

João Vítor Faquineli Cavalcante Mendes de Ávila

**Análise paraconsistente e classificadores baseados em
conhecimento: identificando ceratocone subclínico**

São José do Rio Preto
2022

João Vítor Faquinelí Cavalcante Mendes de Ávila

**Análise paraconsistente e classificadores baseados
em conhecimento: identificando ceratocone
subclínico**

Monografia apresentada ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Bacharel

Orientador: Prof. Dr. Rodrigo Capobianco
Guido

São José do Rio Preto
2021

A958a

Ávila, João Vítor Faquinelí Cavalcante Mendes de
Análise paraconsistente e classificadores baseados em
conhecimento: identificando ceratocone subclínico / João Vítor
Faquinelí Cavalcante Mendes de Ávila. -- São José do Rio Preto, 2021
39 p.

Trabalho de conclusão de curso (Bacharelado - Ciência da
Computação) - Universidade Estadual Paulista (Unesp), Instituto de
Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientador: Rodrigo Capobianco Guido

1. Ceratocone. 2. Machine Learning. 3. Inteligência Artificial. I.
Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de
Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

JOÃO VÍCTOR FAQUINELI CAVALCANTE MENDES DE ÁVILA

**ANÁLISE PARACONSISTENTE E CLASSIFICADORES BASEADOS EM
CONHECIMENTO: IDENTIFICANDO CERATOCONES SUBCLÍNICO**

Monografia apresentada ao Departamento de
Ciências de Computação e Estatística do Instituto de
Biociências Letras e Ciências Exatas da Universidade
Estadual Paulista “Júlio de Mesquita Filho” para a
obtenção do título de Bacharel

Orientador: Prof. Dr. Rodrigo Capobianco Guido

DECLARAÇÃO DO AUTOR

Declaro, para fins de pesquisa acadêmica, didática e técnico-científica, que este Trabalho de Conclusão de Curso pode ser parcialmente utilizado, desde que se faça referência à fonte e ao autor.

São José do Rio Preto, 29 de setembro de 2018.

João Víctor Faquineli Cavalcante Mendes de Ávila

Dedico este trabalho à minha família e à memória dos meus jamais esquecido amigos Rodolfo Rodrigues Simões e Gabriel Willians Gonçalves.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por estar sempre presente e agindo com toda sua glória em minha vida. Agradeço à minha família, principalmente aos meus pais, que sempre me proporcionaram, apoiaram e incentivaram qualquer um dos meus projetos de vida, assim como este. Também agradeço ao meu orientador, prof. Dr. Rodrigo Capobianco Guido, que não só sendo um ótimo professor, me serviu de inspiração de caráter e filosofia. Ao Dr. Gildásio Castello, que prestou todo o apoio e os dados necessários para a continuação do estudo e, por último, mas não menos importante, agradeço ao meu amigo Jogi Suda, que esteve comigo desde os primeiros dias de faculdade, sempre me deu o apoio quando precisei, e que considero hoje como um irmão de outra mãe.

"If you're going to try, go all the way. Otherwise, don't even start." (BUKOWSKI, 1975)

RESUMO

A evolução da computação tem proporcionado a intercomunicação de diversas áreas do conhecimento, para criar novas soluções e paradigmas. É cada vez mais comum o uso da tecnologia para otimizar a evolução em qualquer âmbito, uma das vertentes que vem se mostrando cada vez mais aplicada nesse aspecto é o *Machine Learning*, que utilizando uma grande capacidade de processamento e diferentes métodos estatísticos, oferece uma gama de possibilidades para tratar diversos cenários de aplicação. Neste trabalho, interligamos a inteligência artificial à área de oftalmologia. Observamos um cenário onde existe um banco de dados de diversos pacientes que se submeteram ao exame de ceratocone, porém, devido à uma imprecisão do diagnóstico para determinados casos, onde há incerteza da necessidade do paciente passar por uma cirurgia, foi vista a possibilidade da aplicação do aprendizado de máquina para solvência desse óbice. Foram utilizados métodos classificadores baseados em árvore de decisão e em redes neurais para fazer a análise das características coletadas nos diagnósticos que, posteriormente, permitiram traçar, a partir de métricas pré-definidas, um paralelo de comparação entre os classificadores e elencar sua eficiência no contexto aplicado. Junto a isso, também foi aplicado o método de engenharia paraconsistente, que busca fazer um processo de avaliação nas características coletadas e utilizadas, descartando as que seriam menos convenientes para o cenário. Os classificadores foram aplicados antes e depois da análise paraconsistente, o que também nos permitiu traçar um paralelo de comparação e eficiência do método.

ABSTRACT

The evolution of computing has provided the intercommunication of different areas of knowledge to create new solutions and paradigms. It is increasingly common to use technology to optimize evolution in any context, one of the aspects that has been increasingly applied in this aspect is Machine Learning, which, using a large processing capacity and different statistical methods, offers a range of possibilities to deal with various different scenarios. In this project, we interconnect artificial intelligence to the field of ophthalmology. We observe a scenario where there is a database of several patients who underwent keratoconus examination, however, due to an inaccuracy of the diagnosis for certain cases, where there is uncertainty about the need for the patient to undergo surgery, the possibility of applying machine learning to solve this problem was seen. The classification methods based on decision tree and neural networks were used to analyze the characteristics collected in the diagnoses, which later allowed to draw, from pre-defined metrics, a parallel of comparison between the classifiers and to list their efficiency in the applied context. Alongside this, the paraconsistent engineering method was also applied, which seeks to do an evaluation process on the characteristics collected and used, discarding those that would be less convenient for the scenario. The classifiers were applied before and after the paraconsistent analysis, which also allowed us to draw a parallel of comparison and efficiency of the method.

LISTA DE TABELAS

Tabela 1	– Métricas por grupo anterior à engenharia paraconsistente . . .	30
Tabela 2	– Métricas por grupo posterior à engenharia paraconsistente . .	30
Tabela 3	– Acurácia antes e depois da engenharia paraconsistente respec- tivamente	30
Tabela 4	– Métricas por grupo anterior à engenharia paraconsistente . . .	30
Tabela 5	– Métricas por grupo posterior à engenharia paraconsistente . .	30
Tabela 6	– Acurácia antes e depois da engenharia paraconsistente respec- tivamente	31
Tabela 7	– Precisão em cada cenário por para cada algoritmo	31
Tabela 8	– Acurácia em cada cenário por para cada algoritmo	31
Tabela 9	– Sensibilidade em cada cenário por para cada algoritmo	32

LISTA DE FIGURAS

Figura 1 – Representação do algoritmo CART	17
Figura 2 – Representação do algoritmo MLP	18
Figura 3 – Representação do algoritmo ADAM	19
Figura 4 – Relação entre precisão e acurácia	23
Figura 5 – Representação da patologia ceratocone	25

LISTA DE SIGLAS

CART	– Classification And Regression Trees
MLP	– Multilayer Perceptron
TPR	– True Positive Rate
TNR	– True Negative Rate
PPV	– Positive Predicted Value
MVS	– Máquina de Vetor de Suporte
ACP	– Análise de Componente Principal

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Considerações iniciais	13
1.2	Objetivos	14
1.3	Organização do trabalho	14
2	REVISÃO BIBLIOGRÁFICA	15
2.1	Análise Paraconsistente	15
2.1.1	Introdução	15
2.1.2	Método	15
2.2	Classificador Árvore de decisão	16
2.3	Classificador Multi-layer Perceptron (MLP)	17
2.3.1	Algoritmo ADAM	19
2.4	Trabalhos correlatos	20
2.4.1	Classificações Pattern-matching e Knowledge-based de Sinais Bi- ológicos: identificando casos sub-clínicos	20
2.4.2	Use of a Support Vector Machine for Keratoconus and Subcli- nical Keratoconus Detection by Topographic and Tomographic Data	20
2.4.3	Evaluation of machine learning classifiers in keratoconus detec- tion from orbscan II examinations	21
2.4.4	Comparação aos trabalhos correlatos	21
2.5	Métricas utilizadas	22
2.5.1	Sensibilidade	22
2.5.2	Precisão	22
2.5.3	Acurácia	22
3	DESENVOLVIMENTO	24
3.1	Especificações técnicas	24
3.2	Arcabouço Metodológico	24
3.2.1	Coleta da base	24
3.2.2	Categorização	25
3.2.3	Aplicação dos algoritmos	27
4	RESULTADOS	29
4.1	Engenharia paraconsistente	29
4.1.1	Árvore de decisão anterior x posterior à engenharia paraconsis- tente	30

4.1.2	MLP anterior x posterior à engenharia paraconsistente	30
4.2	Comparação com o trabalho base	31
5	CONCLUSÃO	33
	REFERÊNCIAS	34

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

O ceratocone, termo derivado a partir das palavras gregas *kerato* (córnea) e *conus* (cone), é uma doença degenerativa, não inflamatória, bilateral e assimétrica, progressiva, que leva às inúmeras alterações na superfície da córnea. Foi inicialmente descrita por Benedict Dudell, cerca de três séculos atrás, porém foi John Nottingham, em 1854, que trouxe uma descrição mais abrangente da patologia (TUFT et al., 1994). De acordo com um estudo de 1994 com 2.723 pacientes, a doença ocorre geralmente no final da puberdade e em adultos jovens, com idade média de 22,4 anos (TUFT et al., 1994). Outro estudo realizado em 1961 concluiu que o desenvolvimento do ceratocone ocorre entre os 10 e 20 anos de idade e é uma doença que acomete cerca de 1 a cada 1000 pessoas. (ELIAS et al., 2005) Nem todos os casos de ceratocone se desenvolvem da mesma maneira. Ao que tudo indica, não há nenhum agente que influencie na progressão do ceratocone (ELIAS et al., 2005). Essencialmente, o ceratocone provoca mudanças na estrutura da córnea, devido a um enfraquecimento do tecido, tornando-a mais fina e alterando a sua curvatura, tendendo a um formato mais cônico se comparado à córnea normal. Consequentemente, a imagem que passa pela córnea sofre uma distorção, resultando em alterações na visão de pacientes com ceratocone, o que os leva a desenvolver doenças como miopia e astigmatismo, ambos erros refrativos (IMBORNONI; MCGHEE; BELIN, 2018). O enfraquecimento do tecido da córnea do ceratocone está relacionado ao desequilíbrio das enzimas na córnea. Esse desequilíbrio torna a córnea mais suscetível a danos oxidativos por compostos chamados radicais livres, fazendo com que ela enfraqueça e inche. Essa predisposição está relacionada a fatores genéticos, por isso é muito comum a doença afetar mais de um membro de uma família (IMBORNONI; MCGHEE; BELIN, 2018).

O tratamento inicial para combater a hipovisão causada pela ceratocone é por meio de óculos, quando este não é suficiente, deve-se migrar para lentes especiais para ceratocone, conhecidas como rígidas ou semirrígidas, o que irá depender do estágio da doença em cada paciente (TAI et al., 2013). Caso seja necessária uma intervenção maior, o tratamento moderno para impedir a progressão da ceratocone é o tratamento baseado em riboflavina atuando junto com a radiação UV, que gera enrijecimento molecular da córnea (POLLHAMMER; CURSIEFEN, 2009). O diagnóstico mais comum da ceratocone se dá por meio de aparelhos de topografia da córnea, um dos mais modernos, utilizado para coletar os dados deste trabalho, é o Pentacam, que, ao criar mapas tridimensionais da córnea do paciente, permite tirar a medida de vários parâmetros importantes para a avaliação do diagnóstico final (HASHEMI et al., 2016). Dado isso, conseguimos traçar uma estratégia envolvendo machine learning para fazer um tratamento dos dados extraídos pelos aparelhos de diagnósticos, aplicando métodos classificadores juntos a análise paraconsistente para

auxiliar, principalmente os casos subclínicos, que são casos que carecem de algum tipo de informação para afirmar que determinado paciente que apresenta indícios de ter ceratocone realmente foi acometido por ela.

1.2 OBJETIVOS

A partir do contexto apresentado anteriormente, este trabalho busca fazer um tratamento e uma avaliação comparativa sob a óptica de 2 algoritmos classificadores, dos dados extraídos pelo aparelho Pentacam de exames oftalmológicos que buscam detectar o acometimento da ceratocone. Os algoritmos utilizados serão o classificador de árvore de decisão e o classificador Multilayer Perceptron (MLP). Junto a eles, será aplicada a análise paraconsistente e um posterior comparativo dos resultados obtidos antes e depois de seu uso.

1.3 ORGANIZAÇÃO DO TRABALHO

Os tópicos subsequentes estão organizados da seguinte forma:

- Capítulo 2: Revisão Bibliográfica que traz os principais conceitos para o desenvolvimento do trabalho, apresentando os algoritmos utilizados em todo o estudo.
- Capítulo 3: Demonstração do desenvolvimento do projeto, apresentando e contextualizando especificações técnicas, tecnologias e métodos utilizados.
- Capítulo 4: Apresentação dos testes e resultados do projeto.
- Capítulo 5: Conclusão.

2 REVISÃO BIBLIOGRÁFICA

Nesta seção, são apresentados conceitos essenciais para a compreensão deste trabalho. Os tópicos definem e descrevem a análise paraconsistente, o classificador de árvore de decisão junto ao algoritmo auxiliar CART, e o classificador de perceptron de multicamadas, junto ao seu algoritmo auxiliar ADAM.

2.1 ANÁLISE PARACONSISTENTE

2.1.1 Introdução

A análise paraconsistente é um método que utiliza uma lógica não clássica (lógica paraconsistente) para, essencialmente, ponderar a conveniência de determinadas classes de um N-problema que possui N-classes. Para isso o método utiliza 2 abordagens para levantar 2 parâmetros independentes, por isso, são não complementares, o que justifica o uso da lógica paraconsistente. A análise intraclasse, que classifica a semelhança de características dentro de uma mesma classe, a qual é representada por $0 \leq \alpha \leq 1$, e a análise interclasse, que classificaram a heterogeneidade de determinada característica nas diferentes classes, representada por $0 \leq \beta \leq 1$ (GUIDO, 2018).

2.1.2 Método

Primeiramente é necessário a normalização dos vetores de características para que então possa se realizar as análises. A abordagem da intraclasse, que visa buscar a semelhança entre as características, utiliza um método de distribuir essas características ao longo de um eixo, o qual será usado como parâmetro para comparar a dispersão das mesmas, utilizando a medida de desvio $Y = 1 - A$, na qual A se valerá da amplitude entre essas observações. Posteriormente, é calculada a média dos valores de cada ponto, repetindo-se o processo para todas as classes, para finalmente chegarmos ao valor de α . É considerado o menor valor de $Y(C_i)$, visto que α representa o pior caso (GUIDO, 2018).

$$(\min \bar{Y}(C_1), \bar{Y}(C_2) \dots \bar{Y}(C_M), \text{ onde } \bar{Y}(C_i) = \frac{1}{n} \sum_{k=1}^{n-1} svC_i[k])$$

Na análise interclasse, é buscado a dissimilaridade entre as diferentes classes. Ela se inicia construindo 2 vetores de alcance de tamanho T para cada classe, um com o elemento mínimo e outro com o máximo para cada característica. Em sequência, os vetores de alcance serão comparados com os vetores de características de todas as classes, assim como seus respectivos vetores de alcance, que nos resultará em um parâmetro R, que essencialmente, será o número de sobreposições encontradas de F possíveis. Sendo assim, chegaremos em:

$$\beta = R/F$$

Supondo que teremos um número de classes M , contendo X vetores com n características, teremos:

$$F = M \cdot (M - 1) \cdot X \cdot n$$

2.2 CLASSIFICADOR ÁRVORE DE DECISÃO

Árvores de decisão, são métodos estatísticos de aprendizado de máquina com treinamento supervisionado, podendo ser utilizado como classificação e regressão; no caso usaremos como classificador neste trabalho. É um dos algoritmos de inferência mais populares, tendo um abrangente campo de aplicação, apresentando uma estrutura simples e eficiente (SWAIN; HAUSKA, 1977). Utilizando uma estrutura onde nós guardam determinadas informações, que se subdividem em outros nós chamados “filhos” a partir de uma regra pré-definida, este modelo é sempre referenciado pela sua técnica de dividir e conquistar, o qual transforma um problema maior em mais complexo em vários problemas menores e mais simples, sendo estes resolvidos recursivamente até chegar à resolução principal (SILVA, 2005). O modelo de classificação por árvore de decisão, se sustenta na abordagem de fornecer treinamentos que já possuem uma caracterização e por isso são conhecidos. Com base nesses treinamentos, são criados os modelos classificadores que serão aplicados aos conjuntos de teste, que consistem em registros com rótulos de classes ainda não conhecidas (BENTO; KUX; KÖRTING,). Existem diversos algoritmos tanto para classificação ou regressão para árvore de decisão, neste trabalho, utilizamos o algoritmo CART (Classification and Regression Trees), que também é um dos algoritmos mais populares nesse âmbito por apresentar grande adaptabilidade para lidar com muitos parâmetros de entrada. O algoritmo CART funciona de maneira que, em primeiro momento, ele elenca um conjunto de regras que, baseados no coeficiente de Gini, sendo que a pergunta obtém apenas resposta positiva ou negativa, e assim, define qual será a melhor divisão do nó pai baseada na variável dependente, em seus 2 nós filhos. O processo se repete nos nós filhos até que o processo seja interrompido ou impossível de se dar continuidade. E assim se finaliza com o processo de poda, que buscará a árvore que apresenta menor custo, este que leva em consideração a menor complexidade e taxa de erro (SAFAVIAN; LANDGREBE, 1991).

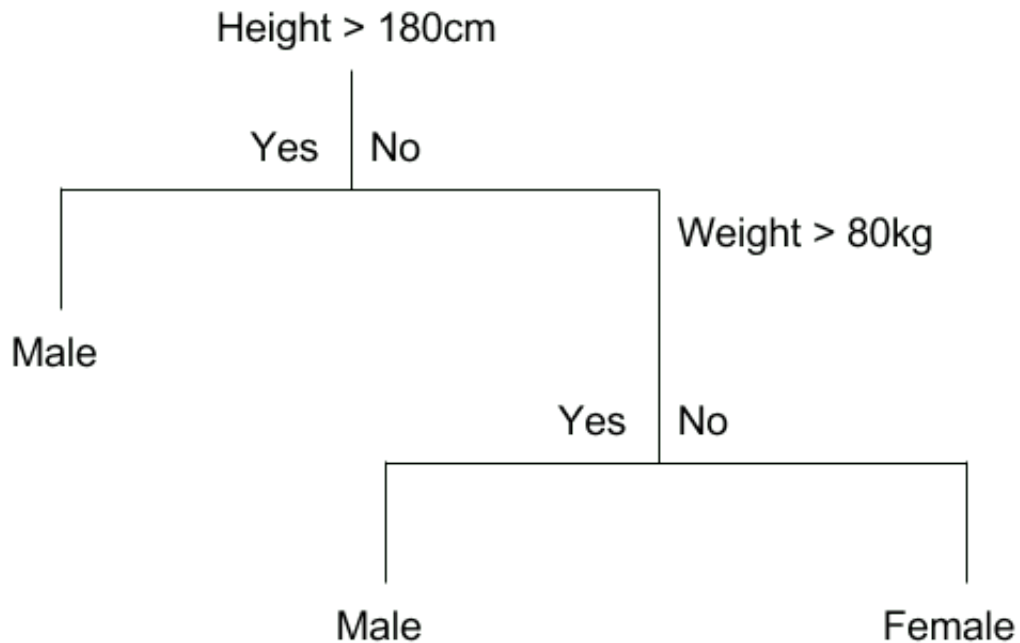


Figura 1 – Representação do algoritmo CART

O Coeficiente de Gini é representado pela equação:

$$i(t) = \sum_{i \neq j} p(i/t)p(j/t)$$

Onde $p(i/t)$ é a probabilidade de uma amostra aleatória X pertencer a classe i , dado que estamos no nó t .

2.3 CLASSIFICADOR MULTI-LAYER PERCEPTRON (MLP)

Multilayer Perceptron (MLP) é um algoritmo de aprendizado supervisionado, que possui 3 ou mais camadas, dispostas em: uma camada de entrada e uma camada de saída com uma ou mais camadas não lineares (ocultas) entre as anteriores. A camada de entrada possui os primeiros neurônios a receberem os valores e os repassarem para a camada oculta, que transforma os mesmos em uma somatória linear com um fator multiplicador para cada resultado e, em sequência, é passada por uma função de ativação não linear. A camada de saída receberá os valores da última camada oculta e os transformará em um resultado (MIA et al., 2015).

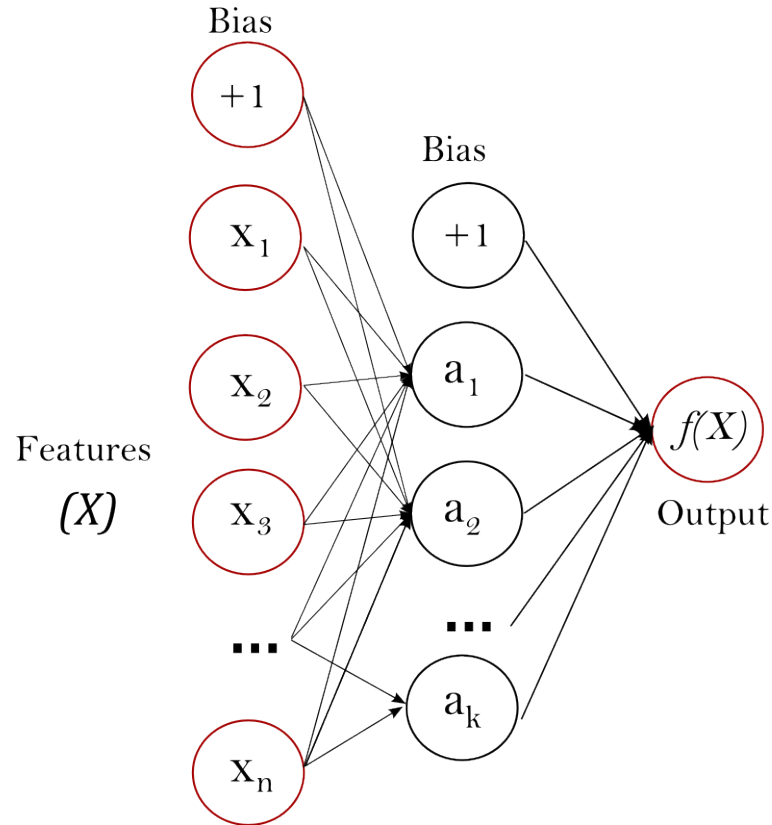


Figura 2 – Representação do algoritmo MLP

O processo se inicia com o recebimento dos valores de entrada da primeira camada e suas correspondentes saídas. Os neurônios da primeira camada geram uma resposta para a camada oculta subsequente, que dá continuidade ao processo para as próximas camadas, se houver, até a camada de saída. A resposta obtida após todo o processo, é comparada com a resposta que era esperada, e a partir disso, é calculado o erro. Conforme a diferença do erro do neurônio de saída, é calculado um parâmetro que será a taxa de mudança do neurônio conforme sua atividade também se altera. Todo esse procedimento foi feito pelo caminho de entrada até a saída, logo após ele ser finalizado, ele recua, refazendo os cálculos em cada camada e alterando os pesos de cada parâmetro, o que busca minimizar o erro da saída (RUDER, 2016). O algoritmo segue essa sequência até chegar novamente na camada de entrada. Quando ele a atinge novamente e os pesos não são alterados, ele atinge um estado estacionário, fazendo-se necessário selecionar os próximos padrões de entradas-alvo, repetindo-se todo o processo novamente. No estudo em questão, O MLP foi configurado para ter 2 camadas ocultas, a primeira tendo 25 e a segunda 10 neurônios. Apesar de os resultados moverem-se para frente, os pesos são calculados se movendo pelo caminho contrário, diante disso, tem-se o nome de algoritmo de retro propagação (MIA et al., 2015).

O algoritmo de retropropagação de erro implementado para o treinamento do MLP, no caso, foi o ADAM, um algoritmo para otimização baseada em gradiente

de primeira ordem de funções objetivo estocásticas. O método utiliza o momento, que essencialmente é o aumento da aceleração do gradiente estocástico na descida e também o seu controle de oscilação conforme seja conveniente. O algoritmo também conta com a taxa de aprendizagem adaptativa, que calcula, para cada parâmetro, um valor individualmente que vai se adaptando conforme o algoritmo é executado.

2.3.1 Algoritmo ADAM

ADAM foi implementado com o intuito principal de combinar as vantagens de dois métodos que eram e ainda são bastante utilizados (AdaGrad e RMSProp), sendo que o primeiro se destaca para operações com gradientes mais esparsos, e o segundo para configurações não-estacionárias. Com a junção desses diferenciais, ADAM surgiu como um dos algoritmos mais populares quando se trata de otimização de redes neurais (KINGMA; BA, 2014). Sebastian Ruder, publicou em seu artigo “An overview of gradient descent optimization algorithms”, que:

“Na medida em que, RMSprop, Adadelta e Adam são algoritmos muito semelhantes que funcionam bem em circunstâncias semelhantes. [...] Sua correção de viés ajuda Adam a superar ligeiramente o RMSprop no final da otimização, conforme os gradientes se tornam mais esparsos. Na medida em que Adam pode ser a melhor escolha geral.” (KINGMA; BA, 2014).

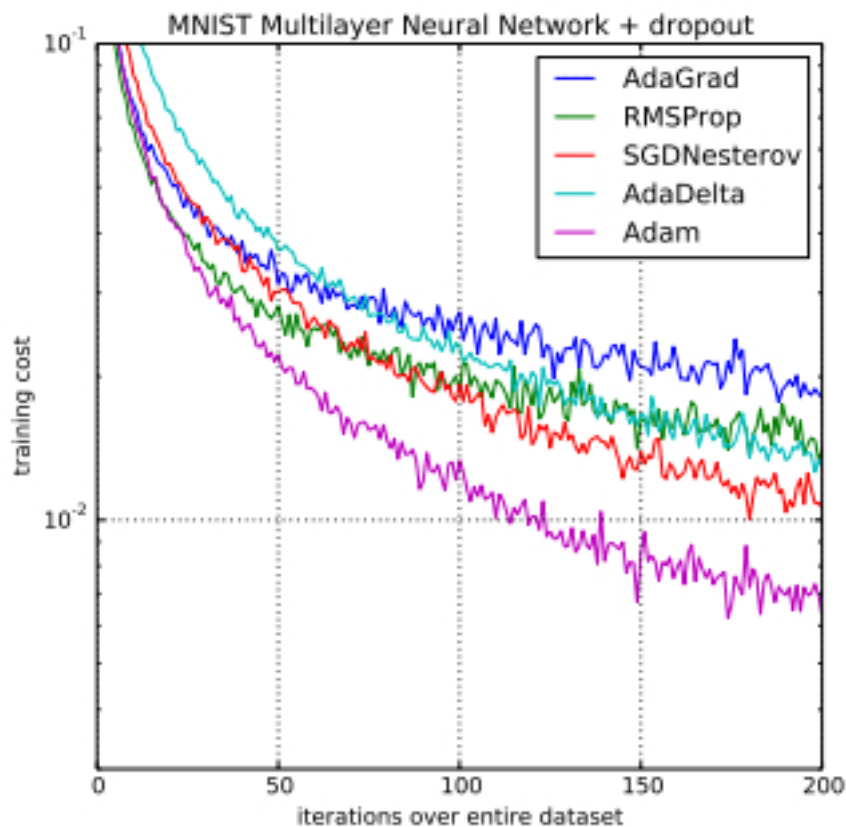


Figura 3 – Representação do algoritmo ADAM

A figura 3 elenca a comparação entre os diversos métodos, demonstrando uma eficiência perceptível para o algoritmo ADAM.

2.4 TRABALHOS CORRELATOS

2.4.1 Classificações Pattern-matching e Knowledge-based de Sinais Biológicos: identificando casos sub-clínicos

Este trabalho foi feito como uma forma de continuação do trabalho "Classificações Pattern-matching e Knowledge-based de Sinais Biológicos: identificando casos sub-clínicos" (NETO, 2018), que utiliza uma base de dados de pacientes examinados pelo mesmo médico e pelos mesmos aparatos. Foi iniciado o estudo de métodos classificadores e análise paraconsistente em seu trabalho, em que no caso foram utilizados os classificadores K-Vizinhos mais próximos (*K-Nearest Neighbors* - KNNs) e Máquina de Vetor de Suporte (MVS).

Ademais, foi incluído como comparação o método ACP, que essencialmente é um método estatístico para traçar um paralelo com os resultados obtidos pelos classificadores. Também foi utilizada a engenharia paraconsistencia associada aos classificadores.

Além dos classificadores diferentes, neste trabalho eliminaremos método ACP e traçaremos os resultados obtidos pelo primeiro trabalho como parâmetro de comparação. Também há uma base de dados com mais pacientes em todos os grupos, o que permitirá um melhor treinamento para os classificadores.

2.4.2 Use of a Support Vector Machine for Keratoconus and Subclinical Keratoconus Detection by Topographic and Tomographic Data

Outro trabalho que apresenta bastante semelhança, Use of a Support Vector Machine for Keratoconus and Subclinical Keratoconus Detection by Topographic and Tomographic Data (ARBELAEZ et al., 2012), foi também proposto um classificador que fosse aplicado em exames de ceratocone, com foco aos casos subclínicos, gerados por um aparelho clínico semelhante ao Pentacam, chamado genericamente de câmera de Scheimpflug, junto a outro aparelho de análise de córnea, que faz diagnósticos também semelhantes ao Pentacam. Neste trabalho, foi utilizada uma base de dados de 877 olhos com ceratocone, 426 olhos com subclínica ceratocone, 940 olhos com história de cirurgia da córnea (definida como anormal) e 1259 olhos de controle saudáveis. Esses dados foram coletados da base de dados de Muscat Eye Laser Center, uma clínica oftalmológica de Omã. Em questão, o foco foi um único classificador, o Máquina de Vetor de Suporte (MVS), que apresentou resultados muito satisfatórios, a partir das mesmas métrica utilizadas neste

trabalho. O autor dividiu a análise em duas partes, sendo a primeira utilizando apenas os parâmetros fornecidos pela câmera de *Scheimpflug*, e depois, foi adicionado os parâmetros extras da análise da córnea para comparar os dois cenários. Ambos conseguiram atingir uma alta sensibilidade, sendo, seguindo o primeiro e segundo cenário, respectivamente, 89,3% a 96,0% em olhos anormais, 92,8% a 95,0% em olhos com ceratocone, 75,2% a 92,0% em olhos com ceratocone subclínico e 93,1% para 97,2% em olhos normais. Além disso, também foram encontrados bons resultados para as outras métricas, o que demonstrou que o algoritmo MVS aplicado aos dados dos diagnósticos de ceratocone, entrega um resultado bastante acertado sobre as condições do paciente.

2.4.3 Evaluation of machine learning classifiers in keratoconus detection from orbscan II examinations

No trabalho *Evaluation of machine learning classifiers in keratoconus detection from orbscan II examinations* (SOUZA et al., 2010), os autores propuseram utilizar os algoritmos Máquina de Vetor de Suporte (MVS), função de base radial e perceptron multicamadas (MLP) para identificar pacientes com ceratocone. O trabalho utiliza os dois métodos em uma base de dados de 172 pacientes normais, 89 pacientes com astigmatismo e 57 pacientes com ceratocone. Essa base de dados foi gerada a partir dos exames clínicos feitos pelo orbscan II, um aparelho que semelhante ao Pentacam, faz análises e métricas de diversos parâmetros da córnea, porém, é um aparelho menos moderno. A partir do processamento dos dados, utilizando a Sensibilidade, métrica também utilizada neste trabalho, eles chegaram em um resultado que ambos os algoritmos tem desempenho melhor do que a análise individual de cada atributo fornecido pelo Orbscan. Além disso, também foi demonstrado que há uma diferença ínfima entre o desempenho dos três algoritmos. Sendo assim, conclui-se que atrelar as técnicas de machine learning às análises e exames tradicionais, pode obter-se um resultado mais confiável e preciso.

2.4.4 Comparação aos trabalhos correlatos

Todos os trabalhos se assemelham muito em seus segmentos e suas metodologias. Foi utilizado um ou mais classificador baseado em *machine learning* que foi aplicado em uma base de dados, coletada por aparelhos tradicionais da área oftalmológica, que tiram medidas dos olhos e principalmente da córnea do paciente. Os trabalhos começam a ter suas primeiras diferenciações nesse momento, uma vez que nem todos utilizaram os mesmos aparelhos, consequentemente, não utilizando os mesmos parâmetros. Por conseguinte, além da base de dados diferentes, cada trabalho tratou um ou mais algoritmos para serem analisados, além de que, apesar de apresentarem algumas métricas diferentes, todos utilizam a sensibilidade como métrica principal para a valoração do desempenho do classificador, neste trabalho foi seguido o mesmo padrão. Este trabalho será melhor

comparado ao trabalho Classificações Pattern-matching e Knowledge-based de Sinais Biológicos: identificando casos sub-clínicos (NETO, 2018), por apresentar a mesma fonte da base de dados, apesar de dados diferentes, com parâmetros semelhantes e pacientes que foram examinados pela mesma máquina, o Pentacam. Além de aplicarem o método de análise paraconsistente. Por isso, são utilizadas também as mesmas métricas, para que seja possível traçar um melhor padrão de comparação.

2.5 MÉTRICAS UTILIZADAS

As métricas utilizadas neste trabalho seguirão o mesmo padrão das utilizadas no trabalho inicial que inspirou a elaboração deste trabalho, para que assim, seja possível traçar um critério de comparação mais fidedigno e confiável sobre os diferentes classificadores aplicados.

2.5.1 Sensibilidade

Este parâmetro serve como forma de medir a capacidade de determinado classificador acertar casos patológicos. Pode ser considerado o mais relevante no caso prático devido ao fato de que é buscado alertar o paciente que ele precisa da cirurgia.

A sensibilidade, ou *TruePositiveRate* (TPR), é definida por:

$$TPR = \frac{TP}{TP+TN}$$

2.5.2 Precisão

A precisão é um parâmetro que mede a capacidade do classificador de manter seu padrão de acertos.

A precisão, ou *PositivePredictedValue*, é definida por:

$$PPV = \frac{TP}{TP+FP}$$

2.5.3 Acurácia

A acurácia (ACC) é o parâmetro utilizado para calcular a taxa de acerto do classificador com relação ao todo.

A acurácia, é definida por:

$$ACC = \frac{TN+TP}{TP+TN+FP+FN}$$

A figura 4 demonstra uma comparação mais visual da relação entre a acurácia e a precisão:

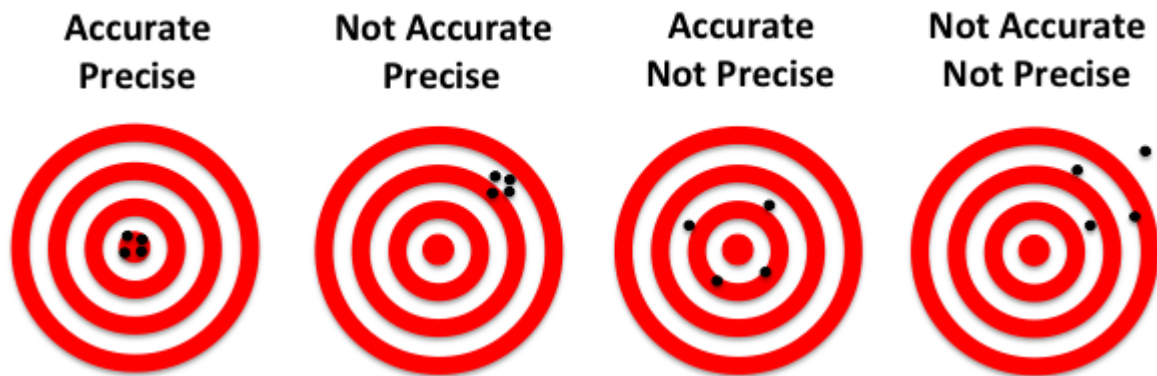


Figura 4 – Relação entre precisão e acurácia

3 DESENVOLVIMENTO

Neste capítulo, será apresentada todo o arcabouço clínica de coleta de dados, incluindo os aparelhos utilizados e as características extraídas nos exames, assim como a classificação dos subgrupos de pacientes. Da mesma forma, serão apresentadas as especificações da máquina que será usada para o processamento desses dados junto aos classificadores, assim como alguns detalhes da implementação utilizada.

3.1 ESPECIFICAÇÕES TÉCNICAS

A implementação do código para o tratamento desta base de dados será *Python*, que apesar de ser uma linguagem substancialmente mais lenta por ser alto nível, apresenta uma série de bibliotecas que auxiliam na implementação dos classificadores, que no caso será a árvore de decisão e o *MLP*, e apresenta uma robustez que é necessária para trabalhos mais complexos. As principais bibliotecas para implementação do código são a *scikitlearn*, que oferecem diversos métodos vinculados à *Machine learning*, e a *pandas* que nos possibilita vincular o arquivo *.xlsx* da base de dados de pacientes ao código.

3.2 ARCABOUÇO METODOLÓGICO

Nesta seção, são apresentadas as informações que representam a base do trabalho, desde a forma de coleta da base de dados, as características selecionadas e a aplicação dos algoritmos.

3.2.1 Coleta da base

Quatrocentos e noventa e sete pacientes patológicos compareceram ao Hospital de Base / Faculdade de Medicina de São José do Rio Preto e no Visum Centro de Olhos (VCO) no período de janeiro de 2012 a abril de 2021. Dois mil duzentos e noventa e seis pacientes com córneas sadias foram selecionados no VCO, com córneas que submetidas a LASIK ou PRK com grau estável com pelo menos 18 meses de acompanhamento. Concomitante, não apresentavam alteração na elevação posterior no exame pentacam e com resultado refracional estável programado atingido com seguimento de pelo menos 18 meses em relação ao prévio a cirurgia. Os seguintes critérios de exclusão também foram respeitados; história de trauma ocular; uso crônico de medicação ocular, principalmente paciente glaucomatoso, doença do olho seco, cicatriz corneana, ceratopatia neurotrófica, disfunção severa das glândulas de meibomius; população vulnerável devido à enfermidade física ou mental e com dificuldade de linguagem. Este estudo aderiu aos princípios da Declaração de Helsinque e foi aprovada pelo Comitê de Ética em Pesquisa da Faculdade de Medicina em São José do Rio Preto (FAMERP – #1.123.652). Todos os pacientes

foram informados sobre os objetivos deste estudo e assinaram formulários escritos de consentimento antes de participarem.

3.2.2 Categorização

A base de dados do trabalho é composta por 2894, os quais foram separados em 3 classes, as mesmas categorizadas no projeto base, sendo elas:

- Grupo controle (A): Grupo composto por 2296 pacientes normais.
- Grupo clínico (KC): Grupo composto por 410 pacientes com ceratocone clínico.
- Grupo subclínico (VAE-NT): Grupo composto por 187 pacientes com aspecto em topografia normal, mas com algumas indicações de ceratocone clínico.

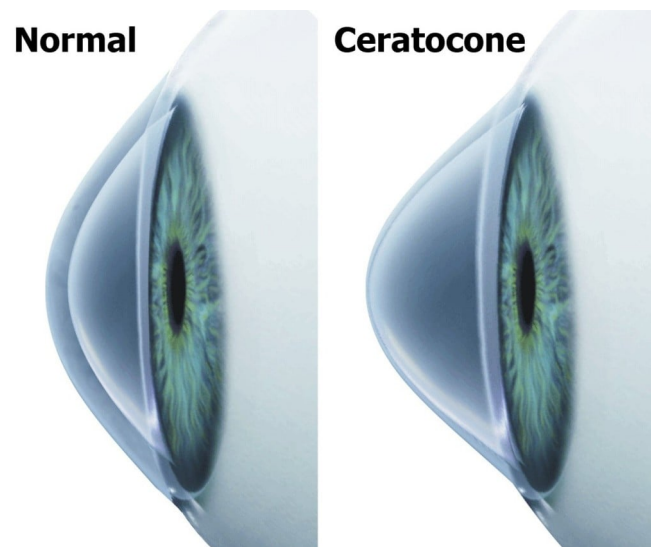


Figura 5 – Representação da patologia ceratocone

Além disso, serão utilizadas vinte e uma características extraídas e utilizadas na base de dados vinculadas à cada um dos pacientes. Sendo elas:

1. Z 3 -1 (CB): índice que representa a aberração de coma vertical de terceira ordem da parte posterior da córnea.
2. RMS Z3,4,5,6 CB: Raiz quadrada média dos seguintes polinômios de Zernike: Z3 3, Z3 1, Z3-1, Z3-3, Z4 4, Z4 2, Z4-4, Z4-4, Z5 5, Z5 3, Z5 1, Z5-1, Z5-3, Z5-5, Z6 6, Z6 4, Z6 2, Z6-2, Z6-4, Z6-6 da parte posterior da córnea, observe que Z4 0 e Z6 0 foram excluídos.

3. RMS Z3,4,5,6 CF: Raiz quadrada média dos seguintes polinômios de Zernike: Z3 3, Z3 1, Z3-1, Z3-3, Z4 4, Z4 2, Z4-4, Z4-4, Z5 5, Z5 3, Z5 1, Z5-1, Z5-3, Z5-5, Z6 6, Z6 4, Z6 2, Z6-2, Z6-4, Z6-6 da frente da córnea, observe que Z4 0 e Z6 0 foram excluídos.
4. Z 3 -1 (CF): Aberração coma vertical de terceira ordem da frente da córnea.
5. IHD: Índice de descentralização mais alta.
6. Z31Cornea-A: Aberração de coma vertical de terceira ordem córnea total.
7. IVA: Índice de assimetria vertical.
8. Ele B BFTE 8mm Thinnest: Parâmetro de elevação derivado da superfície posterior centrada no ponto mais fino usando o elipsoide tórico de 8 mm de melhor ajuste.
9. RMSHOACB: Raiz quadrada média da aberração de alta ordem da parte posterior da córnea.
10. RMSHOACF: Raiz quadrada média da aberração de ordem superior da frente da córnea.
11. ARTMax: Espessura relacional de Ambrósio máxima.
12. AGE NORMAIS: A idade de cada paciente na data do exame.
13. D2.4mm Pachy/PF: Quociente de espessura média da córnea de um círculo de 2,4 mm centrado no local mais fino / Pachy Min (espessura da córnea no ponto mais fino).
14. D2.8mm Pachy/PF: Quociente da espessura média da córnea de um círculo de 2,8 mm centrado no local mais fino / Pachy Min.
15. D3.2mm /PF: Quociente da espessura média da córnea de um círculo de 3,2 mm centrado no local mais fino / Pachy Min.
16. C.Vol. D 2.5mm: Volume da córnea com área de 2,5 mm de diâmetro.
17. C.Vol. D 3.0mm: Volume da córnea com área de 3,0 mm de diâmetro.
18. SD DA MÉDIA DOS RC ORTOGONIAS NO DIAMTERO DE 6,0 MM: DP dos valores médios de CP ortogonais no diâmetro de 6 mm da parte posterior da córnea centrada no ápice corneano. Os valores de CP avaliados foram: r0,0 0 °, r1,0 ((0 ° + 90 °) / 2, (45 ° + 135 °) / 2, (180 ° + 270 °) / 2, (225 ° + 315 °) / 2), r2,0 ((0 ° + 90 °) / 2, (30 °

- + 120 °) / 2, (60 ° + 150 °) / 2, (180 ° + 270 °) / 2, (210 ° + 300 °) / 2, (240 ° + 330 °) / 2) r3,0 ((0 ° + 90 °) / 2, (23 ° + 113 °) / 2, (45 ° + 135 °) / 2, (68 ° + 157 °) / 2, (180 ° + 270 °) / 2, (202 ° + 292 °) / 2, (225 ° + 315 °) / 2, (247 ° + 338 °) / 2). O valor do CP foi expresso em mm.
19. DIFERENÇA MAX-MIN MÉDIA DOS RC ORTOGONAIS ATE 6,0 MM: Diferença do valor máximo da CP ortogonal média menos o valor mínimo da CP ortogonal média no diâmetro de 6 mm da parte posterior da córnea centrada no ápice corneano. Os valores de CP avaliados foram: r0,00 °, r1,0 ((0 ° + 90 °) / 2, (45 ° + 135 °) / 2, (180 ° + 270 °) / 2, (225 ° + 315 °) / 2), r2,0 ((0 ° + 90 °) / 2, (30 ° + 120 °) / 2, (60 ° + 150 °) / 2, (180 ° + 270 °) / 2, (210 ° + 300 °) / 2, (240 ° + 330 °) / 2) r3,0 ((0 ° + 90 °) / 2, (23 ° + 113 °) / 2, (45 ° + 135 °) / 2, (68 ° + 157 °) / 2, (180 ° + 270 °) / 2, (202 ° + 292 °) / 2, (225 ° + 315 °) / 2, (247 ° + 338 °) / 2). O valor do CP foi expresso em mm.
20. SD DOS RC 6,00 MM DE DIAMTERO: Desvio padrão (DP) dos valores de CP no diâmetro de 6 mm da parte posterior da córnea centrada no ápice corneano. Os locais de CP avaliados foram: r0,0 0 °, r1,0 (0 °, 45 °, 90 °, 135 °, 180 °, 225 °, 270 °, 315 °), r2,0 (0 °, 30 °, 60 °, 90 °, 120 °, 150 °, 180 °, 210 °, 240 °, 270 °, 300 °, 330 °) r3,0 (0 °, 23 °, 45 °, 68 °, 90 °, 113 °, 135 °, 157 °, 180 °, 202 °, 225 °, 247 °, 270 °, 292 °, 315 °, 338 °). O valor do CP foi expresso em mm.
21. DIFEERENÇA MAX-MIN DOS DIAMETRO 6 MM: Diferença do valor máximo de potência da córnea (CP) menos o valor mínimo de CP no diâmetro de 6 mm do dorso da córnea centrado no ápice corneano. Os locais de CP avaliados foram: r0,0 0 °, r1,0 (0 °, 45 °, 90 °, 135 °, 180 °, 315 °, 225 °, 270 °), r2,0 (0 °, 30 °, 60 °, 90 °, 120 °, 150 °, 180 °, 210 °, 240 °, 270 °, 300 °, 330 °, 315 °) r3,0 (0 °, 23 °, 45 °, 68 °), 90 °, 113 °, 135 °, 157 °, 180 °, 202 °, 225 °, 247 °, 270 °, 292 °, 338 °). O valor do CP foi expresso em mm.
22. KI/P+F: Quociente de Pachy Min / KI (índice de ceratocone).

3.2.3 Aplicação dos algoritmos

Após a extração das características por meio do Pentacam, o próprio equipamento popula, em uma tabela no formato *xlsx*, todos os dados coletados dos respectivos pacientes. Utilizando a biblioteca *pandas* da linguagem *Python*, a tabela é extraída para o início do tratamento dos dados.

O primeiro passo é fazer o balanceamento dos 3 grupos, uma vez que há uma discrepância substancial entre eles, principalmente sob o grupo dos não patológicos (A), o

que pode acarretar em um menor desempenho dos algoritmos utilizados para aplicar o machine learning. A lógica usada para este tratamento é simples, sendo apenas utilizada uma função, também da biblioteca pandas, para recolher uma amostra estratificada de cada grupo com valores iguais, que será o valor do menor grupo (VAE-NT) que contém 187 indivíduos.

Após o tratamento dos dados, foi separado o grupo de teste e treinamento, sendo utilizado um valor de 25% para a separação do grupo de treinamento. Em sequência, foram aplicados os algoritmos MLP e Árvore de decisão.

O MLP, como mencionado previamente, foi configurado para utilizar o algoritmo ADAM, no esquema de 2 camadas ocultas; learning rate - parâmetro utilizado para controlar o índice de atualização dos pesos - padrão de 0,001; e random state - parâmetro utilizado para inicialização dos pesos - como 1. Já o algoritmo de árvore de decisão, não possui valores a serem configurados, por se tratar de um algoritmo simples que roda seguindo o modelo CART. Ambos trouxeram no momento resultados que serão discutidos mais convenientemente no próximo capítulo.

Na próxima etapa, foi aplicado o algoritmo de análise paraconsistente, que analisa todas as possíveis combinações das vinte e duas *features* e retorna a combinação de n features que apresentar um melhor *score*. Após o resultado, todo o processo descrito anteriormente é refeito para apenas as features selecionadas pelo método paraconsistente, o que gerou novos resultados para serem comparados.

4 RESULTADOS

Neste capítulo, é disposto em forma de tabela a comparação de todos os resultados dos algoritmos, respectivamente, anteriores e posteriores à análise paraconsistente. Cada método foi executado cem vezes, sendo calculado a média aritmética para cada *feature* com duas casas decimais de confiança. As células com os melhores resultados foram realçadas em verde para facilitar a comparação do desempenho anterior e posterior ao método.

4.1 ENGENHARIA PARACONSISTENTE

O algoritmo paraconsistente aplicado nas vinte e duas características, selecionou um grupo quatorze características mais relevantes para a reaplicação dos algoritmos classificadores, sendo elas:

- Z 3 -1 (CB)
- RMS Z3,4,5,6 CB
- RMS Z3,4,5,6 CF
- Z 3 -1 (CF)
- IHD
- Z31Cornea-A
- IVA
- Ele B BFTE 8mm Thinnest
- RMSHOACB
- RMSHOACF
- ARTMax
- AGE NORMAIS
- D2.4mm Pachy/PF

- D2.8mm Pachy/PF

4.1.1 Árvore de decisão anterior x posterior à engenharia paraconsistente

Tabela 1 – Métricas por grupo anterior à engenharia paraconsistente

Grupo	Sensibilidade	Precisão
A	71.75%	72.75%
KC	91.75%	96.88%
VAE-NT	68.75%	67.63%

Tabela 2 – Métricas por grupo posterior à engenharia paraconsistente

Grupo	Sensibilidade	Precisão
A	70.13%	70.38%
KC	94.13%	95.15%
VAE-NT	68.25%	67.38%

Tabela 3 – Acurácia antes e depois da engenharia paraconsistente respectivamente

Acurácia	77.89%	X	Acurácia.	76.78%
----------	--------	---	-----------	--------

O algoritmo de árvore de decisão apresentou um melhor resultado em todas as *features*, exceto para a sensibilidade do grupo KC, anteriormente à aplicação da engenharia paraconsistente. A sensibilidade do grupo VAE-NT, atributo de maior valor de significância deste estudo, teve uma pequena diminuição, ainda estando bem próximos nos dois cenários analisados. Os outros atributos para os outros apresentaram uma mudança levemente mais expressiva que a sensibilidade no grupo subclínico, além de apresentar uma pequena diminuição na acurácia na engenharia paraconsistente.

4.1.2 MLP anterior x posterior à engenharia paraconsistente

Tabela 4 – Métricas por grupo anterior à engenharia paraconsistente

Grupo	Sensibilidade	Precisão
A	85.12%	79.63%
KC	93.25%	98.13%
VAE-NT	74.75%	76.63%

Tabela 5 – Métricas por grupo posterior à engenharia paraconsistente

Grupo	Sensibilidade	Precisão
A	87.5%	78.38%
KC	92.13%	97.75%
VAE-NT	74.13%	78.38%

Tabela 6 – **Acurácia antes e depois da engenharia paraconsistente respectivamente**

Acurácia	84.44%	X	Acurácia	85.11%
----------	--------	---	----------	--------

O algoritmo MLP baseado no ADAM apresentou, em relação ao algoritmo de árvore de decisão, resultados substancialmente superior em todas as features, com exceção da sensibilidade do grupo KC posterior à análise paraconsistente. Foi possível observar uma melhora no campo da sensibilidade para o grupo A, precisão pro grupo VAE-NT, e acurácia do método após a aplicação da engenharia paraconsistente. A sensibilidade para o grupo VAE-NT apresentou uma leve queda.

4.2 COMPARAÇÃO COM O TRABALHO BASE

A partir dos resultados do trabalho base, que utilizou os algoritmos MVS Linear e KNN, foi feita uma comparação de desempenho dos algoritmos para o grupo subclínico (VAE-NT). O trabalho base abrangeu um terceiro cenário, onde foram selecionadas vinte e três principais características, além dos cenários pré e pós análise paraconsistente, assim como este. Os maiores valores locais de cada cenário foram realçados em azul, enquanto o maior global foi realçado em verde.

Tabela 7 – **Precisão em cada cenário por para cada algoritmo**

Grupo	Todos os componentes	Engenharia paraconsistente	ACP
MVS	97.3%	97.8%	96.8%
KNN	96.5%	89.3%	96.5%
ARV	67.63%	67.38%	X
MLP	76.63%	78.38%	X

Tabela 8 – **Acurácia em cada cenário por para cada algoritmo**

Grupo	Todos os componentes	Engenharia paraconsistente	ACP
MVS	91.4%	80.9%	91.4%
KNN	89.6%	82.3%	89.6%
ARV	77.89%	76.78%	X
MLP	84.44%	85.11%	X

Tabela 9 – Sensibilidade em cada cenário por para cada algoritmo

Grupo	Todos os componentes	Engenharia paraconsistente	ACP
MVS	85.4%	57.0%	85.5%
KNN	82.2%	67.7%	81.7%
ARV.	68.75%	68.25%	X
MLP	74.75%	74.13%	X

É possível observar que o algoritmo MVS linear obteve uma maior precisão, em todos os cenários, assim como também obteve uma melhor acurácia com exceção do cenário que foi aplicada a engenharia paraconsistente, onde o MLP obteve um resultado superior, melhor inclusive que o cenário que utiliza todas as características.

Com relação a sensibilidade, que é o caso mais importante do estudo, ainda obtemos um melhor resultado para o MVS, com exceção do cenário da aplicação da análise paraconsistente, onde o MLP também foi superior. Um ponto importante de se observar é que, o cenário no qual foi aplicado a engenharia paraconsistente no trabalho base, obteve um resultado substancialmente menor que os outros dois cenários, enquanto neste trabalho, apesar deste cenário apresentar resultados inferiores, esta diminuição foi bastante sutil, apresentando até uma melhora em alguns casos.

Tendo a sensibilidade como foco da análise deste projeto, é possível perceber que o MLP teve o melhor aumento deste parâmetro em um cenário diferente do que utiliza todas as características, apesar de o MVS ser superior na maioria dos outros.

5 CONCLUSÃO

Este estudo teve como principal objetivo de, a partir do estudo base, aplicar novos algoritmos baseados em *machine learning*, incluindo redes neurais, reaplicar a análise paraconsistente para uma base de dados mais robusta e observar os resultados dessas alterações.

Partindo do princípio da análise e comparação dos algoritmos antes da aplicação da análise paraconsistente, observamos uma nítida diferença entre os dois inicialmente utilizados no projeto. O MLP obteve uma melhor performance em todos os cenários para todas as características e para todos os grupos. O mesmo se repetiu após a aplicação da análise paraconsistente, que gerou um cenário de queda para a maioria das features, porém, o MLP se manteve superior ao CART antes e depois da aplicação do algoritmo. Outro fator que pode ser aprofundado, é que o algoritmo MLP possui uma diversidade de possibilidades de configuração, podendo ser alterado seu algoritmo base, número de neurônios e camadas ocultas, o que pode afetar o resultado final, antes ou depois de aplicada a análise paraconsistente.

Com relação à análise paraconsistente, o algoritmo não obteve os melhores resultados ao causar uma queda em praticamente todas as features, inclusive na sensibilidade do grupo VAE-NT, cenário com o maior foco deste projeto. Apesar disso, foi possível observar que a análise paraconsistente obteve um resultado melhor do que quando aplicada no trabalho base, que obteve uma queda substancial em todos os cenários. Apesar de O MSV demonstrar ser o algoritmo mais eficiente de forma geral, no cenário da análise paraconsistente, o MLP obteve um resultado superior. Desta forma, é perceptível que o algoritmo pode ser aperfeiçoado ao ponto de entregar resultados um melhor resultado com relação ao cenário aplicado, uma vez que, comparado ao estudo anterior, obteve desempenho mais próximo dos melhores casos, que foram aqueles onde os algoritmos foram aplicados em todas as características ou nas vinte e três mais relevantes pelo método ACP no trabalho base.

Por fim, é perceptível que os algoritmos classificadores podem apresentar auxílio para embasar diagnósticos médicos. Junto à análise paraconsistente, que, ainda não apresentou resultados majoritariamente positivos nesse campo de aplicação, mas, que possibilita o aperfeiçoamento do algoritmo para gerar melhores resultados em estudos futuros.

Para trabalhos futuros, é sugerido a aplicação de novos métodos classificadores, assim como novos testes de configurações do algoritmo MLP, que apresentou resultados promissores, assim como o aperfeiçoamento do algoritmo de análise paraconsistente, que está próximo de atingir resultados melhores que os cenários onde todas as características são utilizadas.

REFERÊNCIAS

- ARBELAEZ, Maria Clara et al. Use of a support vector machine for keratoconus and subclinical keratoconus detection by topographic and tomographic data. *Ophthalmology*, Elsevier, v. 119, n. 11, p. 2231–2238, 2012.
- BENTO, Bruna Maria Pechini; KUX, Hermann Johann Heinrich; KÖRTING, Thales Sehn. Assessment of classifiers through decision tree and regression tree algorithms in urban area using worldview-2 image.
- BUKOWSKI, Charles. *Factotum*. [S.l.]: Black Sparrow Press, 1975.
- ELIAS, Rosana Molina Saraiva et al. Ceratocone: fatores prognósticos. *Arquivos Brasileiros de Oftalmologia*, SciELO Brasil, v. 68, p. 491–494, 2005.
- GUIDO, Rodrigo Capobianco. Paraconsistent feature engineering [lecture notes]. *IEEE Signal Processing Magazine*, IEEE, v. 36, n. 1, p. 154–158, 2018.
- HASHEMI, Hassan et al. Pentacam top indices for diagnosing subclinical and definite keratoconus. *Journal of current ophthalmology*, Elsevier, v. 28, n. 1, p. 21–26, 2016.
- IMBORNONI, Lauren Marie; MCGHEE, Charles Ninian John; BELIN, Michael W. Die evolution in der behandlung des keratokonus: von der diagnose zur therapie. *Klinische Monatsblätter für Augenheilkunde*, Ferdinand Enke Verlag, 2018.
- KINGMA, Diederik P; BA, Jimmy. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- MIA, Mohammad Mahmudul Alam et al. An algorithm for training multilayer perceptron (mlp) for image reconstruction using neural network without overfitting. *International Journal of Scientific & Technology Research*, v. 4, n. 02, p. 271–275, 2015.
- NETO, Jogi Suda. Die evolution in der behandlung des keratokonus: von der diagnose zur therapie. Jogi Suda Neto, 2018.
- POLLHAMMER, Matthias; CURSIEFEN, Claus. Bacterial keratitis early after corneal crosslinking with riboflavin and ultraviolet-a. *Journal of Cataract & Refractive Surgery*, Elsevier, v. 35, n. 3, p. 588–589, 2009.
- RUDER, Sebastian. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.
- SAFAVIAN, S Rasoul; LANDGREBE, David. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, IEEE, v. 21, n. 3, p. 660–674, 1991.
- SILVA, LM. Uma aplicação de árvores de decisão, redes neurais e knn para a identificação de modelos arma não sazonais e sazonais. *Rio de Janeiro. 145p. Tese de Doutorado-Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro*, 2005.
- SOUZA, Murilo Barreto et al. Evaluation of machine learning classifiers in keratoconus detection from orbscan ii examinations. *Clinics*, SciELO Brasil, v. 65, p. 1223–1228, 2010.

SWAIN, Philip H; HAUSKA, Hans. The decision tree classifier: Design and potential. *IEEE Transactions on Geoscience Electronics*, IEEE, v. 15, n. 3, p. 142–147, 1977.

TAI, Lai-Yong et al. Central corneal thickness measurements with different imaging devices and ultrasound pachymetry. *Cornea*, LWW, v. 32, n. 6, p. 766–771, 2013.

TUFT, Stephen J et al. Prognostic factors for the progression of keratoconus. *Ophthalmology*, Elsevier, v. 101, n. 3, p. 439–447, 1994.