

UNIVERSIDADE FEDERAL DE ALAGOAS  
INSTITUTO DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

MARIANA SILVA GOIS DE ALMEIDA

**O uso da aprendizagem de máquina para a criação de fórmula biométrica para  
o cálculo da lente intraocular**

Maceió

2023

MARIANA SILVA GOIS DE ALMEIDA

**O uso da aprendizagem de máquina para a criação de fórmula biométrica para o cálculo da lente intraocular**

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Informática do Instituto de Computação da Universidade Federal de Alagoas para exame de obtenção de título de Mestre.

Área de concentração: Ciência da computação

Orientadores: Aydano Pamponet Machado e João Marcelo de Almeida Gusmão Lyra

Maceió

2023

**Catálogo na Fonte  
Universidade Federal de Alagoas  
Biblioteca Central  
Divisão de Tratamento Técnico**

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 - 1767

A447u Almeida, Mariana Silva Gois de.  
O uso da aprendizagem de máquina para a criação de fórmula  
biométrica para o cálculo da lente intraocular / Mariana Silva Gois de  
Almeida. – 2023.  
58 f. : il.

Orientador: Aydano Pamponet Machado.

Orientador: João Marcelo de Almeida Gusmão Lyra.

Dissertação (mestrado em Informática) - Universidade Federal de  
Alagoas. Instituto de Computação. Maceió, 2023.

Bibliografia: f. 49-53.

Anexos: f. 54-58.

1. Biometria ocular. 2. Fórmulas biométricas. 3. Cálculo do poder da  
Lente intraocular. 4. Aprendizagem de máquina. 5. Inteligência artificial. I.  
Título.

CDU: 004.383.8

Dedico este trabalho primeiramente a Deus e à Nossa Senhora, por me darem forças e conhecimento para seguir em frente e alcançar meus objetivos.

Aos meus pais e ao meu irmão, pelo amor incondicional e por contribuírem com todas as etapas da minha formação acadêmica e pessoal.

## **AGRADECIMENTOS**

Aos meus orientadores, Aydano Machado e João Marcelo Lyra, que mesmo com todas as demandas externas me deram todo suporte necessário para realização deste trabalho.

Aos meus professores, em especial Edileuza Virgínio Leão, que sempre acreditaram no meu potencial.

Aos meus colegas de mestrado e profissão que contribuíram de alguma forma na execução do mestrado.

*A vida não é mais do que um único e longo encadeamento de coincidências. (Jostein Gaarder)*

**Programa de Pós-Graduação em Informática do Instituto de Computação, da  
Universidade Federal de Alagoas**

**Banca examinadora da dissertação de Mestrado**

Aluna: Mariana Silva Gois de Almeida

Professores e orientadores: Aydano Pamponet Machado e João Marcelo de Almeida  
Gusmão Lyra

**Membros:**

1. Professor Evandro de Barros Costa
2. Professora Edileuza Virginio Leão
3. Rafael Ferreira Leite de Mello

## RESUMO

**Objetivo:** Elaborar uma nova fórmula biométrica com base em um banco de dados da população brasileira e comparar o seu desempenho com o de outras 6 fórmulas existentes tanto na prática clínica quanto no meio científico. **Local:** Os participantes foram avaliados e submetidos à cirurgia em um hospital oftalmológico de Brasília, DF, Brasil. **Desenho:** Um estudo clínico retrospectivo para comparar o desempenho de fórmulas biométricas. **Metodologia:** A fórmula biométrica foi construída seguindo o processo de descoberta de conhecimento em base de dados (KDD) e os desempenhos de modelos de aprendizagem de máquina e das fórmulas de Barrett Universal II, Haigis, Hoffer Q, Holladay 1, Kane e SRK/T foram analisados comparando-se o erro absoluto mediano e a porcentagem de olhos com menos de 0,5 e 1,0D de erro absoluto. **Resultados:** Foram analisados 1526 olhos únicos, onde as fórmulas de Barrett Universal II, Haigis, Hoffer Q, Holladay 1, Kane e SRK/T e os modelos MLP e SVR obtiveram, respectivamente, os seguintes valores de erro absoluto mediano: 0.393, 0.426, 0.427, 0.387, 0.409, 0.410, 0.405 e 0.370 D. Os desempenhos do modelo SVR e das fórmulas Holladay 1 e Barrett Universal II também foram superiores em todos os subgrupos estudados. **Conclusão:** A fórmula baseada em modelo computacional proposta obteve resultados superiores em comparação com outras fórmulas para cálculo do poder da lente intraocular nesta população e em todos os subgrupos biométricos estudados.

**Palavras-chave:** Biometria ocular; Fórmulas biométricas; Cálculo do poder da LIO; Aprendizagem de Máquina; Inteligência Artificial.

## ABSTRACT

**Objective:** To develop a new biometric formula based on a database of the Brazilian population and compare its performance with that of 6 other formulas existing both in clinical practice and in the scientific world. **Location:** Participants were evaluated and underwent surgery at an ophthalmological hospital in Brasília, DF, Brazil. **Design:** A retrospective clinical study to compare the performance of biometric formulas. **Methodology:** The biometric formula was constructed following the knowledge discovery in database (KDD) and the performances of machine learning models and the formulas of Barrett Universal II, Haigis, Hoffer Q, Holladay 1, Kane and SRK/ T were analyzed by comparing the median absolute error and the percentage of eyes with less than 0.5 and 1.0D of absolute error. **Results:** 1526 single eyes were analyzed, where the Barrett Universal II, Haigis, Hoffer Q, Holladay 1, Kane and SRK/T formulas and the MLP and SVR models obtained, respectively, the following median absolute error values: 0.393, 0.426, 0.427, 0.387, 0.409, 0.410, 0.405 and 0.370 D. The performances of the SVR model and the Holladay 1 and Barrett Universal II formulas were also superior in all subgroups studied. **Conclusion:** The proposed computational model-based formula obtained superior results compared to other formulas for calculating intraocular lens power in this population and in all biometric subgroups studied.

**Keywords:** Ocular biometrics; Biometric formulas; Calculation of IOL power; Machine Learning; Artificial intelligence.

## LISTA DE ILUSTRAÇÕES

|  |    |
|--|----|
| Figura 1. Modelos preditivos: Porcentagem de olhos por faixa de MAE _____  | 37 |
| Figura 2. Modelos preditivos: MAE versus ACD _____   | 38 |
| Figura 3. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Comprimento Axial _____  | 40 |
| Figura 4. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Ceratometria central _____   | 41 |
| Figura 5. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Profundidade de câmara anterior _____  | 41 |
| Figura 6. Modelos preditivos: ceratometria central versus erro absoluto médio _____  | 54 |
| Figura 7. Modelos preditivos: comprimento axial versus erro absoluto médio _____   | 54 |
| Figura 8. Fórmulas e modelos preditivos: comprimento axial versus erro absoluto médio em olhos curtos _____                                      | 55 |
| Figura 9. Fórmulas e modelos preditivos: comprimento axial versus erro absoluto médio em olhos longos _____                                      | 55 |
| Figura 10. Fórmulas e modelos preditivos: profundidade da câmara anterior versus erro absoluto médio em olhos com câmara anterior profunda _____ | 56 |
| Figura 11. Fórmulas e modelos preditivos: profundidade da câmara anterior versus erro absoluto médio em olhos com câmara anterior rasa _____     | 56 |
| Figura 12. Fórmulas e modelos preditivos: ceratometria central versus erro absoluto médio em olhos com córnea plana _____                        | 57 |
| Figura 13. Fórmulas e modelos preditivos: ceratometria central versus erro absoluto médio em olhos com córnea curva _____                        | 57 |

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1. Valores demográficos  | 35 |
| Tabela 2. Modelos preditivos: Erro previsto e cálculo das LIOs em todos os intervalos de AL         | 36 |
| Tabela 3. Fórmulas e modelos preditivos: Erro previsto e cálculo das LIOs em todos intervalos de AL | 39 |
| Tabela 4. Comparação dos resultados   | 46 |
| Tabela 5. Valores das variáveis utilizadas em cada algoritmo da biblioteca scikit-learn             | 53 |

## LISTA DE ABREVIATURAS E SIGLAS

|           |   |
|-----------|---|
| LIO       | Lente intreocular   |
| SRK       | Sanders-Retzlaff-Kraff  |
| ACD       | Anterior Chamber Depth  |
| SRK/T     | Sanders-Retzlaff-Kraff Teórica  |
| Hill-RBF  | Hill Radial Basis Function  |
| PEARL-DGS | Postoperative spherical Equivalent prediction using Artificial intelligence and Linear algorithms – Debellemanière, Gatinel, Saad |
| D         | Dioptria  |
| AL        | Axial length / Comprimento Axial  |
| K         | Ceratometria  |
| ELP       | Effective Lens Position / Posição Efetiva da Lente  |
| WTW       | White-To-White / Branco-A-Branco  |
| IA        | Inteligência Artificial   |
| MLP       | Perceptron Multicamadas   |
| CNN       | Redes Neurais Convolucionais  |
| RNN       | Redes Neurais Recorrentes   |
| GBR       | Gradient Boosting   |
| SVM       | Máquina de Vetores de Suporte   |
| k-NN      | k Nearest Neighbors   |
| CAAE      | Certificado de Apresentação de Apreciação Ética   |
| ULIB      | User group for Laser Interference Biometry  |
| MEA       | Mean Absolute Error   |

|       |                                 |
|-------|---------------------------------|
| MedAE | Median Absolute Error           |
| IC    | Intervalo de Confiança          |
| KDD   | Knowledge Discovery in Database |
| SVR   | Support Vector Regression       |
| K1    | Ceratometria mais plana         |
| K2    | Ceratometria mais curva         |
| mm    | Milímetro                       |
| LR    | Linear Regression               |

## SUMÁRIO

|       |   |    |
|-------|---|----|
| 1     | INTRODUÇÃO.....                           | 15 |
| 2     | REFERENCIAL TEÓRICO .....                 | 18 |
| 2.1   | FÓRMULAS BIOMÉTRICAS.....                 | 18 |
| 2.1.1 | PRIMEIRA GERAÇÃO .....                    | 18 |
| 2.1.2 | SEGUNDA GERAÇÃO .....                     | 19 |
| 2.1.3 | TERCEIRA GERAÇÃO .....                    | 19 |
| 2.1.4 | QUARTA GERAÇÃO .....                      | 20 |
| 2.1.5 | QUINTA GERAÇÃO.....                       | 21 |
| 2.2   | APRENDIZAGEM DE MÁQUINA.....              | 23 |
| 2.2.1 | ÁRVORE DE DECISÃO .....                   | 24 |
| 2.2.2 | RANDOM FOREST.....                        | 25 |
| 2.2.3 | REGRESSÃO LINEAR .....                    | 26 |
| 2.2.4 | REDE NEURAL.....                          | 27 |
| 2.2.5 | REDES NEURAIS PROFUNDAS .....             | 28 |
| 2.2.6 | MÁQUINA DE VETORES DE SUPORTE (SVM) ..... | 28 |
| 2.2.7 | k-NEAREST NEIGHBORS (kNN) .....           | 29 |
| 2.3   | PERGUNTA DE PESQUISA E OBJETIVOS.....     | 30 |
| 3     | METODOLOGIA.....                          | 31 |
| 3.1   | MATERIAIS.....                            | 31 |
| 3.2   | MÉTODO .....                              | 32 |
| 4     | RESULTADOS .....                          | 35 |
| 5     | DISCUSSÃO.....                            | 43 |
| 6     | CONCLUSÃO.....                            | 47 |
|       | REFERÊNCIAS.....                          | 48 |
|       | APÊNDICE .....                            | 53 |

## 1 INTRODUÇÃO

A catarata é a principal causa de cegueira reversível no mundo, representando 32,4 milhões de cegos e 191 milhões com baixa visão em 2010.(LAM et al., 2015) A catarata é mais prevalente na população idosa, já que o cristalino passa por um processo de opacificação com o tempo. O tratamento da catarata consiste na substituição do cristalino opacificado por uma lente intraocular (LIO) de material biocompatível com poder de refração adequado ao olho de cada indivíduo.

A cirurgia de catarata passou por uma revolução em sua técnica na década de 1960, quando foi desenvolvida a técnica de facoemulsificação, na qual a uma caneta de ultrassom passou a ser utilizada para emulsificar e aspirar o cristalino cataratoso por meio de uma incisão cirúrgica de 2 a 3 milímetros de diâmetro. Com a realização de uma cirurgia minimamente invasiva surgiu a necessidade de desenvolver lentes intraoculares maleáveis que pudessem ser implantadas dobradas por meio da incisão de no máximo 3 mm. (GURNANI B, KAUR K., 2023)

A partir da evolução na técnica cirúrgica e nas LIOs, surgiram novas cobranças com relação ao resultado refrativo dos pacientes, com o objetivo de obter cada vez mais independência dos óculos.

Com o desenvolvimento de fórmulas biométricas, equações matemáticas utilizadas para calcular a potência refrativa da LIO, e aparelhos capazes de captar com mais precisão as medidas oculares, o aspecto refrativo da cirurgia de catarata tornou-se cada vez mais importante, e hoje a maioria dos pacientes espera alcançar um alto grau de independência de óculos após a cirurgia.

Várias fórmulas biométricas foram desenvolvidas nas últimas décadas. A primeira fórmula foi proposta na década de 1960 com embasamento teórico e resultados e características muito semelhantes às demais fórmulas de primeira geração publicadas na sequência. Posteriormente, fórmulas de segunda geração foram desenvolvidas com uma abordagem mais empírica, como as técnicas de regressão linear aplicadas

em fórmulas SRK.(KRAFF; SANDERS; LIEBERMAN, 1978; SANDERS; KRAFF, 1980)

A terceira geração de fórmulas biométricas misturou fundamentos teóricos e empíricos e considerou a influência de outros fatores, como a profundidade da câmara anterior (ACD) e variabilidade intercirurgião, em seus cálculos. Algumas fórmulas utilizadas até hoje pertencem a esse grupo, como a Hoffer Q (HOFFER, 1993), Holladay 1 (HOLLADAY et al., 1988) e SRK/T(RETZLAFF; SANDERS; KRAFF, 1990).

As fórmulas de quarta geração foram desenvolvidas principalmente na mesma época da popularização dos biômetros ópticos, que disponibilizaram novas medidas para fórmula construção. Alguns exemplos dessa geração são as fórmulas Haigis (HAIGIS et al., 2000), Holladay 2, Barrett Universal II <sup>1</sup> e Olsen, que são capazes de fornecer excelentes resultados clínicos. No entanto, muitas dessas fórmulas ainda não foram publicadas como artigos revisados por pares na revista científica literatura. Como resultado, não temos acesso detalhado a como eles foram desenvolvidos, limitando a análise da fundamentação teórica e do processo de desenvolvimento. (CONNELL; KANE, 2019; COOKE; COOKE, 2016a; DARCY et al., 2020; KANE et al., 2016; MELLES; HOLLADAY; CHANG, 2018)

Mais recentemente, desenvolveu-se o conceito de uma quinta geração de fórmulas, envolvendo aquelas fórmulas que empregam inteligência artificial ou métodos de otimização por características populacionais, como as fórmulas Hill-RBF<sup>2</sup>, Kane<sup>3</sup> e Pearl-DGS<sup>4</sup>. Essas fórmulas têm sido de grande interesse para a comunidade científica e são frequentemente atualizados com grande quantidade de dados, com bons resultados clínicos comparáveis aos das fórmulas de quarta geração. (CONNELL; KANE, 2019; COOKE; COOKE, 2016a; DARCY et al., 2020; KANE et al., 2016; MELLES; HOLLADAY; CHANG, 2018)

Apesar dos resultados obtidos com as fórmulas mais recentes, o cálculo do poder da LIO em pacientes com olhos com medidas biométricas que divergem dos parâmetros populacionais médios continua sendo um desafio para o cirurgião de catarata. Ajustes

---

<sup>1</sup> Disponível em : [https://calc.apacrs.org/barrett\\_universal2105](https://calc.apacrs.org/barrett_universal2105)

<sup>2</sup> Disponível em: <https://rbfcalculator.com/index.html>

<sup>3</sup> Disponível em: <https://www.iolformula.com/>

<sup>4</sup> Disponível em: <https://www.gatinel.com/recherche-formation/biometrie-oculaire-calculdimplant/pearl-dgs-formula-for-iol-power-calculation/>

para esses casos particulares têm sido propostos (WANG; HOLLADAY; KOCH, 2018), e vários estudos com foco em subgrupos populacionais biométricos específicos têm sido realizados (FARAMARZI; AGHAJANI; GHIASIAN, 2017; GÖKCE et al., 2017; IDROBO-ROBALINO; SANTAELLA; GUTIÉRREZ, 2019; JI et al., 2021; MUSTAFA et al., 2019; SHRIVASTAVA et al., 2018). Apesar da considerável variedade de opções disponíveis para os cirurgiões de catarata, ainda há falta de consenso na literatura sobre a fórmula ideal para cada subgrupo. Como resultado, a tomada de decisão nessa área é complexa e muitas vezes envolve um grau considerável de subjetividade.

Estudos que utilizaram populações latino-americanas são escassos e, em sua grande maioria, utilizaram pequenas bases com poucos indivíduos, o que diminui o poder estatístico e a capacidade de inferir informações de maneira generalizada. Considerando ainda as dificuldades técnicas e socioculturais da região, novos desafios emergem nesta região (DE ALMEIDA FERREIRA et al., 2017), que resultam em limitações no cuidado da população. Em alguns países ainda se estudam técnicas cirúrgicas consideradas defasadas, como a extração extracapsular sem facoemulsificação (RUÍZ et al., 2017), enquanto em países mais desenvolvidos tecnologias como biometria por swept source e laser femtosegundo são de ampla disponibilidade. Tais disparidades são traduzidas em resultados: enquanto os melhores trabalhos publicados na literatura internacional demonstram no máximo cerca de 80% dos pacientes em um equivalente esférico a menos de 0,5D de distância do alvo refrativo (HAHN et al., 2011), as referências latino-americanas apresentam resultados muito mais humildes, atingindo números em torno de 50% dos olhos com erros menores que 0,5D. (GRANADO et al., 2016; MEDINA, 2019; NODALES et al., 2017; RODRIGUES et al., 2015)

Além de estudar e desenvolver uma fórmula biométrica, este trabalho buscou comparar o desempenho de algumas das fórmulas mais utilizadas disponíveis na literatura em uma população de brasileiros. Até onde sabemos, este é o estudo com maior número de olhos operados pelo mesmo cirurgião em uma população com essas características.

## 2 REFERENCIAL TEÓRICO

### 2.1 FÓRMULAS BIOMÉTRICAS

As fórmulas biométricas foram desenvolvidas com o objetivo de calcular o poder dióptrico ideal da lente intraocular (LIO) que deve ser implantada na cirurgia de catarata. Inicialmente baseadas em modelos teóricos do olho e, atualmente, utilizam medidas precisas do olho a partir de aparelhos baseados em interferometria óptica que fazem a captação das medidas oculares, tais como: comprimento axial, curvatura da córnea e profundidade da câmara anterior. A elaboração das fórmulas biométricas para cálculo do poder da LIO pode ser distribuído em cinco gerações.

#### 2.1.1 PRIMEIRA GERAÇÃO

As primeiras fórmulas biométricas foram propostas de forma independente por Fyodorov, Colenbrander, Van der Heijde e Binkhorst, entre as décadas de 1960 e 1970, quando percebeu-se o benefícios de ter o poder refrativo da lente intraocular calculado para cada indivíduo.

Realizando cálculos para obter as medidas biométricas de acordo com modelos teóricos do olho humano (especialmente o modelo esquemático de Gullstrand, desenvolvido em 1905) e ajustando a dioptria da LIO a depender da ametropia do paciente antes de realizar a cirurgia. Apesar de terem sido desenvolvidas por autores diferentes e independentemente, suas fórmulas são algebricamente equivalentes e podem ser resumidas na seguinte fórmula: (COLENBRANDER, 1973; RETZLAFF, 1980)

$$P = \frac{N}{L-C} - \frac{NK}{N-KC} \quad (1)$$

Onde, P: poder da LIO, N: índice de refração do humor aquoso e vítreo, L: comprimento axial, K: curvatura da córnea, C: profundidade da câmara anterior pós operatório, R: raio de curvatura da córnea.

Em 1975, Hoffer descreveu em um editorial a inacurácia das fórmulas existentes e a necessidade de melhorar esse cálculo com o uso de computação para um rápido processamento das variáveis que viessem a ser obtidas para esse fim. (HOFFER, 1975)

### 2.1.2 SEGUNDA GERAÇÃO

Com a incorporação do comprimento axial (AL), curvatura anterior da córnea (K) e técnicas de regressão linear, surgiram as fórmulas de segunda geração: Hoffer, SRK e Olsen (primeira versão). (MONTEIRO; TRINDADE, 2022)

A exemplo da segunda geração, pode-se observar a fórmula Hoffer que foi derivada por meio de regressão linear, gerando uma equação de 1º grau:

$$ELP = (0.292 \times AL) - 2.93 \quad (2)$$

Onde é utilizado o comprimento axial (AL) em centímetros para encontrar a posição efetiva da lente (ELP). (HOFFER, 1993)

A primeira versão da fórmula SRK foi proposta em em 1980 também fazendo uso de regressão linear para encontrar uma relação ente duas variáveis e definir o melhor poder dióptrico da lente intraocular para cada paciente.(RETZLAFF, 1980)

$$P = 111.6 - (2.41 \times L) - (0.87 \times K) \quad (3)$$

Onde P corresponde ao poder da LIO, L ao comprimento axial e K à ceratometria da córnea.

Essa versão foi otimizada por meio de análise de regressão múltipla e desenvolvidos três constantes distintas (A, B e C) das quais apenas a primeira varia de acordo com o modelo da lente. (SANDERS; KRAFF, 1980)

$$P = A - 2.5 \times L - 0,9 \times K \quad (4)$$

Onde P corresponde ao poder da LIO, A à constnte da lente, L ao comprimento axial e K à ceratometria da córnea.

Em geral, tanto as fórmulas de primeira, quanto de segunda geração não atingiam bons resultados refrativos e por isso caíram em desuso.

### 2.1.3 TERCEIRA GERAÇÃO

Já as fórmulas de terceira geração alçaram melhores resultados ao integrar componentes teóricos e empíricos ao cálculo biométrico, entre elas estão Hoffer Q, Holladay 1, SRK II, SRK/T e Barrett Universal. Algumas destas são a principal escolha,

atualmente, para definição da LIO a ser implantada. (LEE; QAZI; PEPOSE, 2008; MONTEIRO; TRINDADE, 2022)

A fórmula SRK passou por mais uma modificação para ter melhor desempenho em olhos com comprimento axial extremo, surgindo, assim, a SRK II que manteve sua simplicidade no cálculo e se destacou nos resultados refrativos quando foi lançada. (SANDERS; RETZLAFF; KRAFF, 1988) Em seguida, foi utilizada otimização baseada na regreeção empírica da fórmula por meio de termos não lineares das fórmulas teóricas, surgiu a SRK/T. (RETZLAFF; SANDERS; KRAFF, 1990)

A Holladay I propôs que, para atingir o resultado refrativo ideal, a fórmula biométrica deveria se basear em três pontos: a avaliação do comprimento axial e da ceratometria para identificar valores improváveis, superar a acurácia das fórmulas existentes para olhos curtos, médios e longos, e o ajuste de uma constante nomeada de “fator cirurgião”. Teve como base 1,200 olhos, submetidos à cirurgia de catarata por 12 cirurgiões, cujos dados biométricos e refrativos foram analisados e a boa performance alcançada se deu pelo cálculo da predição da posição efetiva da lente (ELP). (HOLLADAY et al., 1988)

Em 1993, foi publicada a fórmula Hoffer Q, atualizada com medida personalizada do ACD, para seu desenvolvimento foram utilizados dados de pacientes submetidos a cirurgia por um único cirurgião e com o implante de LIO do mesmo fabricante, totalizando 500 olhos. Sua acurácia foi equivalente à Holladay I e superior à SRK/T. (HOFFER, 1993)

Por fim, a fórmula Barrett também foi otimizada, dando lugar à Barrett II, baseada na relação da câmara anterior com o comprimento axial e a ceratometria, na relação entre a “constante A” e o fator da lente para determinar a câmara anterior e na posição da LIO para o resultado refrativo final. Neste caso foram utilizados dados de 100 pacientes, sem seleção prévia, para comparar com as demais fórmulas de 3ª geração, destacando-se com melhor resultado de erro refrativo previsto em todos os grupos de comprimento axial. (BARRETT, 1993)

#### **2.1.4 QUARTA GERAÇÃO**

As fórmulas de quarta geração surgiram com o advento dos aparelhos de biometria óptica que permitiram avaliação de mais medidas biométricas. Este grupo engloba a

maiora das fórmulas em uso hoje em dia – Haigis, Holladay 2, Barrett Universal II e Olsen.

A Haigis trouxe como inovação o uso de três constantes específicas ( $a_0$ ,  $a_1$  e  $a_2$ ) e a medida da profundidade de câmara anterior (ACD), permitindo determinar de forma mais acurada a posição e o poder da LIO indicada. Entretanto, as constantes utilizadas devem ser derivadas da análise do banco de dados específico do cirurgião, com pelo menos 50 casos com olhos de diversas medidas biométricas. (HAIGIS et al., 2000)

A Holladay 2, disponibilizada em 1998, possui maior facilidade de otimização, porém exige avaliação de sete variáveis – AL, K médio, espessura do cristalino, WTW, ACD, refração pré-operatória e idade do paciente. (LEE; QAZI; PEPOSE, 2008) Os detalhes de como foi feito seu cálculo não foram publicados, a fórmula está disponível online<sup>5</sup>.

Da mesma forma, os detalhes sobre o cálculo da Barrett Universal II<sup>6</sup> não estão disponíveis na literatura, apenas o acesso para realizar o cálculo da LIO. As variáveis utilizadas para o cálculo são: comprimento axial, ceratometria, profundidade da câmara anterior, refração pré-operatória, espessura do cristalino, branco-a-branco e constante A.

A fórmula Olsen publicada em 1995 visando prever a posição efetiva da LIO e consequentemente melhor resultado refrativo, considerou o valor de ACD de acordo com um cálculo de regressão linear utilizando ACD pré-operatório, ceratometria, AL e espessura do cristalino. (OLSEN; CORYDON; GIMBEL, 1995) Após 11 anos, Olsen publicou um novo artigo com a análise de uma base de dados de 900 olhos, que incluía variáveis tais quais refração pós-operatória, ACD pós-operatório efetivo, AL, ACD pré-operatório, espessura do cristalino, ceratometria idade, branco-a-branco e refração pré-operatória. A partir dessas variáveis foi realizado o estudo por regressão multivariada para avaliar covariância, a fim de prever o ACD pós-operatório. (OLSEN, 2006)

### **2.1.5 QUINTA GERAÇÃO**

A existência do grupo de fórmulas da quinta geração ainda não é completamente aceito pela comunidade científica. Também conhecida como nova geração, essas

---

<sup>5</sup> Disponível em: <http://www.hicsoap.com/>

<sup>6</sup> Disponível em: [https://calc.apacrs.org/barrett\\_universal2105/](https://calc.apacrs.org/barrett_universal2105/)

fórmulas utilizam estratégias matemáticas e modelos computacionais para seu desenvolvimento, visando melhor acurácia no cálculo do poder da LIO, entre elas estão: Hill-RBF, Kane, LADAS Superformula e PEARL-DGS. (LI et al., 2022)

A fórmula Hill-RBF (Radial Basis Function) foi publicada em 2016 e foi pioneira por se basear somente em inteligência artificial para obter o poder da LIO. Usa as variáveis AL, ACD e ceratometria, seu algoritmo não foi publicado, mas sua calculadora está disponível online<sup>7</sup> e instalado em alguns modelos de biômetros ópticos.(SAVINI; TARONI; HOFFER, 2020)

Outra fórmula que também não foi divulgada, Kane<sup>8</sup>, é baseada em IA e em óptica teórica, segundo seu autor. Usa como variáveis: AL, K, ACD e gênero. Sua calculadora está disponível online e vem crescendo o número de estudos publicados demonstrando o bom desempenho dessa fórmula. (SAVINI; TARONI; HOFFER, 2020)

A LADAS Superformula<sup>9</sup> segue um padrão diferente de todas as fórmulas apresentadas anteriormente, consiste na indicação da melhor fórmula – dentre a Hoffer Q, Holladay 1, Holladay 2 e SRK/T – que o cirurgião deve seguir para definir a dioptria da LIO em cada olho, de acordo com um modelo tridimensional baseado em IA. (SAVINI; TARONI; HOFFER, 2020)

A fórmula PEARL-DGS <sup>10</sup>(Postoperative spherical Equivalent prediction using Artificial intelligence and Linear algorithms – Debellemanière, Gatinel, Saad) também é baseada em inteligência artificial e não teve sua estrutura publicada.

---

<sup>7</sup> Disponível em: <http://www.rbfcalsculator.com/>

<sup>8</sup> Disponível em: <https://www.iolformula.com/>

<sup>9</sup> Disponível em: <http://www.iolcalc.com/>

<sup>10</sup> Disponível em: <http://www.iolsolver.com/>

## 2.2 APRENDIZAGEM DE MÁQUINA

A aprendizagem de máquina é considerada um subcampo da inteligência artificial (IA), na qual é realizado o desenvolvimento de algoritmos que permitam o sistema aprender a partir da experiência e desenvolver soluções com melhores desempenhos, sendo assim, procura emular a inteligência humana. (EL NAQA; MURPHY, 2015)

Sua aplicação pode ser em mineração de grandes bases de dados com valiosas regularidades implícitas que podem ser descobertas automaticamente, domínios pouco compreendidos nos quais os seres humanos não possuem conhecimento suficiente pra desenvolver algoritmos eficazes e em domínios que o programa deve se adaptar automaticamente às condições em mudança. (MITCHELL, 1997)

Pode-se observar o sucesso do aprendizado das máquinas em diversos campos do conhecimento, reconhecimento de padrões, visão computacional, engenharia espacial, biologia computacional para biomedicina e aplicações de medicações, o processo de descoberta de conhecimento consiste basicamente em sete etapas: (EL NAQA; MURPHY, 2015; REICH, 1996)

- Seleção de dados: reunir conjunto de dados com informações relevantes para a resolução da tarefa proposta;
- Pré-processamento dos dados: remoção de informações incompletas, em formatos diferentes ou com alguma outra irregularidade;
- Escolha do algoritmo: definir se o algoritmo de aprendizagem de máquina a ser utilizado baseado em aprendizagem supervisionada, não supervisionada ou por reforço, a depender da natureza dos dados e do problema em questão;
- Treinamento do modelo: realização de testes com conjuntos de dados de treinamento;
- Avaliação do modelo: realização de teste do modelo com conjunto de dados não utilizados previamente para treinamento;
- Ajuste e otimização: realizar modificações no modelo de acordo com os resultados da etapa anterior;
- Implantação do modelo: implementação do modelo para realizar previsões ou auxiliar em tarefas predeterminadas.

Ao colocar juntas as as técnicas apresentadas, é possível criar um sistema que imita como funcionam seres sencientes através da repetição e criação de parâmetros bem definidos (EL NAQA; MURPHY, 2015). Para um melhor entendimento, a aprendizagem de máquina é geralmente dividida em três tipos: a supervisionada, a não supervisionada e a por reforço.

O aprendizado supervisionado é uma forma de ensinar à máquina a relação existente entre certos parâmetros de entrada e de saída pré-determinados através da repetição com a presença de um tutor. Essa abordagem pode ser utilizada para classificação e regressão por meio do treinamento do modelo com conjunto de dados identificados. (EL NAQA; MURPHY, 2015; RUSSEL; NORVIG, 2013)

Já o aprendizado não supervisionado é outra forma de ensinar a máquina, dando-a mais liberdade de decisão dado o objetivo final (EL NAQA; MURPHY, 2015), por exemplo, ao receber um conjunto de dados sem rótulos e o modelo dever encontrar padrões e estruturas por conta própria, sem a presença de um tutor (RUSSEL; NORVIG, 2013). O resultado neste caso vai ser a identificação de padrões ou associações que vão fazer uma descrição e prover um melhor entendimento dos dados por meio desse tipo de conhecimento.

Por fim, no aprendizado por reforço o agente ou modelo está exposto a um ambiente dinâmico com o qual interage e aprende a partir de reforços – recompensas ou punições – e definir quais ações levam aos melhores resultados.(MAHESH, 2019; RUSSEL; NORVIG, 2013)

Existem diversos algoritmos de aprendizagem de máquina, tais como rede neural, *random forest*, regressão polinomial, redes neurais profundas, regressão linear, entre outros. (SOMVANSHI; CHAVAN, 2016)

### **2.2.1 ÁRVORE DE DECISÃO**

O algoritmo de árvore de decisão é um dos modelos de ensino de algoritmo mais bem-sucedidos dada sua simplicidade, compreensibilidade, falta de parâmetros e conseguir manejar tipos de dados misturados, consiste em um modelo de aprendizagem de máquina supervisionado por meio da divisão dos dados em conjuntos e subconjuntos com base em suas características, de forma hierárquica,

onde cada nó representa uma decisão ou característica e cada folha uma classe ou valor predito, à imagem das ramificações de uma árvore. (SU; ZHANG, 2006)

Representam um modelo útil para tomadas de decisões e de fácil compreensão, em casos de instâncias representadas por pares de atributo-valor, de a função final ter valores de *output* discretos, de descrições disjuntivas, de erros na base de dados de treino e base de dados faltando valores de atributos. (MITCHELL, 1997)

Para análise de dados complexos ou com muitas características, a árvore de decisão pode ser melhorada com técnicas que vão combinar várias árvores, tais como *Random Forest* e *Gradient Boosting (GBR)*. (SUTHAHARAN, 2016)

Em casos de dados com ruído ou do número de exemplo para treinamento muito pequeno para produzir uma amostra representativa da função, pode ocorrer *overfitting* ou sobre ajuste do modelo. *Overfitting* leva à redução da acurácia da árvore de decisão ao utilizá-la com outros exemplos. Para evitar esse prejuízo deve-se parar a construção ou aprendizagem da árvore antes, ou seja, antes de alcançar o ponto que classifica perfeitamente os dados de treinamento, ou permitir que a árvore sobreponha os dados e só depois realizar a poda da árvore. (MITCHELL, 1997)

Outro ponto importante é definir o tamanho correto da árvore de decisão, para isso devem ser utilizados grupos separados de exemplos, diferentes do grupo de treinamento; utilizar todos os dados disponíveis para treinamento, mas aplicar teste estatístico para definir se deve expandir ou podar determinado nó que pode ter uma melhora além do grupo de treinamento; e determinar explicitamente a complexidade para os exemplos de treinamento e a árvore de decisão, de forma a interromper o crescimento da árvore quando esse tamanho for minimizado. (MITCHELL, 1997)

### **2.2.2 RANDOM FOREST**

Principalmente utilizado para tarefas de classificação, uma resposta categórica variável, e regressão, respostas contínuas, utiliza a previsão de modelos mais simples para elaborar uma previsão final mais precisa, isso é alcançado devido a associação de várias árvores de decisão para o treinamento do algoritmo. (CUTLER; CUTLER; STEVENS, 2012)

As *Random Forests* são interessantes por prover medição de importância de variável, diferenciação do peso das classes, imputação de valor faltante, visualização e detector de ponto fora da curva. Assim, consegue melhorar o seu desempenho e é muito utilizada em diagnósticos médicos e previsão do mercado financeiro, por exemplo. (BREIMAN, 2001; CUTLER; CUTLER; STEVENS, 2012)

Para otimização do algoritmo pode ser aplicada a randomização na seleção dos atributos, possibilitando que cada árvore seja diferente, são as chamadas “*extremely randomized tree*”. Além disso, todos os hiperparâmetros das *random forests* podem ser treinados por validação cruzada. (RUSSEL; NORVIG, 2013)

Criar uma *random forest* com  $k$  árvores pode parecer demandar mais tempo, entretanto, isso não é ruim, pois cada ponto de divisão passa mais rápido por estar considerando menos atributos; a etapa de poda de cada árvore pode ser ignorado, já que o conjunto de árvores como um todo reduz o *overfitting*; se houver disponíveis  $k$  computadores, as árvores podem ser construídas em paralelo. (RUSSEL; NORVIG, 2013)

### 2.2.3 REGRESSÃO LINEAR

Esta técnica é utilizada para estabelecer uma relação entre uma variável dependente e outras, podendo ou não ser dependentes ou independentes, a fim de encontrar as melhores previsões sobre o comportamento de sua variável dependente. (SU; YAN; TSAI, 2012) Existem tipos diferentes de regressão linear:

- Regressão logística: utilizada para estimar a probabilidade de uma instância pertencer a uma classe. (GÉRON, 2019)
- Regressão de Lasso (*Least Absolute Shrinkage and Selection Operator*): é uma versão regularizada da regressão linear, tende a eliminar os pesos das características menos importantes, aproximando de zero. (GÉRON, 2019)
- Regressão de Ridge: é uma versão regularizada da regressão linear, ou seja, o algoritmo é forçado a ajustar os dados e manter os pesos dos modelos reduzidos. (GÉRON, 2019)

## 2.2.4 REDE NEURAL

Inspirado em parte no funcionamento do cérebro humano, o algoritmo de rede neural é composto por unidades de neurônios artificiais organizados em camadas, transmitindo sinais por meio das conexões. (MITCHELL, 1997)

Inicialmente, é necessário alimentar a rede com um grande conjunto de dados de treinamento e a partir destes aprender padrões e estruturas nos dados. (HAYKIN, 1999) Ou seja, ao abstrair o cérebro humano para um modelo que propõe o processamento de informações, um modelo simples é estabelecido, compondo variadas redes de acordo com a necessidade das conexões. (WU; FENG, 2018)

As redes neurais se encaixam bem em problemas cujo banco de dados apresenta ruído, dados complexos – por exemplo, inputs de câmeras –, instâncias com muitos pares de atributo-valor, a saída da função pode ser um valor discreto, real ou um vetor de vários atributos de valor real ou discreto, os grupos de treino pode conter erros, tempo longo de treinamento é aceito, pode ser necessária uma avaliação rápida da função final e não é importante o homem entender a função final. (MITCHELL, 1997)

O perceptron é o modelo matemático de um neurônio biológico, que obtém um vetor de entrada de valores reais, calcula a combinação linear das entradas e entrega na saída valores que dependem do ponto de corte estabelecido. (MITCHELL, 1997)

Alguns tipos de redes neurais são:

- Perceptron simples: rede neural de camada única. (RUSSEL; NORVIG, 2013)
- Perceptron multicamadas (MLP): possui três tipos de camadas (entrada, saída e oculta). (ABIRAMI; CHITRA, 2020)
- Redes neurais convolucionais (CNN): possui conexões locais nas primeiras camadas e padrões de peso que são replicados nas unidades de cada camada. (RUSSEL; NORVIG, 2013)
- Redes neurais recorrentes (RNN): permitem ciclos no gráfico computacional, cada ciclo tem um atraso, assim, a unidade pode tomar como entrada um valor calculado a partir da sua saída. (RUSSEL; NORVIG, 2013)

### 2.2.5 REDES NEURAIIS PROFUNDAS

Modelos convencionais de aprendizagem de máquina, ao mostrar-se limitados em sua habilidade para processar dados em sua forma bruta, precisando de considerável especialização e engenharia cuidadosa para transformação dos dados em questão em informação adequada para representação.

Assim, surge a técnica de *deep learning*, ou aprendizagem profunda, que utiliza a estrutura de rede neural organizada em uma grande quantidade de múltiplas camadas para adquirir abstrações mais complexas cada vez que passa por essas camadas. Busca aumentar a mimetização da forma de aprendizado humana e elevar sua capacidade de resolver problemas complexos, tem como vantagem sua capacidade de aprender características ou representações automaticamente. (CHASSAGNON et al., 2020; LECUN; BENGIO; HINTON, 2015)

Esse algoritmo visa treinar circuitos de forma que os caminhos computacionais são longos, permitindo que todas as variáveis de entrada interajam de maneira complexa, permitindo a aprendizagem da representação/abstração antes da aprendizagem da predição. (RUSSEL; NORVIG, 2013)

### 2.2.6 MÁQUINA DE VETORES DE SUPORTE (SVM)

É um algoritmo de aprendizado supervisionado que aprende por exemplificação à atribuir rótulos aos objetos. (NOBLE, 2006) Utilizado principalmente para classificação e regressão em problemas binários, encontrando um hiperplano que maximize a margem, ou distância, entre dois dados mais próximos de cada classe. A margem, neste caso, é o chamado vetor de suporte.

Possui como vantagens a criação de um hiperplano de separação linear e mantém apenas os exemplos que estão mais próximos do plano de separação. Quando a separação pelo hiperplano não é satisfatória o algoritmo se utiliza de funções kernel para a criação de um novo eixo de representação a fim de que os exemplos sejam melhor separados. As SVMs visam minimizar a perda de generalização esperada. (RUSSEL; NORVIG, 2013)

### 2.2.7 k-NEAREST NEIGHBORS (kNN)

O kNN é um algoritmo de aprendizado supervisionado que usa dados rotulados para treinamento e a partir do número de pontos vizinhos (k) mais próximos realiza a previsão de novos dados. Quanto menor o valor de k, mais flexível e sensível a ruído será o modelo, se, por outro lado, o k possuir valor maior, melhor e mais suave a previsão, porém pode não captar com eficiência os padrões do banco de dados. (KRAMER, 2013)

Este algoritmo não constrói um modelo prévio – como as árvores de decisão ou redes neurais –, o que faz é calcular a vizinhança por meio do cálculo de distância/similaridade entre o objeto a ser classificado e a base de exemplos no momento da classificação e utiliza a vizinhança encontrada para a tomada de decisão/predição.

Para otimizar o algoritmo, além do ajuste do k, pode-se aumentar o peso de cada característica e/ou exemplo vizinho utilizado de acordo com a distância do ponto em questão, dando maior peso aos vizinhos mais próximos. Entretanto, considerar todos os exemplos implicará no funcionamento lento do algoritmo. (MITCHELL, 1997)

### **2.3 PERGUNTA DE PESQUISA E OBJETIVOS**

Temos como pergunta de pesquisa deste trabalho a seguinte: como a inteligência artificial pode ajudar a desenvolver uma fórmula biométrica com melhor performance, em olhos com medidas biométricas extremas?

Para tal, este estudo tem à disposição uma grande base de dados de população brasileira. Deste modo, além utilizar algoritmos de aprendizagem de máquina para previsão do poder refrativo da lente intraocular a ser implantada para construir uma nova fórmula biométrica, este estudo realiza a comparação e a análise do desempenho de diversas fórmulas biométricas em uma base significativa e com um perfil diferente dos trabalhos encontrados atualmente na literatura.

Esta abordagem tem o intuito de auxiliar os cirurgiões oftalmológicos do Brasil e do exterior a providenciar melhores cuidados aos seus pacientes portadores de catarata, tornando mais democrático os níveis dos resultados dos grandes centros de reconhecimento internacional para todo e qualquer cirurgião.

### **3 METODOLOGIA**

Este projeto foi avaliado e autorizado pelo Conselho de Ética em Pesquisa da Universidade Federal de Alagoas, registrado sob o número CAAE 89418718.9.0000.5013, sendo aprovado antes da coleta de dados. O estudo seguiu as diretrizes da Declaração de Helsinki. Não foi necessário consentimento informado por se tratar de um estudo retrospectivo em prontuários, sem alterações no tratamento.

#### **3.1 MATERIAIS**

Os dados para o cálculo das fórmulas foram coletados de prontuários e medidas biométricas de um IOLMaster® 500 (Carl Zeiss Meditech AG, Alemanha) no Hospital Oftalmológico de Brasília (Brasília, DF, Brasil). Um único cirurgião oftalmológico experiente realizou todas as cirurgias (W.T.H.).

Todos os pacientes maiores de 18 anos operados para tratamento de catarata entre 2010 e 2018 neste hospital, recebendo implante de LIO não tórica, seguindo protocolo de exame biométrico pré-operatório e avaliação clínica com medição de refração manifesta, acuidade visual não corrigida e corrigida e biomicroscopia, além de avaliações clínicas pós-operatórias realizadas entre 1 e 6 meses após a cirurgia foram incluídos neste estudo. Assim, foram coletados dados de 5124 pacientes, sendo avaliados 7403 olhos, após a aplicação dos critérios de inclusão e exclusão obteve-se o total de 1526 olhos elegíveis.

Os critérios de inclusão e exclusão foram aplicados seguindo em parte o editorial publicado por Hoffer et al (HOFFER et al., 2015), o qual orienta sobre as melhores práticas em estudos de fórmulas de LIOs.

Assim, os critérios de exclusão foram: registros incompletos, defeituosos ou extremamente desviantes (outliers) cujos dados clínicos não puderam ser recuperados; pacientes com complicações pós-operatórias, traumatismos oculares, comorbidades oculares que pudessem influenciar os resultados ou com cirurgias oculares prévias; casos com acuidade visual pior que 20/40 após a cirurgia; e olhos que receberam implante de LIO cujo modelo ocorreu menos de 50 vezes no banco de dados.

Para o cálculo das predições realizadas pelas fórmulas biométricas existentes foram seguidas duas abordagens diferentes: a implementação direta da fórmula quando esta tinha a descrição disponível na literatura ou a utilização de sites e ferramentas disponibilizadas para o devido cálculo do resultado.

De acordo com publicações e errata na literatura, as fórmulas Haigis, Holladay 1, Hoffer Q e SRK/T foram implementadas na linguagem de programação Python. (HAIGIS et al., 2000; HOFFER, 1993; HOLLADAY et al., 1988; RETZLAFF; SANDERS; KRAFF, 1990) A precisão e a confiabilidade dessas implementações foram extensivamente validadas por meio de uma análise comparativa de seus resultados com os obtidos em biômetros comerciais (IOLMaster® 500).

Os resultados das fórmulas de Barrett e Kane foram obtidos inserindo os dados individualmente nos respectivos sites que hospedam calculadoras para essas fórmulas. Os parâmetros usados para calcular as fórmulas foram comprimento axial (AL), ceratometria (K), ACD e distância branco a branco (WTW). Não foram inseridos parâmetros opcionais (por exemplo, espessura da lente, espessura central da córnea) devido à falta de disponibilidade consistente no banco de dados.

Para o desenvolvimento dos modelos preditivos também foi utilizada a linguagem Python junto com a biblioteca scikit-learn, contando com o auxílio das bibliotecas pandas, numpy e seaborn para a manipulação dos dados e construção dos gráficos apresentados neste documento.

### **3.2 MÉTODO**

Seguindo a descoberta de conhecimento em base de dados (KDD), os dados biométricos e epidemiológicos dos pacientes elegíveis foram selecionados, passaram pelo processo de limpeza e enriquecimento, seguido pela transformação e mineração desses dados.

Finalizadas as primeiras etapas de seleção, pré-processamento e transformação dos dados do KDD, foram utilizados como entrada para a construção dos modelos preditivos os seguintes dados biométricos do olho: comprimento axial (AL), profundidade de câmara anterior (ACD) e ceratometria média (K). Sendo esta última calculada a partir da média das curvaturas K1 e K2.

Como não havia na literatura uma indicação, com evidência, de quais algoritmos de aprendizagem de máquina seriam mais adaptados para o nosso problema foram testados os algoritmos mais comumente utilizados para a tarefa de regressão. Um outro ponto considerado na escolha dos algoritmos foi a busca por diferentes características, tanto para representação do conhecimento utilizado quanto para a estratégia de convergência no momento da aprendizagem do modelo preditivo.

Deste modo, foram testados os algoritmos: regressão linear (clássica, Lasso e Ridge), *k*-NN Regressor, *Gradient Boosting Regressor* (GBR), *Support Vector Regression* (SVR) e a Rede Neural Perceptron Multicamadas (MLP).

Todos os algoritmos tiveram seus hiperparâmetros otimizados de acordo com o teste de todas as combinações das variáveis demonstradas na Tabela 5, do apêndice. O treinamento e a avaliação dos modelos foi realizado utilizando o processo de validação cruzada (*k-fold cross-validation*), dividindo a base de forma aleatória em 5 partes proporcionais de modo que o treinamento foi realizado com o equivalente a 80% da base e o teste/validação feito com os 20% restantes, seguindo o processo de validação cruzada.

As constantes das fórmulas de Barrett, Hoffer Q, Holladay 1 e SRK/T foram calculadas por otimização recorrente, visando centralizar o erro médio o mais próximo possível de zero para cada modelo de LIO no banco de dados. Devido a dificuldades de automação e complexidade, as constantes usadas para calcular as fórmulas de Kane e Haigis foram obtidas do banco de dados *User Group for Laser Interference Biometry* (ULIB).<sup>11</sup>

Os dados demográficos da população foram descritos por frequência, para dados categóricos (sexo, lateralidade do olho, subgrupo biométrico e subgrupo de modelo de LIO) e em média e desvio padrão para dados numéricos (idade, erro de previsão de poder de LIO, AL, ACD, K).

Para fins de análise, o banco de dados também foi dividido em subgrupos por AL, ACD e K. Olhos curtos foram classificados como aqueles com AL menor que 22,5 mm, enquanto olhos longos foram definidos como aqueles com AL maior que 25,5 mm.

---

<sup>11</sup> Disponível em: <http://ocusoft.de/ulib/>

Olhos com ACD menor que 2,5 mm foram considerados como tendo câmara anterior rasa, enquanto aqueles com medida maior que 3,5 mm foram considerados como tendo câmara anterior profunda. As córneas foram classificadas como planas se a leitura de K fosse menor que 42D e muito curvas se a leitura de K fosse maior que 46D.

A análise estatística para comparar o desempenho das fórmulas biométricas foi realizada seguindo as recomendações publicadas por Hoffer et al. (HOFFER et al., 2015) apresentando erro absoluto médio (MAE), erro absoluto mediano (MedAE), erro padrão amostral e intervalo de confiança (IC) de 95% em torno da média. O relatório também fornece uma análise detalhada da porcentagem de indivíduos do banco de dados e seus subgrupos com erros de refração menores que 0,5 D e 1,0 D de equivalente esférico absoluto.

## 4 RESULTADOS

Foram coletadas informações dos prontuários de 5124 pacientes, sendo avaliados 7403 olhos, após a aplicação dos critérios de inclusão e exclusão obteve-se o total de 1526 olhos elegíveis. Os olhos selecionados foram então separados de acordo com o comprimento axial nos seguintes subgrupos: curto (< 22.5mm), médio (22.5mm a 25.5mm) e longo (> 25.5mm); de acordo com a profundidade da câmara anterior em: rasa (< 2.5mm), média (2.5 a 3.5mm) e ampla (> 3.5mm); de acordo com a curvatura anterior da córnea em: leve (< 42D), média (42 a 46D) e acentuada (> 46D).

Foram observados idade, gênero, profundidade da câmara anterior (ACD), comprimento axial (AL), medidas ceratométricas (K1, K1 mm, K1°, K2, K2 mm, K2°) e erro absoluto das fórmulas: Barrett, Haigis, Hoffer Q, Holladay 1, Kane e SRK/T.

Dentre o total de 1526 pacientes, o gênero feminino representou 843 (55.24%) dos casos, o masculino 598 (39.19%) e 85 (5.57%) pacientes não tiveram o gênero divulgado. A idade média no momento da cirurgia foi 67.06 anos. (Tabela 1)

Na análise das variáveis, o valor médio do ACD foi igual a 3.11 com desvio padrão (DP) de  $\pm 0.43$ , 6.16% (94) apresentaram câmara anterior rasa e 15.27% (233), profunda. A média do AL foi de  $23.49 \pm 1.14$ , com 16.78% (256) sendo olhos curtos e 4.85% (74), longos. De acordo com a curvatura da córnea, a média do K foi igual a  $43.92 \pm 1.72$ , onde 11.01% (168) tinham córnea plana e 10.35% (158) tinham córnea com curvatura acentuada.

Tabela 1. Valores demográficos

|                            | <b>Média <math>\pm</math> Desvio padrão</b> | <b>Intervalo</b> |
|----------------------------|---|------------------|
| <b>Idade (anos)</b>        | 67.06 $\pm$ 8.65                            | [45.00, 94.00]   |
| <b>AL (mm)</b>             | 23.49 $\pm$ 1.14                            | [20.15, 30.96]   |
| <b>ACD (mm)</b>            | 3.11 $\pm$ 0.43                             | [2.03, 5.23]     |
| <b>K (D)</b>               | 43.92 $\pm$ 1.72                            | [32.92, 50.18]   |
| <b>WTW (mm)</b>            | 11.98 $\pm$ 0.43                            | [9.90, 13.30]    |
| <b>Dioptria da LIO (D)</b> | 21.23 $\pm$ 3.03                            | [6.00, 34.00]    |

Foram realizados testes preditivos com os seguintes modelos: regressão linear, regressão Ridge, regressão Lasso, kNN, *Gradient Boosting Regression (GBR)*,

*Support Vector Regression* (SVR) e *Multilayer Perceptron* (MLP). Dentre esses, o modelo baseado em SVR foi o que apresentou o melhor desempenho quando comparados de acordo com o erro absoluto médio e erro absoluto mediano. (Tabela 2)

Tabela 2. Modelos preditivos: Erro previsto e cálculo das LIOs em todos os intervalos de AL

|                                      | <b>LR</b> | <b>Ridg<br/>e</b> | <b>Lasso</b> | <b>kNN</b> | <b>GBR</b> | <b>SVR</b> | <b>MLP</b> |
|--------------------------------------|-----------|-------------------|--------------|------------|------------|------------|------------|
| <b>Poder médio refrativo da LIO</b>  | 21.04     | 21.04             | 21.04        | 21.11      | 21.05      | 21.05      | 21.05      |
| <b>Erro Absoluto Médio (MAE)</b>     | 0.428     | 0.428             | 0.428        | 0.830      | 0.403      | 0.370      | 0.405      |
| <b>Erro Absoluto Mediano (MedAE)</b> | 0.312     | 0.312             | 0.435        | 0.601      | 0.316      | 0.300      | 0.311      |
| <b>Desvio padrão</b>                 | 0.469     | 0.469             | 0.469        | 0.923      | 0.375      | 0.308      | 0.363      |
| <b>Olhos com erro previsto (%)</b>   |           |                   |              |            |            |            |            |
| <b>± 0.5 D</b>                       | 68.5%     | 68.6<br>%         | 68.6%        | 43.6<br>%  | 70.7<br>%  | 73.9%      | 70.0%      |
| <b>± 1.0 D</b>                       | 92.5%     | 92.5<br>%         | 92.5%        | 73.0<br>%  | 94.3<br>%  | 95.8%      | 93.6%      |

Tendo como parâmetro o percentual de olhos com erro previsto entre -0.5D e +0.5D, o modelo baseado em SVR apresentou 73.9% dos olhos nesse intervalo, demonstrando melhor desempenho dentre os demais (Figura 1).

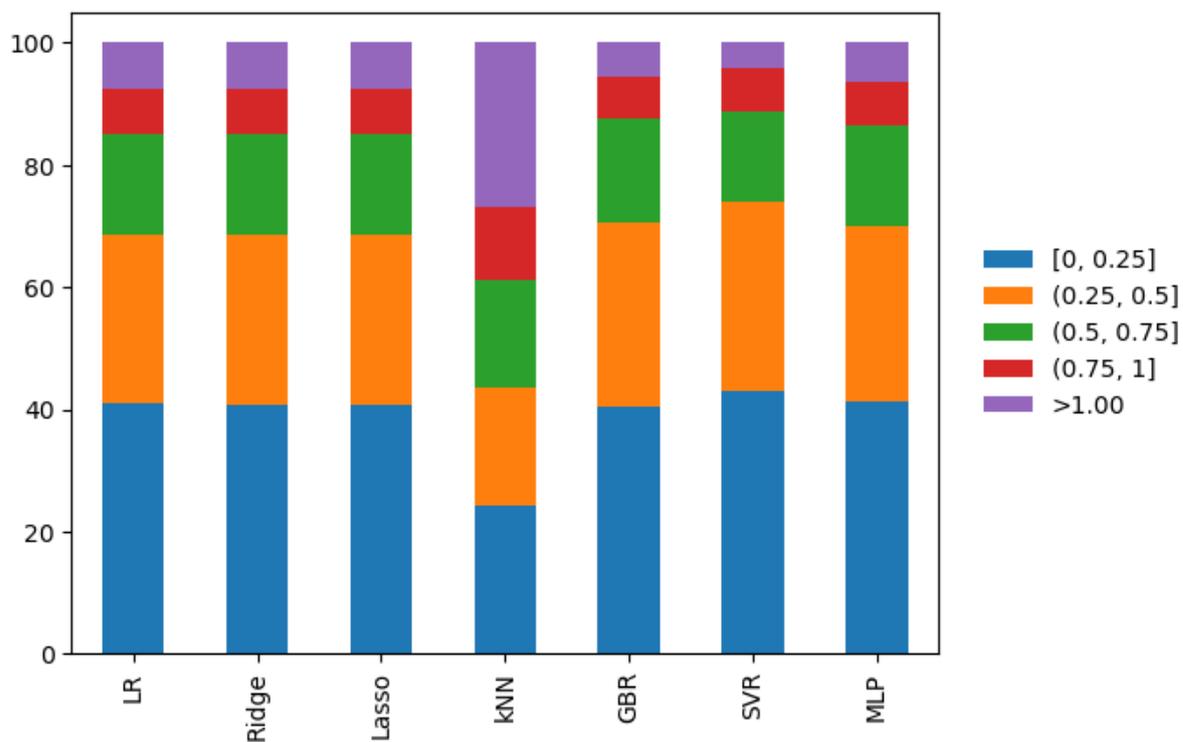


Figura 1. Modelos preditivos: Porcentagem de olhos por faixa de MAE

Ao avaliar os subgrupos de olhos de acordo com o comprimento axial e a curvatura da córnea, o modelo SVR também apresentou os melhores resultados em todos os intervalos. Já em relação à profundidade da câmara anterior, o modelo MLP demonstrou melhor performance nos resultados em olhos com  $ACD < 2.38$  mm e  $ACD > 4.12$  mm (Figura 2).

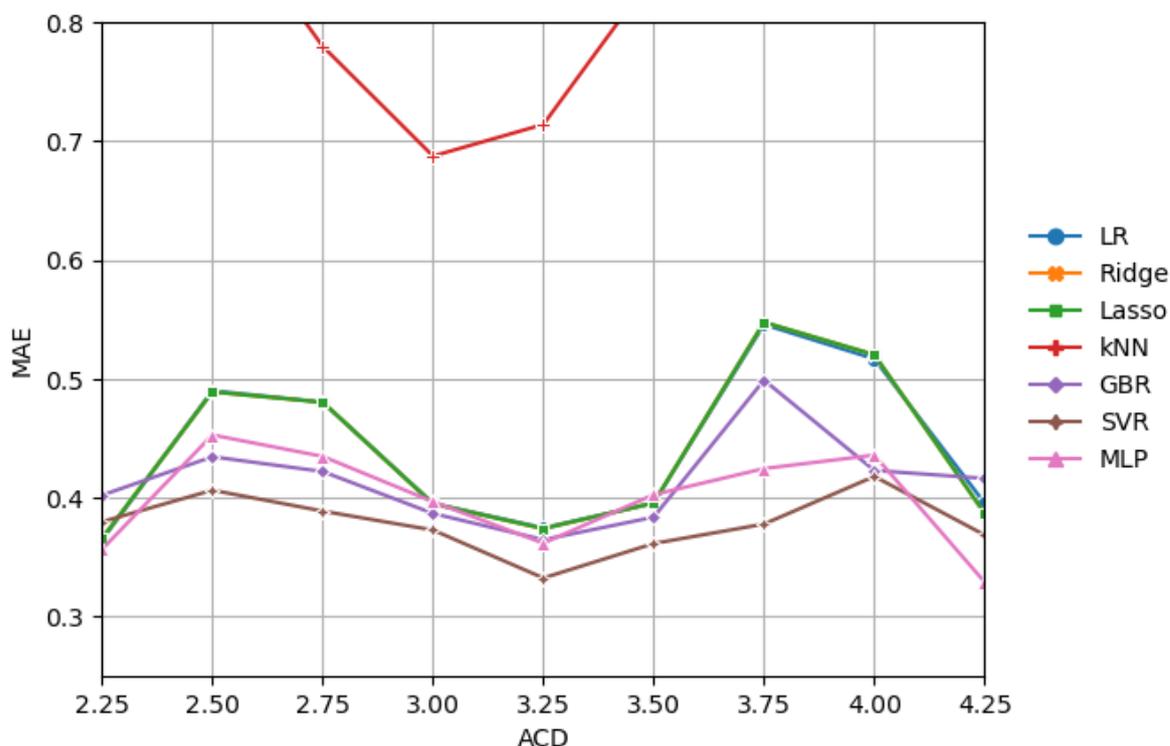


Figura 2. Modelos preditivos: MAE versus ACD

Diante do uso das fórmulas para o cálculo do poder refrativo da lente intraocular a ser implantada, o valor médio foi semelhante entre todas as fórmulas e modelos analisados, entre 21.04D e 21.16D.

Ao utilizar a fórmula Barrett para o cálculo do poder de convergência das LIOs, foi obtido erro absoluto médio (MEA) de  $0.393 \pm 0.329$ . Através da Haigis foi obtido  $0.426 \pm 0.349$ . Utilizando-se a fórmula Hoffer Q, o MAE foi de  $0.427 \pm 0.355$ . No caso da Holladay 1, obteve-se  $0.387 \pm 0.347$ . A fórmula Kane atingiu um erro absoluto médio de  $0.409 \pm 0.360$  e a SRK/T demonstrou um MAE de  $0.410 \pm 0.361$ . Enquanto isso, os dois melhores modelos desenvolvidos, MLP e SVR, apresentaram respectivamente MAE de  $0.405 \pm 0.363$  e  $0.370 \pm 0.308$  (Tabela 3).

Ao avaliar o erro refracional previsto, o melhor resultado foi observado com o modelo SVR, seguido pelas fórmulas Barrett e Holladay 1.

Tabela 3. Fórmulas e modelos preditivos: Erro previsto e cálculo das LIOs em todos intervalos de AL

|                                    | <i>Barrett</i> | <i>Haigis</i> | <i>Hoffer Q</i> | <i>Holladay 1</i> | <i>Kane</i> | <i>SRK/T</i> | <i>Nova fórmula</i> |            |
|------------------------------------|----------------|---------------|-----------------|-------------------|-------------|--------------|---------------------|------------|
|                                    |                |               |                 |                   |             |              | <i>MLP</i>          | <i>SVR</i> |
| <b>LIO</b>                         | 21.04          | 21.03         | 21.04           | 21.04             | 21.16       | 21.04        | 21.05               | 21.05      |
| <b>MAE</b>                         | 0.393          | 0.426         | 0.427           | 0.387             | 0.409       | 0.410        | 0.405               | 0.370      |
| <b>MedAE</b>                       | 0.333          | 0.348         | 0.338           | 0.304             | 0.306       | 0.319        | 0.310               | 0.300      |
| <b>Desvio padrão</b>               | 0.329          | 0.349         | 0.355           | 0.347             | 0.360       | 0.361        | 0.363               | 0.308      |
| <b>Olhos com erro previsto (%)</b> |                |               |                 |                   |             |              |                     |            |
| <b>± 0.5 D</b>                     | 73.04%         | 60.13%        | 58.62%          | 72.8%             | 70.9%       | 68.9%        | 70%                 | 73.9%      |
| <b>± 1.0 D</b>                     | 95.01%         | 87.14%        | 86.82%          | 89.9%             | 93.25%      | 88.85%       | 93.6%               | 95.8%      |

Ao realizar análise comparativa dos resultados mostrados na Tabela 3, observou-se que o modelo SVR, seguido pelas fórmulas Barrett e Holladay 1 demonstraram desempenho superior na população geral, conforme evidenciado pelo MedAE e pelas métricas de desempenho clínico, que mediram a porcentagem de indivíduos com erro absoluto de menor que 0,5 D.

No caso de olhos curtos, a fórmula Holladay 1 teve melhor desempenho, seguida do modelo SVR. Em olhos longos, o melhor resultado se deu com o modelo SVR, Barrett e Kane, respectivamente.

Ao avaliar a profundidade da câmara anterior, é possível observar que olhos com ACD na média (entre 2.5mm e 3.5mm) possuem melhores resultados com o modelo SVR, Barrett e Holladay 1 (Figura 5).

Com relação ao ACD, observou-se que o modelo SVR foi melhor em olhos com câmara anterior de até cerca de 4mm e o modelo MLP se destacou nos olhos com valores de ACD maiores que 4mm. Já em olhos com câmara anterior rasa (ACD < 2.5 mm), tanto as fórmulas quanto os dois melhores modelos apresentaram resultados variáveis, mas todos com MAE abaixo de 0.5D.

O comportamento das fórmulas e modelos, quando comparado o erro absoluto médio e a ceratometria central da córnea, repete-se em olhos com o valor de K dentro da normalidade, pois apresentam os melhores resultados com o modelo SVR, enquanto nos extremos existe uma grande variação de resultados (Figura 4).

Nos casos de córneas mais planas, a fórmula Barrett e o modelo SVR tiveram destaque nos resultados e em córneas com maior curvatura, os melhores resultados foram obtidos pela Barrett, Holladay 1 e SVR.

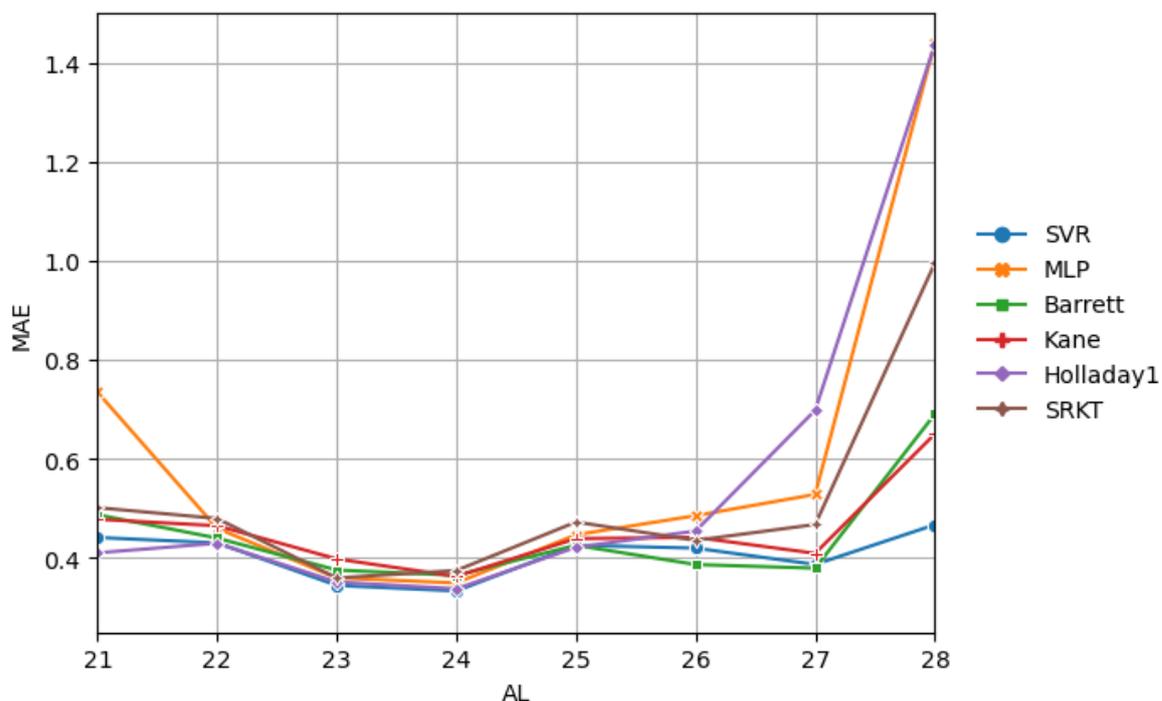


Figura 3. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Comprimento Axial

Vale ressaltar que há maior variabilidade nos resultados quando se trata de valores extremos de AL, K e ACD, conforme ilustrado nas Figuras 3 a 6.

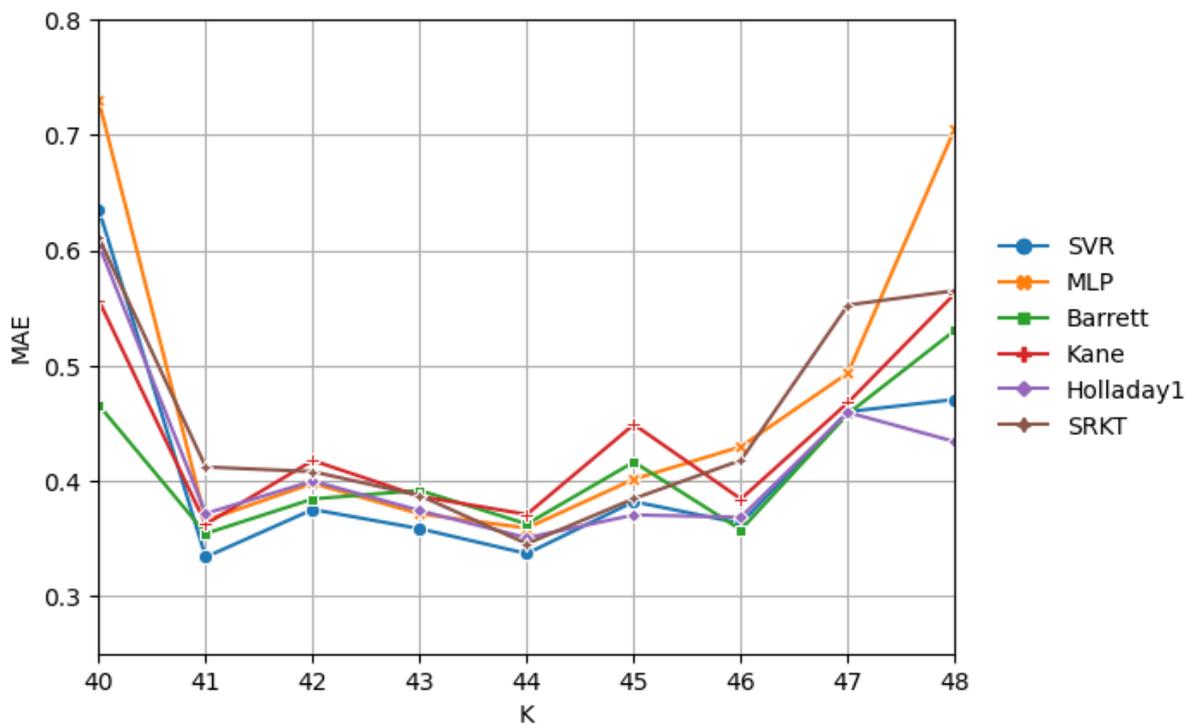


Figura 4. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Ceratometria central

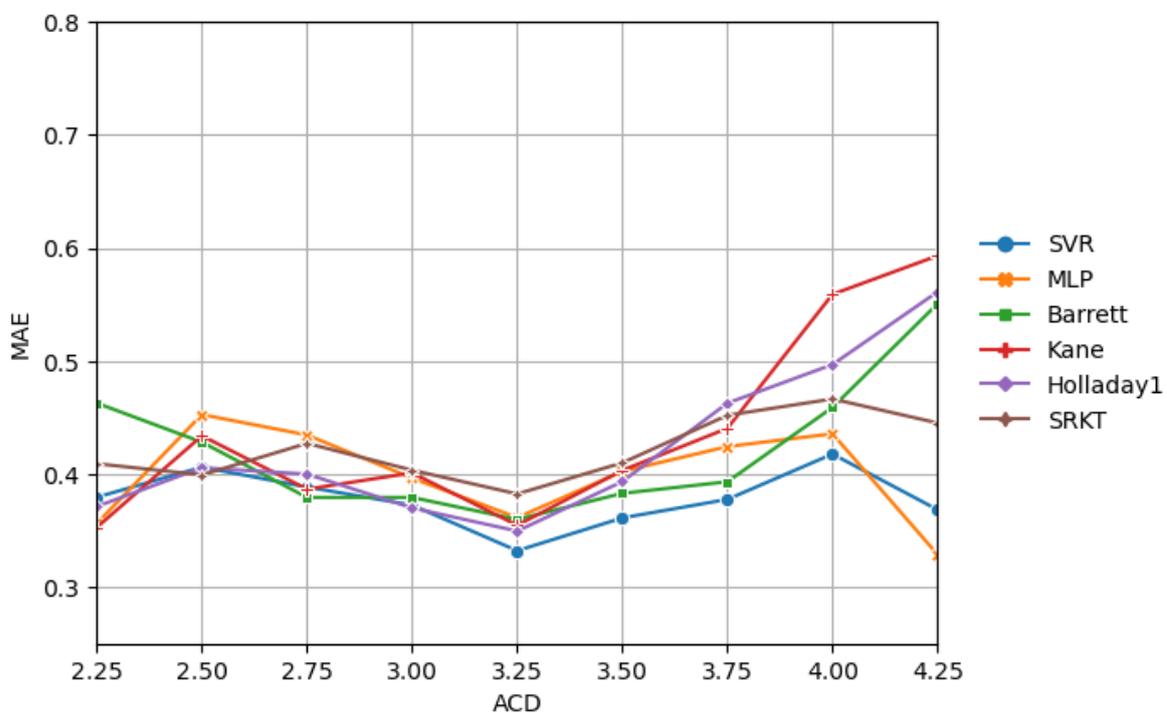


Figura 5. Fórmulas e Modelos preditivos: Erro Absoluto Médio versus Profundidade de câmara anterior

O equivalente esférico a menos de  $\pm 0.5D$  de distância do alvo pode ser observado na Tabela 2 e os resultados foram de 73.9% dos olhos no modelo SVR, 72.8% na

fórmula Holladay 1, 70.9% na Kane, 70.8% na Barrett, 70% no modelo MLP, 60.13% na Haigis, 68.9% na SRK/T e 58.62% na Hoffer Q.

A população em estudo teve maior prevalência de olhos com comprimento axial médio, seguida por olhos curtos e, por último, longos.

## 5 DISCUSSÃO

As características populacionais deste estudo foram semelhantes a outros latino-americanos e globais, nos quais o gênero feminino representa a maioria dos pacientes e a faixa etária média submetida à cirurgia de catarata é de idosos. (MEDINA, 2019; MELLES; HOLLADAY; CHANG, 2018)

De acordo com o tamanho do olho, usando como base o comprimento axial, a distribuição encontrada neste estudo foi semelhante à encontrada por Kane et al e Darcy et al. (DARCY et al., 2020; KANE et al., 2017) Nos estudos realizados por Melles et al e Cooke et al também houve maior prevalência de olhos médios, entretanto a quantidade de olhos longos foi maior que a de curtos. (COOKE; COOKE, 2016b; MELLES; HOLLADAY; CHANG, 2018)

Ao analisar dados de estudos que buscam um melhor resultado refrativo no pós-operatório de catarata, pode ser observado que a porcentagem de pacientes com equivalente esférico menor ou igual a  $\pm 0.5D$  varia entre 40% e 60%, sendo poucos os que atingem valores próximos ou superiores a 80% nesse intervalo do erro refrativo residual. (HAHN et al., 2011; KUGELBERG; LUNDSTRÖM, 2008)

Observou-se que a fórmula desenvolvida por meio de uma regressão vetorial (SVR) utilizando máquina de vetores de suporte (SVM) apresentou melhor desempenho geral, seguida pelas fórmulas Holladay 1 e Barrett, mediante o MAE e as recomendações internacionais do padrão da cirurgia de catarata.

No que diz respeito às fórmulas avaliadas, Melles et al obteve melhor desempenho geral por meio da Barrett, tanto de acordo com o MAE quanto com a porcentagem de olhos com visão boa ( $\pm 0.5D$ ), seguida pela Haigis, Holladay 1, SRK/T e Hoffer Q. Para olhos curtos, a Barrett apresentou melhor resultado e o pior foi com a Hoffer Q. No caso de olhos com comprimento axial médio, dentre as fórmulas avaliadas, Melles encontrou desempenho semelhante com a Barrett, a SRK/T, a Haigis e a Holladay 1. (MELLES; HOLLADAY; CHANG, 2018)

A distribuição do MAE de acordo com os valores de K central encontradas na população em questão foi semelhante à observada por Melles, tendo melhores resultados em valores entre 43D e 44D.

Com relação à profundidade da câmara anterior, dentre as fórmulas apresentadas, Melles encontrou que a Barrett e a Haigis obtiveram melhores resultados nos casos com ACD de 2.75mm a 3.75mm.

Já de acordo com Darcy et al, dentre as fórmulas estudadas, a Kane com melhor performance geral, em olhos curtos, médios e longos. A Barrett apresentou pior desempenho em olhos curtos, porém foi se destacou com bons resultados em olhos médios e longos. Enquanto a Haigis ficou em último lugar em olhos médios e a Holladay 1 em olhos longos.(DARCY et al., 2020)

Dentre as fórmulas analisadas neste estudo, Kane et al encontrou menor MAE e melhor desempenho quanto ao resultado dióptrico  $\pm 0.5D$  com a fórmula Barrett. (KANE et al., 2017) A Holladay 1 teve o próximo melhor desempenho a partir do MAE e 70.1% de olhos com  $\pm 0.5D$ . Por fim, a Hoffer Q apresentou maior erro absoluto médio e 67.9% de olhos com visão boa. (Tabela 4)

Kane et al encontrou melhores resultados com a Kane em olhos de comprimento axial médio. Em olhos com comprimento axial longo, as fórmulas Haigis e Kane apresentaram melhor MAE. Já em olhos curtos, a Holladay 1 permaneceu com melhor desempenho. (CONNELL; KANE, 2019)

De acordo com Cooke et al, utilizando biometria por interferometria óptica, foram alcançados os seguintes resultados: a Haigis foi melhor, seguida pela Holladay 1, Hoffer Q e SRK/T, todas com ambos os parâmetros congruentes. Demonstrou melhor desempenho para olhos curtos com a Haigis e a Holladay 1, enquanto para olhos longos tal resultado foi atingido pela Haigis e pela SRK/T. (COOKE; COOKE, 2016b)

No estudo de Hipólito-Fernandes et al, ficou evidenciado que, de acordo com o MAE, os melhores resultados foram atingidos respectivamente pelas fórmulas Kane e Barrett. Ao observar os olhos com erro previsto entre  $\pm 0.5D$ , a Barrett apresentou melhor desempenho, seguida pela Kane. (HIPÓLITO-FERNANDES et al., 2021)

Deve-se destacar o fato de que mais de 86% dos pacientes apresentaram um resultado de equivalente esférico entre  $\pm 1.0D$  em todas as fórmulas estudadas e nos dois melhores modelos preditivos apresentados. Além disso, a quantidade de pacientes entre  $\pm 0.5D$  variou de 58.6% a 73.0%, mostrando melhores resultados que os apresentados em trabalhos latino-americanos semelhantes.

González et al também propôs em seu estudo uma nova fórmula baseada em aprendizagem de máquina, além das variáveis utilizadas neste estudo, foram usadas a média da ceratometria posterior e espessura do cristalino. Nos resultados refrativos, a fórmula proposta (Karmona) obteve 90.38% dos olhos com equivalente esférico entre  $\pm 0.5D$ , seguido pela Haigis com 86.54%. Entretanto, a amostra do estudo foi restrita, contando com um total de 260 olhos para realizar o treinamento e a validação do modelo de aprendizagem, além de não ter estratificado os olhos em grupos de acordo com o comprimento axial, profundidade de câmara anterior e ceratometria, por exemplo. (CARMONA GONZÁLEZ; PALOMINO BAUTISTA, 2021)

Tabela 4. Comparação dos resultados

| SRK/T | Kane  | Holladay 1 | Hoffer Q | Haigis | Barrett | Fórmulas |    |                   | Melles et al <sup>9</sup> |      |       | Darcy et al <sup>11</sup> |                   |       | Kane et al <sup>12</sup> |       |       | Cooke et al <sup>52</sup> |       |       | Hipólito-Fernandes et al <sup>54</sup> |       |                   |       |       |       |       |       |
|-------|-------|------------|----------|--------|---------|----------|----|-------------------|---------------------------|------|-------|---------------------------|-------------------|-------|--------------------------|-------|-------|---------------------------|-------|-------|--|-------|-------------------|-------|-------|-------|-------|-------|
|       |       |            |          |        |         | MAE      | DP | Erro previsto (%) |                           |      | MAE   | DP                        | Erro previsto (%) |       |                          | MAE   | DP    | Erro previsto (%)         |       |       | MAE                                    | DP    | Erro previsto (%) |       |       |       |       |       |
|       |       |            |          |        |         |          |    | ±0.25             | ±0.5                      | ±1.0 |       |                           | ±0.25             | ±0.5  | ±1.0                     |       |       | ±0.25                     | ±0.5  | ±1.0  |  |       | ±0.25*            | ±0.5  | ±1.0  | ±0.25 | ±0.5  | ±1.0  |
| 0.360 | -     | 0.351      | 0.369    | 0.338  | 0.311   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 41.7%             | 70.7% | 94.7%                    | 0.350 | 0.469 | 47.9%                     | 75.2% | 96.4% | -                                      | -     | -                 | 0.326 | 0.398 | 42.3% | 85.6% | 99.0% |
| 0.463 | -     | 0.453      | 0.473    | 0.437  | 0.404   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 49.8%             | 80.8% | 97.8%                    | 0.350 | 0.469 | 47.9%                     | 75.2% | 96.4% | -                                      | -     | -                 | 0.326 | 0.398 | 42.3% | 85.6% | 99.0% |
| 43.3% | -     | 44.7%      | 42.5%    | 46.1%  | 49.8%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 49.8%             | 80.8% | 97.8%                    | 0.350 | 0.469 | 47.9%                     | 75.2% | 96.4% | -                                      | -     | -                 | 0.326 | 0.398 | 42.3% | 85.6% | 99.0% |
| 74%   | -     | 75%        | 73.0%    | 77.1%  | 80.8%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 80.8%             | 97.8% | 97.8%                    | 0.350 | 0.469 | 47.9%                     | 75.2% | 96.4% | -                                      | -     | -                 | 0.326 | 0.398 | 42.3% | 85.6% | 99.0% |
| 96.5% | -     | 96.8%      | 96.2%    | 97.3%  | 97.8%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 97.8%             | 97.8% | 97.8%                    | 0.350 | 0.469 | 47.9%                     | 75.2% | 96.4% | -                                      | -     | -                 | 0.326 | 0.398 | 42.3% | 85.6% | 99.0% |
| 0.403 | 0.377 | 0.397      | 0.410    | 0.405  | 0.390   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.390 | 0.505                     | 0.390             | 0.505 | 0.390                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.344 | 0.420 | 41.2% | 76.3% | 97.9% |
| 0.522 | 0.490 | 0.512      | 0.527    | 0.521  | 0.505   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.505 | 0.521                     | 39.9%             | 69%   | 94.3%                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.344 | 0.420 | 41.2% | 76.3% | 97.9% |
| 39.9% | 42.6% | 40.2%      | 39.0%    | 39.9%  | 41.7%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.505 | 0.521                     | 39.9%             | 69%   | 94.3%                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.344 | 0.420 | 41.2% | 76.3% | 97.9% |
| 69.1% | 72.0% | 69.6%      | 68.1%    | 69%    | 70.7%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.505 | 0.521                     | 69%               | 94.3% | 94.3%                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.344 | 0.420 | 41.2% | 76.3% | 97.9% |
| 93.9% | 95.2% | 94.4%      | 94.0%    | 94.3%  | 94.7%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.505 | 0.521                     | 94.3%             | 94.3% | 94.3%                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.344 | 0.420 | 41.2% | 76.3% | 97.9% |
| 0.369 | 0.329 | 0.358      | 0.381    | 0.358  | 0.350   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.350 | 0.469                     | 47.9%             | 75.2% | 96.4%                    | 0.358 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.341 | 0.432 | 39.2% | 73.2% | 99.0% |
| 0.489 | 0.445 | 0.475      | 0.499    | 0.480  | 0.469   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.469 | 0.469                     | 47.9%             | 75.2% | 96.4%                    | 0.480 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.432 | 0.432 | 39.2% | 73.2% | 99.0% |
| 44.7% | 52.4% | 45.1%      | 40.3%    | 45.5%  | 47.9%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.469 | 0.469                     | 47.9%             | 75.2% | 96.4%                    | 0.480 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.432 | 0.432 | 39.2% | 73.2% | 99.0% |
| 72.1% | 77.9% | 73.4%      | 73.6%    | 75.5%  | 75.2%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.469 | 0.469                     | 47.9%             | 75.2% | 96.4%                    | 0.480 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.432 | 0.432 | 39.2% | 73.2% | 99.0% |
| 95.9% | 96.6% | 96.1%      | 95.5%    | 96.1%  | 96.4%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.469 | 0.469                     | 47.9%             | 75.2% | 96.4%                    | 0.480 | 0.480 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.432 | 0.432 | 39.2% | 73.2% | 99.0% |
| 0.346 | -     | 0.326      | 0.341    | 0.319  | -       | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | -     | -                         | -                 | -     | -                        | 0.319 | 0.401 | 45.5%                     | 75.5% | 96.1% | 0.319                                  | 0.401 | -                 | 0.341 | 0.432 | 39.2% | 73.2% | 99.0% |
| 0.440 | -     | 0.414      | 0.432    | 0.401  | -       | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | -     | -                         | -                 | -     | -                        | 0.401 | 0.401 | 45.5%                     | 75.5% | 96.1% | 0.401                                  | 0.401 | -                 | 0.432 | 0.432 | 39.2% | 73.2% | 99.0% |
| -     | -     | -          | -        | -      | -       | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | -     | -                         | -                 | -     | -                        | -     | -     | 45.5%                     | 75.5% | 96.1% | -                                      | -     | -                 | -     | -     | -     | -     | -     |
| 75.1% | -     | 79.5%      | 77.0%    | 79.8%  | -       | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | -     | -                         | -                 | -     | -                        | 79.8% | 79.8% | 45.5%                     | 75.5% | 96.1% | 79.8%                                  | 79.8% | -                 | 0.336 | 0.410 | 45.4% | 74.2% | 99.0% |
| 98.1% | -     | 98.7%      | 97.4%    | 98.7%  | -       | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | -     | -                         | -                 | -     | -                        | 98.7% | 98.7% | 45.5%                     | 75.5% | 96.1% | 98.7%                                  | 98.7% | -                 | 0.336 | 0.410 | 45.4% | 74.2% | 99.0% |
| 0.363 | 0.306 | 0.336      | 0.357    | 0.344  | 0.326   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.326 | 0.410                     | 45.4%             | 74.2% | 99.0%                    | 0.344 | 0.420 | 45.5%                     | 75.5% | 96.1% | 0.344                                  | 0.420 | -                 | 0.336 | 0.410 | 45.4% | 74.2% | 99.0% |
| 0.436 | 0.383 | 0.410      | 0.435    | 0.420  | 0.398   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.398 | 0.398                     | 42.3%             | 85.6% | 99.0%                    | 0.420 | 0.420 | 45.5%                     | 75.5% | 96.1% | 0.420                                  | 0.420 | -                 | 0.410 | 0.410 | 45.4% | 74.2% | 99.0% |
| 35.1% | 46.4% | 45.4%      | 39.2%    | 41.2%  | 42.3%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.398 | 0.398                     | 42.3%             | 85.6% | 99.0%                    | 0.420 | 0.420 | 45.5%                     | 75.5% | 96.1% | 0.420                                  | 0.420 | -                 | 0.410 | 0.410 | 45.4% | 74.2% | 99.0% |
| 74.2% | 81.4% | 74.2%      | 73.2%    | 76.3%  | 85.6%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.398 | 0.398                     | 85.6%             | 99.0% | 99.0%                    | 0.420 | 0.420 | 45.5%                     | 75.5% | 96.1% | 0.420                                  | 0.420 | -                 | 0.410 | 0.410 | 45.4% | 74.2% | 99.0% |
| 97.9% | 99.0% | 99.0%      | 99.0%    | 97.9%  | 99.0%   | MAE      | DP | ±0.25             | ±0.5                      | ±1.0 | 0.398 | 0.398                     | 99.0%             | 99.0% | 99.0%                    | 0.420 | 0.420 | 45.5%                     | 75.5% | 96.1% | 0.420                                  | 0.420 | -                 | 0.410 | 0.410 | 45.4% | 74.2% | 99.0% |

## 6 CONCLUSÃO

Este estudo apresenta grande relevância devido ao número de olhos estudados e à heterogeneidade da amostra analisada. Traduzindo-se em uma alta capacidade de generalização para esta e outras populações.

Os resultados encontrados neste estudo foram compatíveis com os relatados na literatura médica em populações mundiais quando avaliadas a fórmulas biométricas existentes. Entretanto, o modelo preditivo desenvolvido e apresentado neste estudo demonstrou melhores resultados na escolha da LIO indicada na cirurgia de catarata, pois obteve menor MAE e maior porcentagem de olhos com erro previsto dentro do intervalo  $\pm 0.5D$ .

Assim, a fórmula proposta neste estudo permite o cirurgião selecionar a melhor fórmula considerando a biometria ocular de cada paciente.

## REFERÊNCIAS

- ABIRAMI, S.; CHITRA, P. Chapter Fourteen - Energy-efficient edge based real-time healthcare support system. Em: RAJ, P.; EVANGELINE, P. (Eds.). **The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases**. Advances in Computers. [s.l.] Elsevier, 2020. v. 117p. 339–368.
- BARRETT, G. D. An improved universal theoretical formula for intraocular lens power prediction. **Journal of Cataract and Refractive Surgery**, v. 19, n. 6, p. 713–720, nov. 1993.
- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 1 out. 2001.
- CARMONA GONZÁLEZ, D.; PALOMINO BAUTISTA, C. Accuracy of a new intraocular lens power calculation method based on artificial intelligence. **Eye**, v. 35, n. 2, p. 517–522, 1 fev. 2021.
- CHASSAGNON, G. et al. Deep learning: definition and perspectives for thoracic imaging. **European Radiology**, v. 30, n. 4, p. 2021–2030, 1 abr. 2020.
- COLENBRANDER, M. C. Calculation of the power of an iris clip lens for distant vision. **British Journal of Ophthalmology**, v. 57, n. 10, p. 735–740, 1 out. 1973.
- CONNELL, B. J.; KANE, J. X. Comparison of the Kane formula with existing formulas for intraocular lens power selection. **BMJ Open Ophthalmology**, v. 4, n. 1, 2019.
- COOKE, D. L.; COOKE, T. L. Comparison of 9 intraocular lens power calculation formulas. **Journal of Cataract & Refractive Surgery**, v. 42, n. 8, 2016a.
- COOKE, D. L.; COOKE, T. L. Prediction accuracy of preinstalled formulas on 2 optical biometers. **Journal of Cataract & Refractive Surgery**, v. 42, n. 3, 2016b.
- CUTLER, A.; CUTLER, D. R.; STEVENS, J. R. Random Forests. Em: ZHANG, C.; MA, Y. (Eds.). **Ensemble Machine Learning: Methods and Applications**. New York, NY: Springer New York, 2012. p. 157–175.
- DARCY, K. et al. Assessment of the accuracy of new and updated intraocular lens power calculation formulas in 10 930 eyes from the UK National Health Service. **Journal of cataract and refractive surgery**, v. 46, n. 1, p. 2–7, jan. 2020.
- DE ALMEIDA FERREIRA, G. et al. Outcomes of and barriers to cataract surgery in Sao Paulo State, Brazil. **BMC Ophthalmology**, v. 17, n. 1, p. 259, 22 dez. 2017.
- EL NAQA, I.; MURPHY, M. J. What Is Machine Learning? Em: EL NAQA, I.; LI, R.; MURPHY, M. J. (Eds.). **Machine Learning in Radiation Oncology: Theory and Applications**. Cham: Springer International Publishing, 2015. p. 3–11.
- FARAMARZI, A.; AGHAJANI, A.; GHIASIAN, L. Accuracy of Various Intraocular Lens Power Calculation Formulas in Steep Corneas. **Journal of Ophthalmic & Vision Research**, v. 12, p. 385–389, 2017.

GÉRON, A. **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow**. [s.l.] Alta Books, 2019.

GÖKCE, S. E. et al. Intraocular lens power calculations in short eyes using 7 formulas. **Journal of Cataract & Refractive Surgery**, v. 43, n. 7, 2017.

GRANADO, J. D. et al. Resultados refractivos y selección del poder del lente Intraocular en cirugía de catarata por facoemulsificación. 2016.

GURNANI B, KAUR K. **PhacoemulsificationIn**. [s.l.] StatPearls Publishing, 2023.

HAHN, U. et al. Determination of Valid Benchmarks for Outcome Indicators in Cataract Surgery A Multicenter, Prospective Cohort Trial. **Ophthalmology**, v. 118, p. 2105–12, 18 ago. 2011.

HAIGIS, W. et al. Comparison of immersion ultrasound biometry and partial coherence interferometry for intraocular lens calculation according to Haigis. **Graefe's Archive for Clinical and Experimental Ophthalmology**, v. 238, n. 9, p. 765–773, 1 set. 2000.

HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. 2. ed. Singapore: Pearson Education, 1999.

HIPÓLITO-FERNANDES, D. et al. Intraocular lens power calculation formulas accuracy in combined phacovitrectomy: an 8-formulas comparison study. **International Journal of Retina and Vitreous**, v. 7, n. 1, p. 47, 18 ago. 2021.

HOFFER, K. et al. Protocols for Studies of Intraocular Lens Formula Accuracy. **American journal of ophthalmology**, v. 160, 24 jun. 2015.

HOFFER, K. J. Mathematics and computers in intraocular lens calculation. **American Intra-Ocular Implant Society Newsletter**, v. 1, n. 1, p. 4–5, maio 1975.

HOFFER, K. J. The Hoffer Q formula: A comparison of theoretic and regression formulas. **Journal of Cataract & Refractive Surgery**, v. 19, n. 6, 1993.

HOLLADAY, J. T. et al. A three-part system for refining intraocular lens power calculations. **Journal of Cataract & Refractive Surgery**, v. 14, n. 1, 1988.

IDROBO-ROBALINO, C. A.; SANTAELLA, G.; GUTIÉRREZ, Á. M. T2 formula in a highly myopic population, comparison with other methods and description of an improved approach for estimating corneal height. **BMC Ophthalmology**, v. 19, n. 1, p. 222, 12 nov. 2019.

JI, J. et al. Comparison of six methods for the intraocular lens power calculation in high myopic eyes. **European Journal of Ophthalmology**, v. 31, n. 1, p. 96–102, 2021.

KANE, J. X. et al. Intraocular lens power formula accuracy: Comparison of 7 formulas. **Journal of Cataract & Refractive Surgery**, v. 42, n. 10, 2016.

KANE, J. X. et al. Accuracy of 3 new methods for intraocular lens power selection. **Journal of Cataract & Refractive Surgery**, v. 43, n. 3, 2017.

KRAFF, M. C.; SANDERS, D. R.; LIEBERMAN, H. L. Biometric analysis of intraocular lens power required to produce emmetropia: results of 450 implants. **American Intra-Ocular Implant Society Journal**, v. 4, n. 1, p. 45–47, 1978.

KRAMER, O. K-Nearest Neighbors. Em: **Dimensionality Reduction with Unsupervised Nearest Neighbors**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. p. 13–23.

KUGELBERG, M.; LUNDSTRÖM, M. Factors related to the degree of success in achieving target refraction in cataract surgery: Swedish National Cataract Register study. **Journal of Cataract & Refractive Surgery**, v. 34, n. 11, p. 1935–1939, 2008.

LAM, D. et al. Cataract. **Nature Reviews Disease Primers**, v. 1, n. 1, p. 15014, 11 jun. 2015.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 1 maio 2015.

LEE, A. C.; QAZI, M. A.; PEPOSE, J. S. Biometry and intraocular lens power calculation. **Current Opinion in Ophthalmology**, v. 19, n. 1, 2008.

LI, H. et al. Comparing the accuracy of the new-generation intraocular lens power calculation formulae in axial myopic eyes: a meta-analysis. **International Ophthalmology**, v. 43, n. 2, p. 619–633, 5 set. 2022.

MAHESH, B. **Machine Learning Algorithms -A Review**. [s.l.: s.n.].

MEDINA, K. A. M. **Evaluación visual posterior a la cirugía de catarata en el Hospital Central FAP, 2018**. Lima: Universidad Nacional Federico Villareal, 2019.

MELLES, R. B.; HOLLADAY, J. T.; CHANG, W. J. Accuracy of Intraocular Lens Calculation Formulas. **Ophthalmology**, v. 125, n. 2, p. 169–178, fev. 2018.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997.

MONTEIRO, D. W. DE L.; TRINDADE, B. L. C. (EDS.). **Advances in Ophthalmic Optics Technology**. [s.l.] IOP Publishing, 2022.

MUSTAFA, M. Z. et al. Accuracy of biometric formulae in hypermetropic patients undergoing cataract surgery. **European Journal of Ophthalmology**, v. 29, n. 5, p. 510–515, 2019.

NOBLE, W. S. What is a support vector machine? **Nature Biotechnology**, v. 24, n. 12, p. 1565–1567, 1 dez. 2006.

NODALES, E. M. B. et al. Defectos esféricos residuales: diferencias entre el valor dióptrico del lente antes y después de la operación de catarata. v. 18, n. 1, 2017.

OLSEN, T. Prediction of the effective postoperative (intraocular lens) anterior chamber depth. **Journal of Cataract and Refractive Surgery**, v. 32, n. 3, p. 419–424, mar. 2006.

OLSEN, T.; CORYDON, L.; GIMBEL, H. Intraocular lens power calculation with an improved anterior chamber depth prediction algorithm. **Journal of Cataract and Refractive Surgery**, v. 21, n. 3, p. 313–319, maio 1995.

REICH, Y. Machine Learning Techniques for Civil Engineering Problems. **Microcomputers in Civil Engineering**, v. 12, 10 dez. 1996.

RETZLAFF, J. A new intraocular lens calculation formula. **American Intra-Ocular Implant Society Journal**, v. 6, n. 2, p. 148–152, abr. 1980.

RETZLAFF, J. A.; SANDERS, D. R.; KRAFF, M. C. Development of the SRK/T intraocular lens implant power calculation formula. **Journal of Cataract & Refractive Surgery**, v. 16, n. 3, 1990.

RODRIGUES, F. W. et al. Ocular biometry, mathematical estimation and spherical variation after facetectomy. v. 74, n. 6, p. 350–354, 1 dez. 2015.

RUÍZ, C. et al. **Variación del estado Refractivo y Agudeza Visual, en pacientes sometidos a Cirugía de Cataratas por las técnicas de Facoemulsificación y Extracción Extracapsular con implante de lente Intraocular Esféricas, atendidos en la Clínica Ojo Sano de la Ciudad de Managua, de Enero del 2015 a Junio del 2016**. Managua: Universidad Nacional Autónoma de Nicaragua, 2017.

RUSSEL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3. ed. [s.l.] Prentice-Hall, 2013.

SANDERS, D.; KRAFF, M. Improvement of intraocular lens power calculation using empirical data. **Journal - American Intra-Ocular Implant Society**, v. 6, n. 3, p. 263—267, jul. 1980.

SANDERS, D. R.; RETZLAFF, J.; KRAFF, M. C. Comparison of the SRK II<sup>TM</sup> formula and other second generation formulas. **Journal of Cataract and Refractive Surgery**, v. 14, n. 2, p. 136–141, mar. 1988.

SAVINI, G.; TARONI, L.; HOFFER, K. J. Recent developments in intraocular lens power calculation methods—update 2020. **Annals of Translational Medicine**, v. 8, n. 22, p. 1553–1553, nov. 2020.

SHRIVASTAVA, A. K. et al. Precision of intraocular lens power prediction in eyes shorter than 22 mm: An analysis of 6 formulas. **Journal of Cataract & Refractive Surgery**, v. 44, n. 11, 2018.

SOMVANSHI, M.; CHAVAN, P. A review of machine learning techniques using decision tree and support vector machine. **2016 International Conference on Computing Communication Control and automation (ICCUBEA)**, p. 1–7, 2016.

SU, J.; ZHANG, H. **A Fast Decision Tree Learning Algorithm**. Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1. **Anais...: AAAI'06**. AAAI Press, 2006.

SU, X.; YAN, X.; TSAI, C.-L. Linear Regression. **WIRES Comput. Stat.**, v. 4, n. 3, p. 275–294, maio 2012.

SUTHAHARAN, S. Decision Tree Learning. Em: **Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning**. Boston, MA: Springer US, 2016. p. 237–269.

WANG, L.; HOLLADAY, J. T.; KOCH, D. D. Wang-Koch axial length adjustment for the Holladay 2 formula in long eyes. **Journal of Cataract & Refractive Surgery**, v. 44, n. 10, 2018.

WU, Y.; FENG, J. Development and Application of Artificial Neural Network. **Wireless Personal Communications**, v. 102, n. 2, p. 1645–1656, 1 set. 2018.

## APÊNDICE

Tabela 5. Valores das variáveis utilizadas em cada algoritmo da biblioteca scikit-learn

| Algoritmo | Variável           | Valores  |
|-----------|--------------------|--|
| LR        |                    |  |
| Ridge     | Alpha              | 0, 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 20, 50, 100, 500, 1000 |
|           | fit_intercept      | True, False  |
|           | Solver             | 'svd', 'cholesky', 'lsqr', 'sparse_cg', 'sag', 'saga'  |
| Lasso     | Alpha              | 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 20, 50, 100, 500, 1000,   |
|           | fit_intercept      | True, False  |
| kNN       | K                  | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30  |
| SVR       | C                  | 0.1, 1, 10, 100, 1000, 10000, 100000, 200000   |
|           | Kernel             | 'linear', 'rbf', 'poly', 'sigmoid'   |
|           | Gamma              | 1, 0.1, 0.01, 0.001,   |
| MLP       | Solver             | 'lbfgs', 'sgd', 'adam'   |
|           | activation         | 'logistic', 'tanh', 'relu'   |
|           | Alpha              | 0.01, 0.001, 0.0001, 0.00001   |
|           | hidden_layer_sizes | (3,), (5,), (7,), (10,)  |
|           | max_iter           | 500, 1000, 1500, 3000  |
| GBR       | learning_rate      | 0.001, 0.01, 0.05  |
|           | subsample          | 0.9, 0.5, 0.2, 0.1   |
|           | n_estimators       | 500, 1000, 1500  |
|           | max_depth          | 4, 6, 8  |

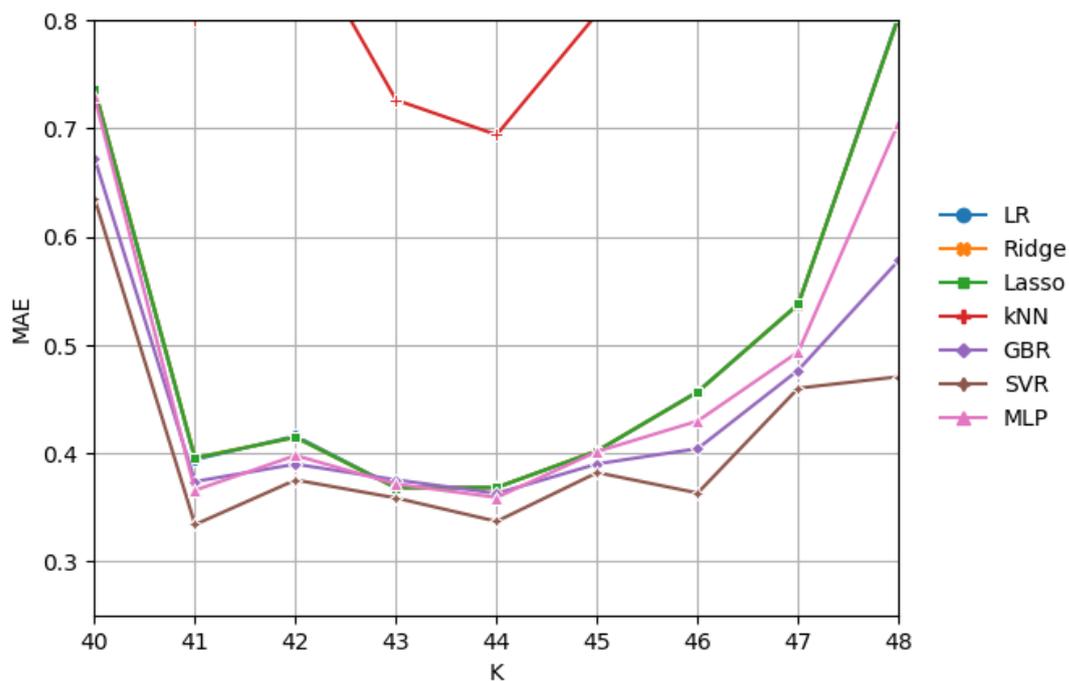


Figura 6. Modelos preditivos: ceratometria central versus erro absoluto médio

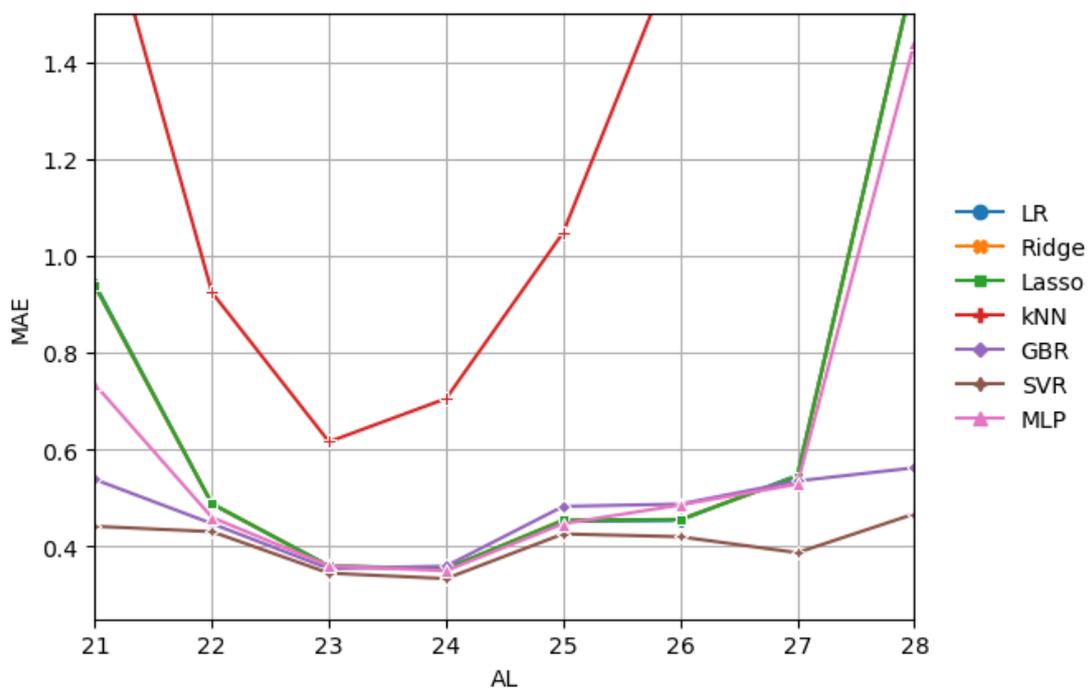


Figura 7. Modelos preditivos: comprimento axial versus erro absoluto médio

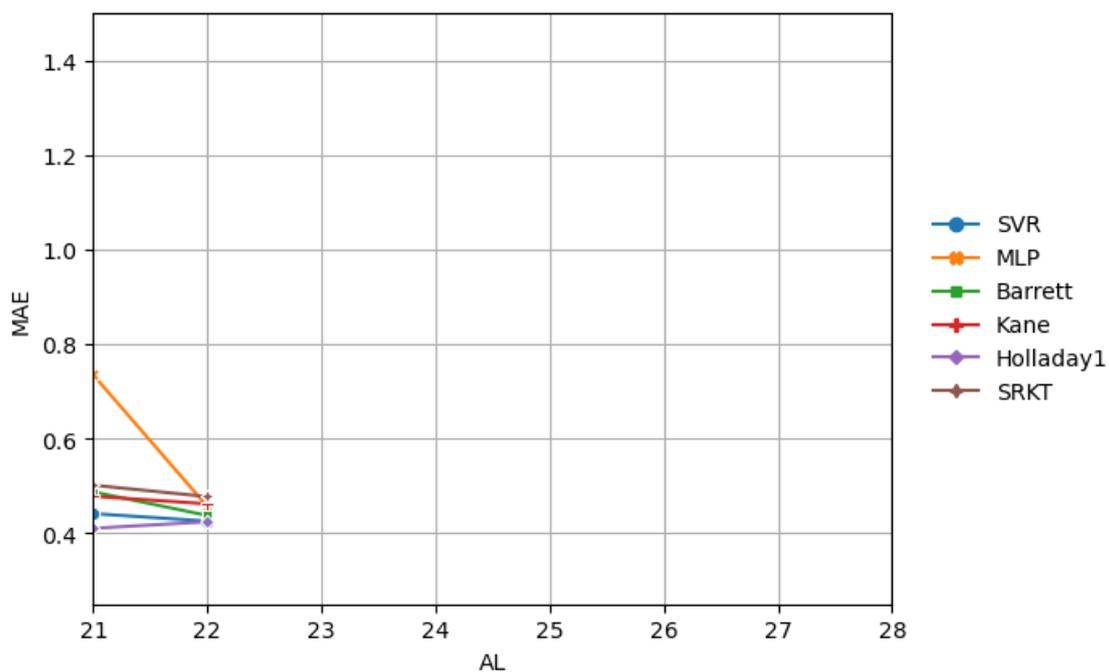


Figura 8. Fórmulas e modelos preditivos: comprimento axial versus erro absoluto médio em olhos curtos

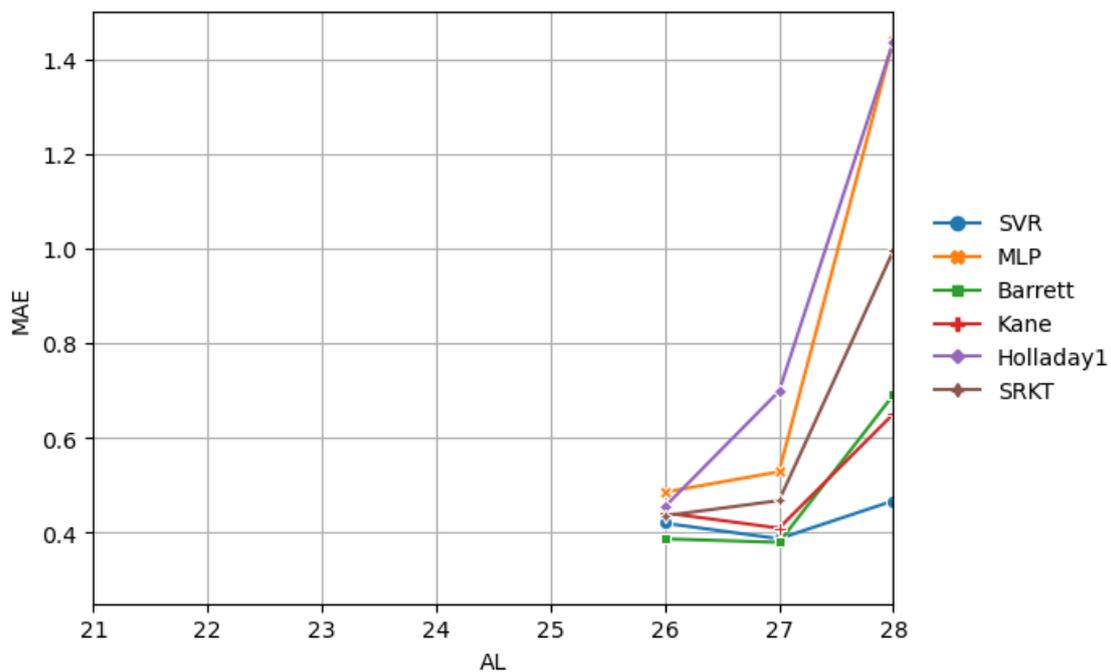


Figura 9. Fórmulas e modelos preditivos: comprimento axial versus erro absoluto médio em olhos longos

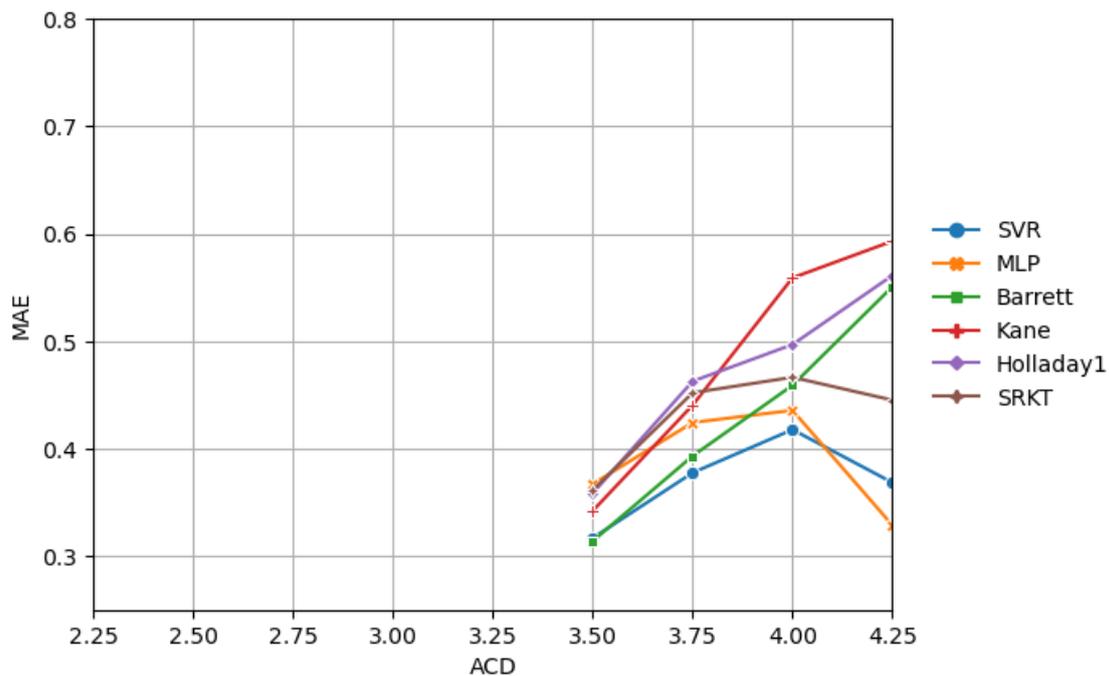


Figura 10. Fórmulas e modelos preditivos: profundidade da câmara anterior versus erro absoluto médio em olhos com câmara anterior profunda

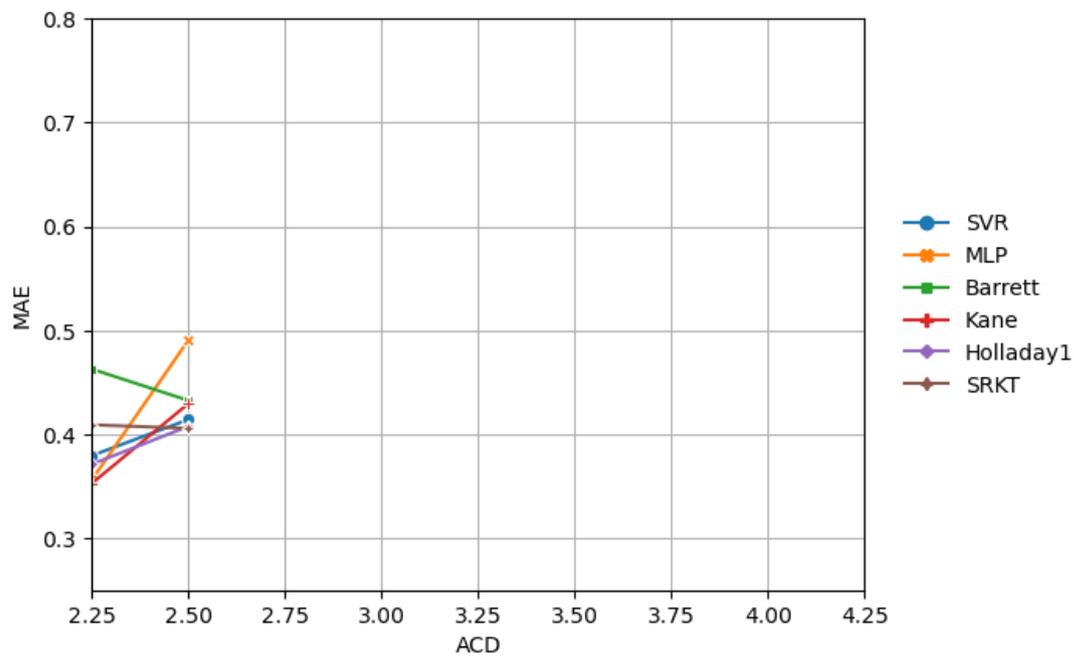


Figura 11. Fórmulas e modelos preditivos: profundidade da câmara anterior versus erro absoluto médio em olhos com câmara anterior rasa

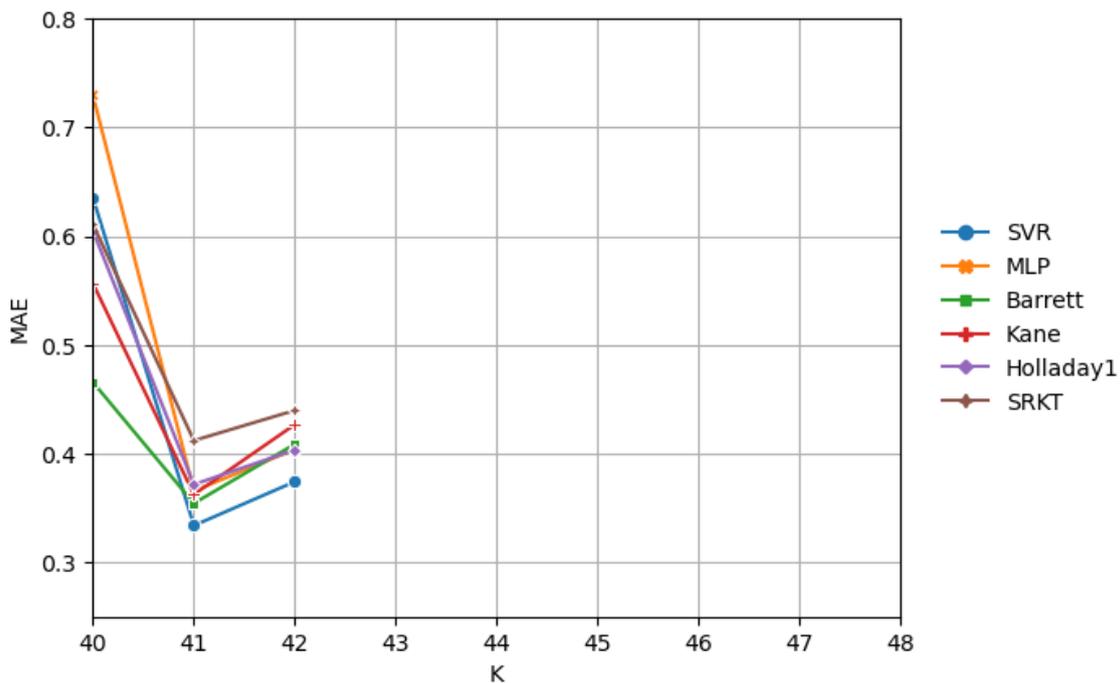


Figura 12. Fórmulas e modelos preditivos: ceratometria central versus erro absoluto médio em olhos com córnea plana

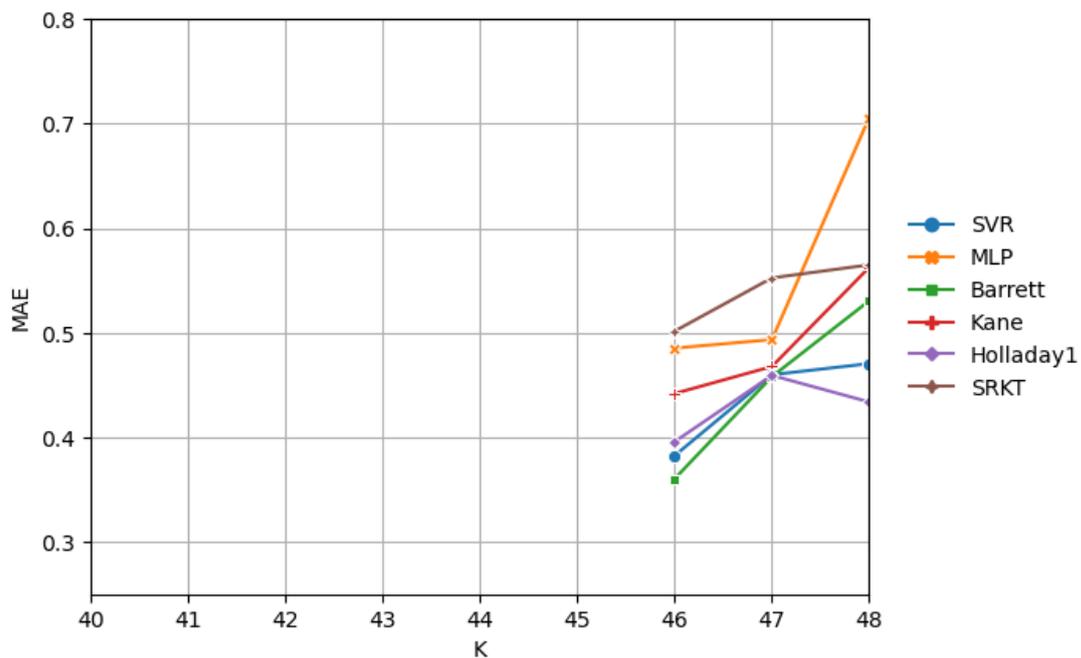


Figura 13. Fórmulas e modelos preditivos: ceratometria central versus erro absoluto médio em olhos com córnea curva