models. Table 1 summarizes the performance of each mentioned learning algorithm, with and without cost-sensitive classification

(CSC), regarding mean accuracy (Ac), precision (Pr), sensitivity (Sn) and, area under the ROC curve (AUC). It is important to mention that SMO was used with (RBFkernel) and without (linear classification) kernel. Furthermore, the number of neighbors was set to 3 for k NN. Except for these modifications, the default parameters of WEKA were kept.

Classifier	Type	Ac	Pr	Sn	AUC	Confidence (95%) interval of Ac
NB	Normal	65.1%	64.7%	65.1%	0.660	[50.16%, 80.04%]
MLP	Normal	90.7%	91.9%	90.7%	0.903	[75.76%, 100.00%]
SMO (no kernel)	Normal	62.8%	58.4%	62.8%	0.513	[47.86%, 77.74%]
SMO (RBFKernel)	Normal	62.8%	39.4%	62.8%	0.500	[47.86%, 77.74%]
k NN	Normal	72.1%	75.1%	72.1%	0.667	[57.16%, 87.04%]
C4.5	Normal	65.1%	64.7%	65.1%	0.668	[50.16%, 80.04%]
RF	Normal	62.8%	61.1%	62.8%	0.686	[47.86%, 77.74%]
NB	CSC	58.1%	60.4%	58.1%	0.660	[43.16%, 73.04%]
MLP	CSC	81.4%	81.4%	81.4%	0.852	[66.46%, 96.34%]
SMO (no kernel)	CSC	67.4%	71.1%	67.4%	0.690	[52.46%, 82.34%]
SMO (RBFKernel)	CSC	44.2%	46.7%	44.2%	0.428	[29.26%, 59.14%]
k NN	CSC	72.1%	75.1%	72.1%	0.635	[57.16%, 87.04%]
C4.5	CSC	67.4%	71.1%	67.4%	0.667	[52.46%, 82.34%]
RF	CSC	60.5%	59.0%	60.5%	0.699	[45.56%, 75.44%]

Table 1: Statistical evaluation (10-fold cross validation) of models generated from several learning algorithms with and without cost-sensitive classification (CSC).

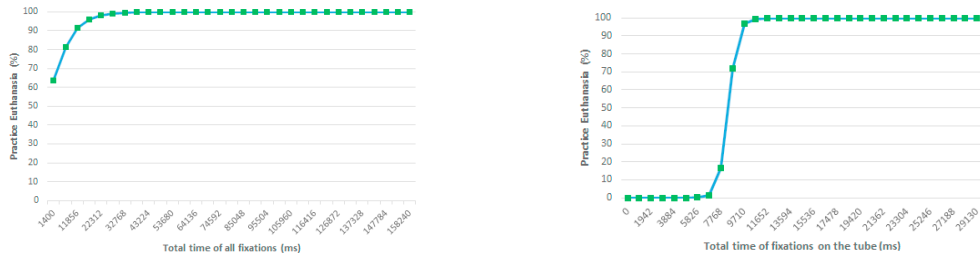
CSC is widely used as a way of dealing with imbalanced datasets with which the generated model may present a bias toward the majority class [5, 25]. In our case, the positive class is the majority, but the imbalance is not severe. Still, we decided to also test CSC to check whether the numerical difference between the classes would cause any sort of bias in the models.

As can be seen, the model generated by the MLP network with homogeneous costs stands out considerably in comparison with the other models, delivering better performance across all metrics. The high AUC value demonstrates a high predictive power and means that the probabilities generated by the model are robust. These probabilities are fundamental to find the patterns reported in the next section.

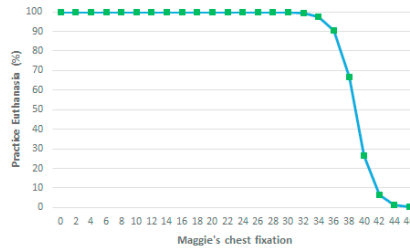
3.2 Extracting patterns by means of simulations with MLP models

In the context of knowledge extraction, NICEsim was used to simulate various scenarios by combining the values of attributes (simulated instances) in order to provide a better understanding of the relation between attributes and the class. After performing the simulations, five plots were generated showing the relation between the attributes and the probability of practicing euthanasia.

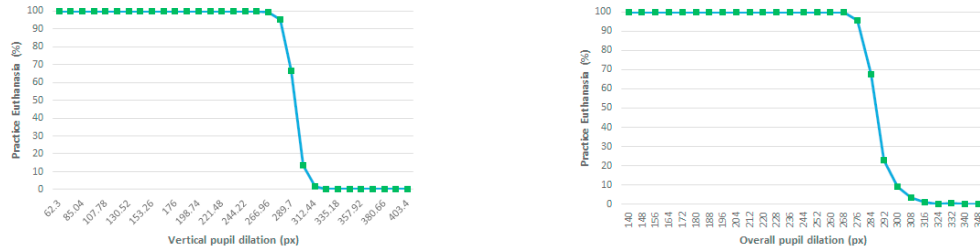
Figures 2a and 2b illustrate the relation between the duration of fixations and the probability of practicing euthanasia. The duration of eye fixation is a measure that examines the relation between the cognitive load and the visualization scene [14]. This measure corresponds to the duration of the cognitive processing of a person in relation to a determined area of interest (AOI) [17]. As can be seen in Figure 2a, for fixations lasting around 1400 ms, the participants present a little over 60% chance of practicing euthanasia. This value is close to 100% when the duration of fixations goes above 32768 ms. Figure 2b, on the other hand, show that fixations lasting less than 7768 ms on the tracheal tube (tracheotomy) lead to virtually no chance (0%) of practicing euthanasia, whereas values above 11652 ms show that the chance is close to 100%. Both figures show that the longer the overall fixations, the higher the chance of the participants practicing euthanasia. It is believed that the participants that took longer looking at AOIs -



(a) Overall duration of fixations. (b) Duration of fixations on the tracheal tube.



(c) Number of fixation on Maggie's chest.



(d) Mean vertical pupil dilation. (e) Mean overall pupil dilation.

Figure 2: Predicted probability of practicing euthanasia versus five key variables.

such as the tracheotomy - became more appalled by the patient’s situation and, consequently, chose to interrupt the suffering.

Figure 2c indicates that attention to Maggie’s chest is a determinant factor, showing that this is definitely an AOI, because important elements receive fixations more often [10]. In this analysis, a low number of fixations (below 32) represent a high probability of practicing euthanasia (approximately 100%), while values above 44 lower the chance to approximately 0%. At first glance, the tendency shown in Figure 2c seems antagonistic to the conclusions taken from Figures 2a and 2b. But, it is important to notice that the patient’s chest area has a string with a crucifix. One plausible explanation is that the more the participants fixated their eyes on the region of this religious object, seeking information, the more their religious bias was taken into account. Even though we have removed the participants who declared to be religious from our dataset, when it comes to cognitive processing, nothing keeps the object in the frame from perhaps influencing the participants’ decision to some extent.

For pupil behavior, either vertical (Figure 2d) or overall (Figure 2e), the results show that the larger the pupil diameter, the lower the probability of practicing euthanasia. Values above 312 pixels indicate a very low probability of practicing euthanasia, whereas values below 268 pixels indicate a chance close to 100%. The pupil accurately represents the cognitive effort as well as the judgment of one’s interest in carrying out a given task, and it is considered an important

cognition element [4, 14]. Then, a high pupil dilation indicates a high level of cognitive effort. When the stimulus decreases, the pupil contracts until returning to its resting size [4] - a process also known as miosis [16]. In this context, it is believed that the participants' who had a very high energy demand opted for a more rational than emotional decision, respecting the oath and code of ethics of nursing professionals. Furthermore, euthanasia is considered homicide according to articles 121 and 122 of the Brazilian Penal Code [8].

Our experiments show that the chosen variables are adequate to build predictive models to the studied ethical decision-making. It was expected, as the components assessed by MET are associated with the participant's cognitive and emotional capacities [13]. Therefore, the relation of these variables to the ethical decision-making shows the significant role that the vision plays in the studied case.

3.3 Extracting patterns by means of decision tree rules

Although Table 1 shows that C4.5 is far less accurate when compared to MLP, an experiment was carried out to verify whether the extracted rules of a C4.5 model would match the tendencies found from the MLP model. Figure 3a and Figure 3b show the rules along with the decision tree from where the rules could be extracted, and demonstrate that the found patterns matched the results of the MLP model. Notice that the rules were taken from a cost-sensitive C4.5 approach, because the C4.5 model, in its original form, presented worse performance (Table 1).

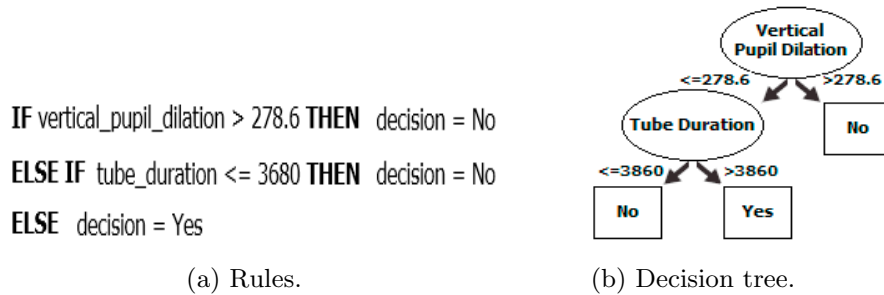


Figure 3: Patterns generated by a cost-sensitive C4.5 model.

The obtained rules show that the attributes referring to vertical pupil dilation and duration of fixation on the patient's tracheotomy are factors that best predict the decision to be made.

Pupil dilation may be indicative of the cognitive load at the decision-making moment [13]. The increase in cortical activation in the limbic system, due to the levels of psychological resources used to solve problems (such as attention, reasoning, and judgment), promotes an increase in pupillary diameter - related to the activation of the sympathetic nervous system [7] - a process also known as mydriasis [16]. The result shown in Figure 3 evidences once more a correlation between pupil dilation and the participant's decision-making process in the present experimental protocol, which shows that the cognitive load stimulated by the level of difficulty of the task primarily influences the participant's decision making. Namely, according to the decision tree, to reach a decision 'yes', the vertical pupil dilation has to be necessarily high (meaning high cognitive load).

The results shown in Figure 2 and Figure 3 indicate that it is also very important to analyze the region where the participant seeks information to make decisions, so that we can understand whether the cognitive activities are related to emotion or reason. Similarly to Figure 2, the rules shown in Figure 3 provide evidence that a high cognitive load, signaled by the high number of

fixations on the tracheal tube, relates to the decision ‘yes’. In this case, it is likely an emotional load due to the fact that the tube demonstrates that the patient is indeed suffering, which is one of the core elements for ethical decision-making in cases of euthanasia [1].

4 Conclusions

The present investigation proposes a computational approach, based on ML and eye tracking techniques, that aims to investigate eye data generated from the appreciation of a piece of cinematographic work in order to aid studies investigating the influence of visual search in the bioethical decision-making process in a very relevant and controversial issue of clinical practice.

To that end, ML algorithms were applied - namely, supervised classification techniques - to a TS built from eye data of 75 (although it is apparently low, this number is expressive in relation to previous studies) students, in order to understand how visual search impacts the ethical decision-making of those future professionals. For the dataset used, MLP was the most accurate among all classification algorithms employed, presenting 90.7% of accuracy and mean AUC of 0.90.

The computational approach hereby presented proved to be useful for researches on cognition since it builds and validates a highly precise predictive model to simulate the effects the variables concerning the number of fixations, the duration of fixation on AOIs, and pupil dilation have on the probability of healthcare professionals practicing euthanasia. Therefore, the results of this study clearly show a connection between cognitive behavior and decision-making. Unfortunately, the proposed approach could not be compared with any other previous study since it is a pioneer in the field.

This study involved multiple areas of knowledge, which contributed to reveal the role of visual attention on the decision-making process, checking to what extent sight is a determining factor in decision-making, particularly in clinical bioethics when dealing with end-of-life issues.

Acknowledgments

The authors thanks the financial support of CAPES and FAPEMIG, Brazilian research agencies.

References

- [1] M. Berghe, B. D. De Casterle, and C. Gastmans. The complexity of nurses’ attitudes toward euthanasia: a review of the literature. *Journal of Medical Ethics*, 31(8):441–446, 2005.
- [2] L. Bottou and C. Lin. Support vector machine solvers. *Large scale kernel machines*, pages 301–320, 2007.
- [3] T. Brosch, K. R. Scherer, D. M. Grandjean, and D. Sander. The impact of emotion on perception, attention, memory, and decision-making. *Swiss Medical Weekly*, 143:1–10, 2013.
- [4] F. da S. L. Cardoso and I. T. da Costa. Behavior pupillary as indicative of specific knowledge of football player (Portuguese). *Revista Mineira de Educação Física*, Edição Especial(9):1087–1094, 2013.
- [5] F. R. Cerqueira, T. G. Ferreira, A. de P. Oliveira, D. A. Augusto, E. Krempser, H. J. C. Barbosa, S. do C. C. Franceschini, B. A. C. de Freitas, A. P. Gomes, and R. Siqueira-Batista. Nicesim: an open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making. *Artificial Intelligence in Medicine*, 62(3):193–201, 2014.

- [6] H. Chen, S. S. Fuller, C. Friedman, and W. Hersh. Knowledge management, data mining, and text mining in medical informatics. In *Medical Informatics*, pages 3–33. Springer, 2005.
- [7] H. D. Critchley, J. Eccles, and S. N. Garfinkel. Interaction between cognition, emotion, and the autonomic nervous system. *Handb Clin Neurol*, 117:59–77, 2013.
- [8] R. E. F. Dodge. Euthanasia - judicial aspects (Portuguese). *Revista Bioética*, 7(1):1–7, 2009.
- [9] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37, 1996.
- [10] P. M. Fitts, R. E. Jones, and J. L. Milton. Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review*, 9:1–6, 1950.
- [11] F. Galgani, Y. Sun, P. L. Lanzi, and J. Leigh. Automatic analysis of eye tracking data for medical diagnosis. In *Computational Intelligence and Data Mining, 2009. CIDM'09. IEEE Symposium on*, pages 195–202. IEEE, 2009.
- [12] J. R. Goldim. Bioethics: origins and complexity (Portuguese). *Revista Hospital de Clínicas de Porto Alegre*, 26(2):86–92, 2006.
- [13] E. Granholm and S. R. Steinhauer. Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, 52(1):1–6, 2004.
- [14] L. A. Granka, T. Joachims, and G. Gay. Eye-tracking analysis of user behavior in www search. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 478–479. ACM, 2004.
- [15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.
- [16] E. H. Hess and J. M. Polt. Pupil size in relation to mental activity during simple problem-solving. *Science*, 143(3611):1190–1192, 1964.
- [17] D. E. Irwin. Fixation location and fixation duration as indices of cognitive processing. *The interface of language, vision, and action: Eye movements and the visual world*, pages 105–133, 2004.
- [18] G. H. John and P. Langley. Estimating continuous distributions in bayesian classifiers. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 338–345. Morgan Kaufmann Publishers Inc., 1995.
- [19] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1):273–324, 1997.
- [20] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell. *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.
- [21] P. Pärnamets, P. Johansson, L. Hall, C. Balkenius, M. J. Spivey, and D. C. Richardson. Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences*, 112(13):4170–4175, 2015.
- [22] J. R. Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.
- [23] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [24] R. Siqueira-Batista, F. Cardoso, A. P. Gomes, J. de O. Fonseca, A. F. Bernini, L. V. Barros, C. R. Souza, V. Esperidião-Antonio, and I. T. da Costa. Decision-making in (bio)ethics: a preliminary study using mobile eye tracking. *Revista Brasileira de Educação Médica*, 39(4):496–501, 2015.
- [25] P. Tan, M. Steinbach, and V. Kumar. *Introduction to data mining*. Pearson Education, India, 2006.
- [26] G. Tourassi, S. Voisin, V. Paquit, and E. Krupinski. Investigating the link between radiologists' gaze, diagnostic decision, and image content. *Journal of the American Medical Informatics Association*, 20(6):1067–1075, 2013.
- [27] C. Yu, D. Yurovsky, and T. L. Xu. Visual data mining: An exploratory approach to analyzing temporal patterns of eye movements. *Infancy*, 17(1):33–60, 2012.