DOI:10.36517/resdite.v5.n3.2020.a1 **ARTIGO**

Emprego de técnicas de data mining na identificação de padrões relacionados às doenças oculares em pacientes pediátricos atendidos em um Hospital Universitário de uma cidade do sudoeste do Brasil

USE OF DATA MINING TECHNIQUES TO IDENTIFY PATTERNS RELATED TO EYE DISEASES IN PEDIATRIC PATIENTS TREATED AT A UNIVERSITY HOSPITAL IN A CITY IN SOUTHWEST BRAZIL

Adriana Faria Gappo Prata[1], Martius Vicente Rodriguez y Rodriguez[2], Adauto Dutra Moraes Barbosa[3]

[1] Mestre. Universidade Federal Fluminense.
ORCID: https://orcid.org/0000-0001-9778-8962
Email: adrianaprata@id.uff.br
[2] Doutor. Universidade Federal Fluminense.
ORCID: https://orcid.org/0000-0001-8270-7488
Email: martiusrodriguez@id.uff.br
[3] Doutor. Universidade Federal Fluminense.
ORCID: https://orcid.org/0000-0003-2841-9877
Email: adautodutra@id.uff.br

Correspondência: Rua Mariz e Barros, 62 apto 701, Icaraí- Niterói- RJ. CEP 24220-121.

Conflito de interesses: os autores declaram que não há conflito de interesses.

Como citar este artigo

Prata AFG; Rodriguez MVR; Barbosa ADM. Emprego de técnicas de data mining na identificação de padrões relacionados às doenças oculares em pacientes pediátricos atendidos em um Hospital Universitário de uma cidade do sudoeste do Brasil. Revista de Saúde Digital e Tecnologias Educacionais. [online], volume 5, n. 3. Editor responsável: Luiz Roberto de Oliveira. Fortaleza, dezembro de 2020, p. 01-13. Disponível em: http://periodicos.ufc.br/resdite/index. Acesso em "dia/mês/ano".

Data de recebimento do artigo: 21/10/2020
Data de aprovação do artigo: 05/12/2020
Data de publicação: 31/12/2020

Resumo

Objetivo: Empregar técnicas de data mining no tratamento de dados coletados de prontuários de pacientes, identificando atributos mais relevantes e ferramentas de data mining mais adequadas para análise de dados na área da saúde. **Métodos:** Estudo transversal realizado com dados secundários de pacientes

pediátricos atendidos no setor de oftalmologia de um Hospital Universitário, localizado no sudeste do Brasil, de janeiro de 2018 a dezembro de 2019. O programa IBM SPSS Statistics v.25 foi utilizado para caracterizar a amostra quanto às características demográficas e consulta oftalmológica dos pacientes. Utilizou-se a ferramenta R versão 4.0.0 para criação de um modelo de classificação pelo algoritmo Naves Bayes, cuja função era avaliar as variáveis para prever o diagnóstico de cada paciente. **Resultados:** A amostra foi de 196 olhos. A média de idade dos pacientes foi de 10,04 anos. 53% eram do sexo masculino e 66,7% eram pardos. As principais queixas que levaram os pacientes a procurar atendimento oftalmológico foram: Olho torto (26,9%) e Baixa Acuidade Visual (25,3%). O modelo de classificação criado obteve uma taxa de acerto do diagnóstico de 73%. **Conclusões:** Foi identificada a importância da informatização do sistema hospitalar e da formação de profissionais de saúde na área de ciência de dados.

Palavras-chave: Mineração de dados. Doenças oculares. Saúde infantil.

Abstract

Objective: *To employ data mining techniques to treat data collected from patient records, identifying the most relevant attributes and the most appropriate data mining tools for analyzing health data.* ***Methods:*** *Cross-sectional study was conducted with secondary data from pediatric patients seen in a University Hospital's ophthalmology sector located in southeastern Brazil from January 2018 to December 2019. The IBM SPSS Statistics v.25 was used to characterize the sample regarding demographic characteristics and ophthalmological consultation of patients. The tool R version 4.0.0 was used to create a classification model by the Naves Bayes algorithm, whose function was to evaluate the variables to predict each patient's diagnosis.* ***Results:*** *The sample consisted of 196 eyes. The mean age of the patients was 10.04 years. 53% were male, and 66.7% were brown. The main complaints that led patients to seek eye care were: Crooked eye (26.9%) and Low Visual Acuity (25.3%). The classification model created obtained a diagnosis accuracy rate of 73%.* ***Conclusions:*** *The importance of computerization of the hospital system and the training of health professionals in data science were identified.*

Keywords: *Data mining. Eye diseases. Children's health.*

1. Introdução

Com a informatização crescente na área da saúde, grandes bancos de dados vêm sendo alimentados com prontuários de pacientes, exames complementares, dados de planos de saúde, e dados advindo do uso de *wearables* (*smartphones e smartwhatches*)[1-4], tornando-se relevante para analisá-los o uso de inteligência artificial e técnicas de *Data Mining*, para encontrar padrões consistentes, descobrir relações entre os dados, realizar classificações, fazer previsões, testar hipóteses, dentre outras[5,6]. Com isso podemos auxiliar na tomada de decisões na área médica, visando a prevenção de potenciais doenças tanto no âmbito individual (medicina de precisão), quanto no de saúde pública.

Data mining é a quarta etapa das cinco que compõe um processo conhecido como KDD (*Knowledge Discovery in Databases*)[1,2,6,7], onde são definidos o método (descoberta não supervisionada de relações, testagem de hipóteses ou modelagem matemática dos dados), a técnica (classificação, regressão, associação, segmentação ou *clustering*, sumarização). Dentre as ferramentas de *data mining* mais adequadas para a análise e tratamento dos

dados coletados destacam-se as que utilizam Sistema R, *Pyton*, WEKA, etc[1,5,8-10]. Os padrões encontrados podem ser representados de diversas formas, tais como árvores de decisão, regras de associação, redes neurais e algoritmos genéticos[1,8].

O objetivo deste estudo foi o de empregar técnicas adequadas de *data mining* no tratamento de dados coletados nos prontuários físicos de pacientes pediátricos, que identificasse aqueles mais relevantes, através de um modelo capaz de predizer a melhor percentagem de acerto diagnóstico.

2. Métodos

Trata-se de estudo transversal realizado com dados secundários de pacientes em idade pediátrica[11], que foram atendidos no setor de Oftalmologia do Hospital Universitário Antônio Pedro (HUAP), no município de Niterói, estado do Rio de Janeiro, no período de janeiro de 2018 a dezembro de 2019.

Para caracterização da amostra, foi utilizado o programa SPSS Statistics® versão 25, no cálculo de frequências, medidas de tendência central, distribuição e dispersão de dados da amostra.

2.1 Seleção dos Dados (1ª etapa)

A coleta dos dados foi realizada no arquivo médico do HUAP, entre os meses de outubro a dezembro de 2019, de acordo com a relação de prontuários fornecido pelo setor de marcação de consultas do hospital. Com isso, obteve-se um "n amostral" de 396 olhos.

Como o registro dos dados dos pacientes ainda é feito em prontuários de papel, os dados extraídos foram digitados em planilha Excel® v. 2019, de forma que cada linha da planilha correspondesse aos dados de cada paciente e cada coluna às variáveis coletadas.

Foram coletados dados das seguintes variáveis: número do prontuário (para evitar dupla digitação); características demográficas referentes ao sexo (feminino, masculino), idade, raça/cor (branca, preta, parda, asiática, indígena); antecedentes perinatais referentes a gestação (realização de pré-natal, intercorrências na gravidez, uso de drogas ilícitas, tabagismo, consumo de álcool, ocorrência de TORCHS), ao parto (parto vaginal ou cesárea, ocorrência de sofrimento fetal, APGAR no 1º e 5º minutos e idade gestacional), dados antropométricos ao nascer (peso, comprimento e perímetro cefálico), período neonatal imediato (se houve intercorrências, qual o tipo e tempo de hospitalização); e dados de variáveis referentes à consulta oftalmológica (queixa principal, história da doença atual, história patológica pregressa, história familiar, acuidade visual sem correção, refração,

acuidade visual com correção, biomicroscopia, pressão intraocular, fundoscopia, diagnóstico e presença ou não de ambliopia).

2.2 Pré-processamento dos Dados (2ª etapa)

Após a criação do banco de dados em planilha Excel®, iniciou-se a limpeza dos dados para a eliminação dos ruídos, erros de preenchimento, dados inválidos, incompletos e irrelevantes.

Os dados referentes ao número de prontuário foram considerados irrelevantes para o processo de *data mining*, e por isso foram excluídos do banco de dados.

Através da adição de filtro nas colunas da planilha, permitiu-se analisar os dados à procura de erros de digitação.

Partiu-se então para a eliminação das variáveis que contivessem 60% ou mais de dados faltantes, ou seja, células em branco na planilha, pois não contribuiriam para a análise, considerando não haver nenhuma informação sobre aquela variável pesquisada no prontuário daquele paciente. Isso permitiu excluir 19 variáveis, ou seja, todas referentes aos antecedentes perinatais (gestação, parto, dados antropométricos e período neonatal imediato) e algumas variáveis referentes à consulta oftalmológica (HDA, HPP, HF e PIO).

A seguir, procedeu-se à limpeza das linhas, ou seja, foram excluídas aquelas possuindo ao menos uma variável com dado faltante. Desse modo, após esta fase, o "n" amostral caiu para 186 olhos.

2.3 Transformação dos Dados (3ª etapa)

Esta etapa teve como objetivo alterar a forma de como esses dados estavam representados de modo que pudessem ser analisados pelo algoritmo de *data mining*, porém sem modificar os seus significados: retirada dos acentos das palavras e digitação em caixa alta para padronização das mesmas; substituição de textos por uma palavra ou termo médico que possuísse o mesmo significado nas variáveis "Queixa Principal", "Biomicroscopia", "Fundoscopia" e "Diagnóstico"; agrupamento dos dados das variáveis "Acuidade Visual sem Correção" e "Acuidade Visual com Correção", de forma que, foi considerado como tendo "Boa visão" os pacientes que apresentaram as acuidades visuais de 20/15, 20/20 e 20/30, "Visão Média" as acuidades visuais de 20/40, 20/50, 20/60 e 20/70, "Visão Ruim" as acuidades visuais de 20/80, 20/100 e 20/200 e "Cegueira Legal" a percepção de vultos, percepção luminosa e ausência de fixação pelo paciente; e por fim, o agrupamento de dados da variável "Refração" e a sua divisão em "Correção Esférica"

(miopias e hipermetropias) e "Correção Cilíndrica" (astigmatismos), de maneira que houvesse somente um único valor em cada célula do *dataset* que será analisado, conforme tabela 01. Isso permitiu que se abolisse o sinal de (-) e (+) que serviam apenas para identificar se aquele valor dióptrico era uma miopia, hipermetropia ou astigmatismo e que poderia ser confundido, na análise pelo algoritmo, como um número natural maior ou menor que zero.

2.4 Data Mining (4ª etapa)

A partir do banco de dados pré-processado e transformado, foi gerada uma nuvem de palavras utilizando-se o site https://www.wordclouds.com/ para que se pudesse obter uma avaliação inicial sobre possíveis relações encontradas (figura 1). Quanto mais destacada a palavra, maior sua frequência no texto analisado[9].

Figura 1 – Nuvem de Palavras com dados da amostra após as fases de pré-processamento e transformação.

Fonte: Prata AFG, Rodriguez MVRY, Barbosa ADM. – elaborado no site https://www.wordclouds.com/

O método escolhido para a análise dos dados foi o método da Modelagem Matemática dos Dados[8,9], em face do vasto conhecimento sobre condições e doenças oftalmológicas capazes de causar ambliopia ou cegueira. A técnica utilizada nesta análise foi a da Classificação porque, segundo Amaral[9], é uma técnica que utiliza dados históricos

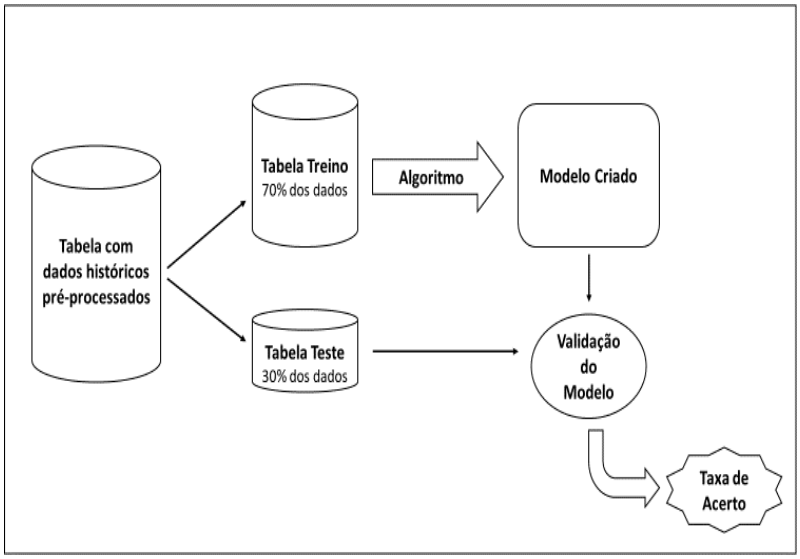
(no caso, os dados secundários dos pacientes) para construir um modelo e tentar predizer a classe de interesse (Diagnóstico).

Utilizou-se a ferramenta R versão 4.0.0 (24/04/2020), disponível no site da *The R Foundation for Statistical Computing* - https://www.r-project.org/foundation/ - , e o algoritmo *Naive Bayes* para a construção do modelo de classificação cuja função foi avaliar o quanto cada valor das variáveis contribuiu para aquele paciente ser "classificado" como tendo aquela doença (variável Diagnóstico)[9].

O *Naive Bayes*[9] é um algoritmo de estatística e aprendizado de máquina. Para que ele possa predizer um resultado, uma classe ou uma variável de interesse, a tabela com os dados históricos é dividida aleatoriamente em tabela de treino e tabela de teste. A tabela de teste contém 70% dos dados da tabela original escolhidos aleatoriamente recebendo o valor "1". A tabela de teste contém os 30% dos dados restantes, que receberam o valor "2" também de forma aleatória[9].

O modelo então é construído com base na tabela treino e validado com a tabela teste, ou seja, gera-se uma Matriz de Confusão[9] onde os valores dos dados da classe de treino serão comparados com os valores previstos pelo modelo, apresentando uma taxa de acerto (figura 2). Após a criação do modelo, quaisquer instâncias (linhas) com dados ainda não analisados poderão ser classificadas.

Figura 2 – Criação do Modelo de Classificação.



Fonte: Prata AFG, Rodriguez MVRY, Barbosa ADM.

3. Resultados

A amostra final obtida após a limpeza dos dados faltantes, dados inválidos, incompletos e irrelevantes foi de 196 olhos. A média de idade dos pacientes foi de 10,04 ± 4,22 anos, e a mediana de 10 anos.

Na tabela 2 estão compiladas as frequências absolutas e relativas dos resultados obtidos quanto às variáveis: sexo, raça, queixa principal, correção esférica, correção cilíndrica, diagnóstico e presença de ambliopia encontradas na amostra estudada.

Ao "rodar" a programação no sistema R, o algoritmo criou um modelo de classificação baseado na tabela treino e o validou através da tabela de teste, gerando o que se chama de Matriz de Confusão[9]. Ela pode ser vista parcialmente na figura 3 (anexo), obtendo com isso uma taxa de acerto do diagnóstico de 73%.

Discussão

A mineração de dados tem enorme vantagem quando temos um grande volume de dados complexos para serem analisados[5,12], pois é capaz de mapear dados brutos (geralmente muito volumosos e de difícil entendimento a princípio) e transformá-los em outras formas mais compactas (por exemplo: um relatório curto), mais abstratas (por exemplo: uma descrição aproximada ou modelo do processo que gerou os dados), ou mais útil (por exemplo: modelo preditivo para estimar o valor de casos futuros), através da aplicação de métodos específicos de mineração de dados para a descoberta e extração de padrões[2].

De acordo com Sharma e Mansotra[13], a mineração de dados é muito útil na área da saúde, pois o conhecimento obtido através desta técnica permite potencializar o auxílio no controle das infecções, diagnósticos e tratamentos de várias doenças, além de melhorar a gestão de recursos de saúde, gestão hospitalar e administração da saúde pública[13] e já está sendo utilizada em larga escala por muitas organizações de saúde[14]

Entre os resultados encontrados, após a mineração dos dados quanto à idade dos pacientes atendidos, observamos reduzido número na faixa etária entre zero e cinco anos de idade entre aqueles atendidos pelo setor. Esse fato nos chamou muito a atenção, considerando que o desenvolvimento do córtex visual da criança ocorre a partir de seu nascimento e estará completamente formado em torno dos oito anos de idade[15,16]. Apresentando o paciente alguma condição ou patologia capaz de afetar esse desenvolvimento visual, quanto mais perto da idade de oito anos ela estiver para começar o

tratamento, pior o prognóstico. A grande maioria dos pacientes se encontravam na faixa etária de 6 a 14 anos, sendo a média de 10 anos de idade, fazendo-nos pensar que os pais só procuram atendimento oftalmológico quando o seu filho relata ter algum problema visual, ou quando o problema se torna visível para os pais ou professores. Tal rotina deveria ser modificada o quanto antes.

Após analisar as principais queixas que levaram os pacientes a procurar assistência oftalmológica, a tese anterior ganha mais força, pois 26,9% dos casos apresentaram queixa de olho torto, observada pelos pais, e em 25% dos casos foi de Baixa Acuidade Visual, percebida pelos professores em sala de aula, ou quando a criança já atinge idade na qual consegue discernir ter um dos olhos incapaz de enxergar tão bem quanto o outro.

Entre os diagnósticos feitos nesta amostra, após emprego da técnica de mineração de dados, a maioria foi considerada normal (30,65%), vindo a seguir o grupo dos Estrabismos (21,6%). Esses resultados, embora não possuam validade externa, considerando o Hospital Universitário Antônio Pedro ser referência para casos de estrabismo, facilitam ao gestor reconhecer a necessidade de se preparar para recepcionar outras doenças oftalmológicas, porque além de receber vários pacientes encaminhados de outros setores, com doenças sistêmicas, serve como área de triagem para saber se apresenta ou não doença ou condição oftalmológica relacionada a sua doença de base.

Apesar do setor de oftalmologia do Hospital Universitário Antônio Pedro (HUAP) atender muitas crianças oriundas de várias cidades do estado do Rio de Janeiro, gerando um volume de dados gigantesco e de valor científico inestimável, a forma de coleta de dados primários e seu armazenamento em prontuários em papel, não é mais respaldado como meio adequado, atualmente[5,17]. A grande dificuldade de obtenção de dados confiáveis e consistentes através dos prontuários em papel, encolheu muito a amostra, limitando as análises com técnicas de mineração de dados, impedindo a máxima extração de conhecimento desses dados, dificultando a identificação de padrões ocultos e a geração de novos conhecimentos. Entretanto, o modelo conseguiu obter uma taxa de acerto de 73%, ou seja, ele conseguiu prever o diagnóstico corretamente em 73% dos casos, o que pode ser considerada uma taxa adequada, visto não existir modelo com 100% de acertamento[9].

Uma das limitações desse estudo foi o tamanho da amostra. Contribuíram para a amostra estudada ter um tamanho reduzido: prontuários não encontrados no arquivo, preenchimento incompleto dos dados do paciente e dos exames realizados, não realização de exames importantes como por exemplo, a acuidade visual e refração em crianças pré-verbais, o que é relevante para prevenção da ambliopia, já que muitos podem

apresentar erros refrativos e por não se expressarem verbalmente, o médico pode deixar essa condição passar despercebida.

4. Conclusão

Através da técnica de data mining foi possível identificar terem sido subestimados mais de um quarto dos dados do setor de oftalmologia. Tal achado confirma a necessidade de informatizar o sistema hospitalar através do uso de Prontuário Eletrônico de Paciente (PEP) ou de Registro Eletrônico de Saúde (RES), interoperável e em conformidade com as especificações do CFM/SBIS, preferencialmente certificado, permitindo aumentar a precisão diagnóstica e contribuindo também para melhorar o atendimento à criança com uma oftalmopatia.

Conflito de interesse: os autores declaram não haver conflitos de interesse pessoal ou comercial na realização deste trabalho.

5. Referências

1. Pereira J. Modelos de data mining para multi-previsão: aplicação à medicina intensiva [Internet] [Dissertação de mestrado]. [Braga - Portugal]: Universidade do Minho; 2005 [citado 14 de março de 2019]. Disponível em: https://core.ac.uk/download/pdf/55626354.pdf
2. Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data. Commun ACM. 1996;39(11):27–34.
3. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. J Integr Bioinforma [Internet]. 25 de setembro de 2018 [citado 3 de julho de 2019];15(3). Disponível em: http://www.degruyter.com/view/j/jib.2018.15.issue-3/jib-2017-0030/jib-2017-0030.xml
4. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst. dezembro de 2014;2(1):3.
5. Carvalho LAV de. Data mining. A mineração de dados no marketing, medicina, economia. 1a ed. Rio de Janeiro: Ciência Moderna; 2005. 256 p.
6. Rodriguez MVRY. Gestão empresarial em oranizações aprendizes: a arte de gerir mudanças. Rio de Janeiro: Qualitymark; 2002. 352 p.
7. Fayyad U, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery in databases. AI Mag. 1996;17(3):37–54.
8. Côrtes S da C, Porcaro RM, Lifschitz S. Mineração de Dados – Funcionalidades, Técnicas e Abordagens. Internet Rio Jan PUC. 2002;34.
9. Amaral F. Introdução à ciência de dados: mineração de dados e Big Data. Rio de Janeiro: Alta Books; 2016. 320 p.
10. Clésio F. Mineração de Dados com Software Livre [Internet]. Data Mining / Machine Learning / Data Analysis. 2012 [citado 21 de fevereiro de 2020]. Disponível em:

https://mineracaodedados.wordpress.com/2012/07/12/mineracao-de-dados-com-software-livre/
11. Brasil. Dispõe sobre o Estatuto da Criança e do Adolescente e dá outras providências [Internet]. Lei no 8.069 jul 13, 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8069.htm
12. Pyle D. Data Preparation for Data Mining. Morgan Kaufmann; 1999. 564 p.
13. Sharma A, Mansotra V. Emerging applications of data mining for healthcare management - A critical review. In: 2014 International Conference on Computing for Sustainable Global Development (INDIACom) [Internet]. New Delhi, India: IEEE; 2014 [citado 14 de março de 2019]. p. 377–82. Disponível em: http://ieeexplore.ieee.org/document/6828163/
14. Koh HC, Tan G. Data mining applications in healthcare. J Healthc Inf Manag. 2005;19(2):64–72.
15. Fialho FA, Dias IMÁV, Salvador M, Pacheco ZML, Nascimento L. A enfermagem avaliando a acuidade visual de estudantes do ensino fundamental. Rev Baiana Enfermagem25;2012. (1):16. 40–33. Rocha MNAM, Ávila MP de, Isaac DLC, Mendonça LS de M, Nakanishi L, Auad LJ. Prevalence of eye diseases and refractive errors in children seen at a referral center for ophthalmology in the central-west region, Brazil. Rev Bras Oftalmol. 2014;73(4):7.
17. Brasil. Manual de Certificação para Sistemas de Registro Eletrônico em Saúde (S-RES) versão 4.2 [Internet]. 2016 [citado 13 de março de 2019]. Disponível em: http://www.sbis.org.br/certificacao/Manual_Certificacao_SBIS-CFM_2016_v4-2.pdf

Anexos

Tabela 01. Agrupamento dos dados da variável "Refração"

	Dioptrias esféricas	Dioptrias cilíndricas
Plano	Zero	Zero
Miopia		
Leve	-0,25 a -3,00	Zero
Moderada	-3,25 a -6,00	Zero
Alta	≥ -6,25	Zero
Hipermetropia		
Leve	+0,25 a +3,00	Zero
Moderada	+3,25 a +6,00	Zero
Alta	≥ +6,25	Zero
Astigmatismo		
Leve	Zero	-0,25 a -1,00
Moderado	Zero	-1,25 a -3,00
Alto	Zero	≥ -3,25
Impossível verificar	Não passa faixa na esquiascopia	Não passa faixa na esquiascopia

Fonte: Prata AFG, Rodriguez MVRY, Barbosa ADM.

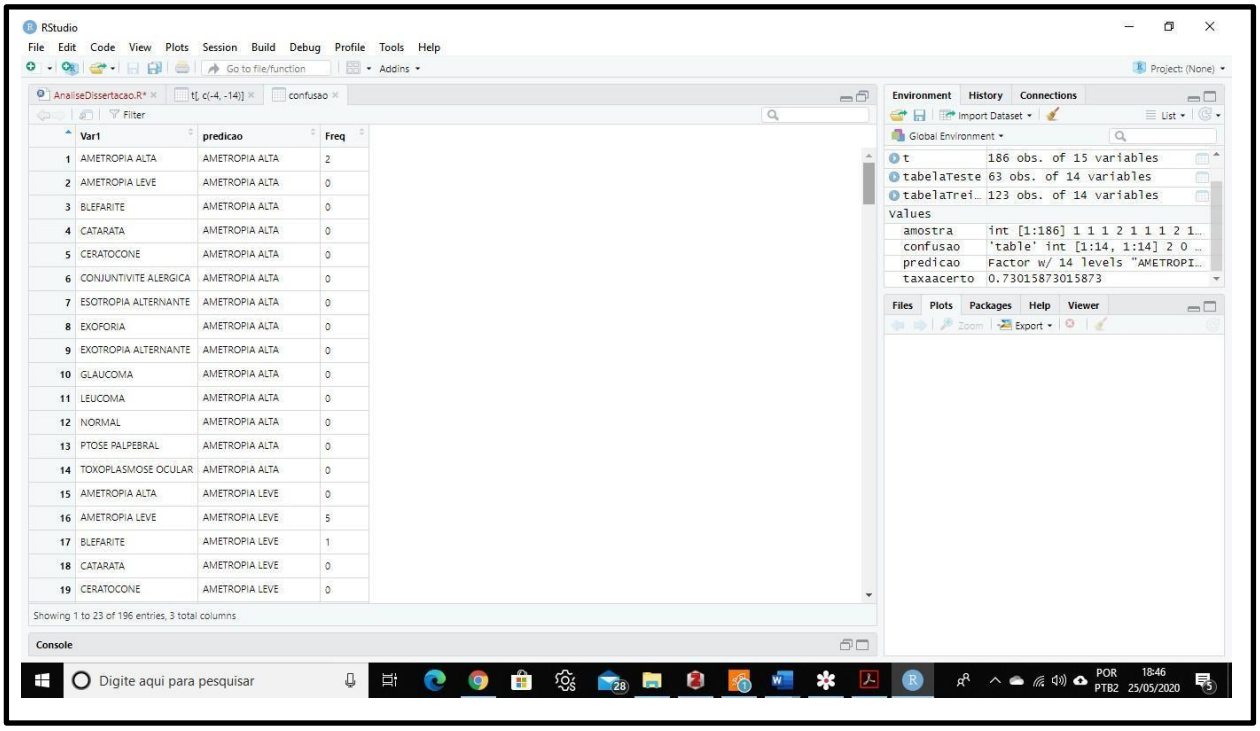
Tabela 2 - Frequências absoluta e relativa por sexo, raça, queixa principal, correção esférica, correção cilíndrica, diagnóstico e presença de ambliopia da amostra

	Frequência Absoluta	Frequência Relativa
Sexo		
Feminino	86	46,2
Masculino	100	53,8
Raça		
Branca	46	24,7
Parda	124	66,7
Preta	16	8,6
Queixa Principal		
Avaliação oftalmológica	37	19,9
Baixa acuidade visual	47	25,3
Cefaleia	26	14,0
Suspeita de Ceratocone	2	1,1
Dor ocular	2	1,1
Fotofobia	6	3,2
Suspeita de Glaucoma	2	1,1
Lesões ulceradas periocular	1	0,5
Olho torto	50	26,9
Pinta no olho	1	0,5
Prurido ocular	8	4,3
Queda da pálpebra	3	1,6
Trauma	1	0,5
Correção Esférica		

	Hipermetropia alta	6	3,2
	Hipermetropia leve	23	12,4
	Miopia alta	1	0,5
	Miopia leve	34	18,3
	Miopia Moderada	5	2,7
	Não passa faixa	3	1,6
	Plano	114	61,3
Correção Cilíndrica			
	Astigmatismo alto	1	0,5
	Astigmatismo leve	34	18,3
	Astigmatismo moderado	14	7,5
	Não passa faixa	3	1,6
	Plano	134	72,0
Diagnóstico			
	Ametropia alta	7	3,8
	Ametropia leve	28	15,1
	Blefarite	6	3,2
	Catarata	8	4,3
	Ceratocone	4	2,2
	Conjuntivite alérgica	8	4,3
	Esotropia alternante	23	12,4
	Exoforia	7	3,8
	Exotropia alternante	10	5,4
	Glaucoma	14	7,5
	Leucoma	3	1,6
	Normal	57	30,6
	Ptose palpebral	3	1,6
	Toxoplasmose ocular	8	4,3
Ambliopia			
	Não	158	84,9
	Sim	28	15,1
Total		186	100,0

Fonte: Prata AFG, Rodriguez MVRY, Barbosa ADM.

Figura 3 – Criação da Matriz Confusão pelo modelo.



Fonte: Prata AFG, Rodriguez MVRY, Barbosa ADM.