

CLASSIFICAÇÃO DE INDIVÍDUOS ANSIOSOS POR ÁRVORE DE DECISÃO

Ribeiro, Andréa Vieira¹; Junior, Homero da Silva Nahum^{2,3};

Barreto, Ana Cristina Lopes y Glória²; Brasil, Roxana Macedo²

64

Resumo

Objetivando estimar dois modelos de classificação com 62 indivíduos ansiosos, aplicou-se o algoritmo CART para árvore de decisão, visando identificar o Problema Clínico (Ansiedade e Ansiedade + Depressão) e Estilo de Vida (Sedentário e Ativo). Utilizou-se o RStudio 2022.07.2+576 for Windows® para codificação e a aprendizagem de máquina para reduzir a interferência humana, separando em grupos de teste e treino, contendo 70,00% e 30,00% dos voluntários respectivamente. O modelo para predição do Problema Clínico apresentou-se adequadamente com acurácia de 90,70%, enquanto aquele para predizer a Condição de Sedentário deteve ajuste perfeito, acurácia de 100,00%. Então, concluiu-se que ambos os modelos conseguiram classificações adequadas à realidade apresentada.

Palavras-chave: Aprendizagem de máquina; Modelo supervisionado; Inteligência artificial; Ciências de dados; Mineração de dados.

Abstract

Aiming to estimate two classification models with 62 anxious individuals, the CART algorithm was applied to the decision tree, aiming to identify the Clinical Problem (Anxiety and Anxiety + Depression) and Lifestyle (Sedentary and Active). RStudio 2022.07.2+576 for Windows® was used for coding and machine learning to reduce human interference, separating into test and training groups, containing 70.00% and 30.00% of volunteers respectively. The model for predicting the Clinical Problem presented itself adequately with an accuracy of 90.70%, while the one for predicting the Sedentary Condition had a perfect fit, with an accuracy of 100.00%. Therefore, it was concluded that both models achieved classifications appropriate to the reality presented.

Keywords: Machine learning; Supervised model; Artificial intelligence; Data science; Data mining.

Introdução

Algoritmos de classificação seriam procedimentos de aprendizagem supervisionada, cujo objetivo seria predizer uma classe, por exemplo Ativo ou Sedentário. Diferenciando-os dos modelos de regressão, por esses estimarem determinado valor absoluto ou probabilidade (Silva e Silva Neto, 2022). A classificação poderia ser binária, quando somente duas classes seriam possíveis, ou multiclasse, na existência de mais possibilidades (Breiman *et al.*, 2017). A árvore de decisão classificaria a variável

¹ Graduanda em Educação Física no Centro Universitário Celso Lisboa – RJ/Brasil

² Docente do Curso de Educação Física do Centro Universitário Celso Lisboa – RJ/Brasil

³ Docente da Escola de Saúde da Universidade Candido Mendes – RJ/Brasil

dependente (instância) com base nos valores das independentes (atributos), obedecendo o caminho definido da raiz às folhas (Breiman, 2001).

Os critérios de avaliação seriam métricas de variabilidade e informação. Dentre elas, a Entropia estimaria a homogeneidade dos dados, portanto quanto maior, mais distintos (impuros) seriam (Silva e Silva Neto, 2022). O cálculo considerou que os modelos deveriam classificar os indivíduos em dois grupos ($i = 2$), os quais teriam respectivas quantidades de pessoas, cuja proporção em relação à totalidade dos dados seria p_i , logo a Entropia foi tomada por (Silva, 2005):

$$Entropia = \sum_{i=1}^2 -p_i \cdot \log_2^{p_i}$$

Então, quanto maior a Entropia, menor seria o Ganho de Informações, porque a Entropia = 0, significaria que os dados seriam puros, então o Ganho seria máximo. Em última análise, estimaria o quanto que a variável dependente poderia ser explicada a partir da independente (Simonetto, 2022). O cálculo seria a diferença entre as entropias do nó pai e a ponderação daquelas dos nós filhos (Silva, 2005):

$$Ganho = Entropia(pai) - \sum \frac{Ocorrências\ do\ Nó\ Filho}{Total\ de\ Dados} \cdot Entropia(nó\ filho)$$

O Índice Gini estimaria a impureza, portanto o valor zero indicaria máxima pureza (Ariza *et al.*, 2022), portanto Silva (2005) expos que o emprego dessa métrica isolaria o grupo mais frequente em um determinado ramo, enquanto que a Entropia distribuiria os grupos em cada ramo. Para tanto, o cálculo seria:

$$Gini = 1 - \sum_{i=1}^2 p_i^2$$

Na classificação binária possível seria avaliar o desempenho dos algoritmos a partir do acerto das predições, criando a matriz de confusão (Tonidandel, Silva e Brun, 2023) com os valores de: Verdadeiros Positivos (VP), quantidade de predições corretas da classe 1; Verdadeiros Negativos (VN), número de predições corretas da classe 0; Falsos Positivos (FP), quantidade de predições erradas da classe 1; e Verdadeiros Negativos (VN), número de predições erradas da classe 0. Com esses valores, possível seria estimar a acurácia que refletiria o percentual de predições corretas.

Nas áreas de saúde e Educação Física, as aplicações de árvore de decisão com aprendizagem de máquina foram realizadas para predição de resultados no futebol (Silva, Barros e Albuquerque, 2020), predição de comprometimentos na saúde psíquica (Li, 2023),

avaliação alimentar com comprometimento cardiovascular (Silva *et al.* 2022) e para predizer risco cardíaco e nível de atividade física em adolescente (Soares Júnior, 2023). Então, o corrente estudo objetivou classificar indivíduos ansiosos com relação ao Problema Clínico e à Condição de Sedentário.

Materiais e Métodos

O algoritmo foi codificado em RStudio 2022.07.2+576 "Spotted Wakerobin" Release (e7373ef832b49b2a9b88162cfe7eac5f22c40b34, 2022-09-06) para Windows®, com a utilização dos pacotes rpart 4.1.19, rpart.plot 3.1.1, coin 1.4-2 e readxl 1.4.3, e aplicado para classificar 62 indivíduos (16 diagnosticados com Ansiedade e Depressão, e 46 pessoas, somente com Ansiedade), caracterizados por 14 variáveis (Sexo, Idade, Escolaridade, Estado Civil, Ocupação, Renda Mensal, Número de Moradores, Tempo de Diagnóstico de Ansiedade, Sintomas de Ansiedade, Sintomas de Depressão, Condição de Sedentário, Tempo de Prática, Total de Pontos de Sofrimento Psíquico e Problema Clínico) e organizados em banco de dados.

Para reduzir a interferência humana na modelagem foi adotado o aprendizado de máquina (*machine learning*), o qual consistiria na automação da análise de dados e desenvolvimento de modelos analíticos (Mezávila, Dias e Franco, 2021; Carmo e Silva, 2023; Lisboa, 2023). Essencialmente, pelos dados os sistemas seriam capazes de tomar decisões ou reconhecer padrões, portanto ramo da inteligência artificial seria (Silva L *et al.*, 2023; Giraffa e Khols-Santos, 2023; Barbosa, 2023).

A aprendizagem de máquina exigiu a divisão automática desse em grupos de treino e teste, contendo respectivamente 70,00% e 30,00% do total de pessoas (Benevento e Meirelles, 2023; Andrade *et al.*, 2023; Barros e Freitas Junior, 2023). Procedimento adotado nos modelos para predição do Problema Clínico e Condição de Sedentário. O algoritmo utilizado foi o *Classification And Regression Trees* (CART – Árvores de Classificação e Regressão) apresentado por Breiman em 1984 (Simonetto, 2022; Silva, 2005), pois desenvolvido foi para gerar predições com base em valores conhecidos, no caso corrente a ideia era decidir se determinado indivíduo pertencia a um dentre dois grupos (Gouveia *et al.*, 2015; Dias e Dias, 2023), o que caracterizou a aplicação como árvore de classificação (Ariza *et al.*, 2023; Costa Júnior *et al.*, 2023). Os critérios de avaliação adotados foram a Entropia, o Índice Gini e Ganho de Conhecimento.

Discussão

O modelo de classificação do Problema Clínico (Ansiedade ou Ansiedade + Depressão) revelou que das 13 variáveis independentes, Sexo não conquistou efetiva capacidade preditiva (Tabela 1) e dessa, aproximadamente, 58,00%, se concentraram em Tempo de Prática, Sintomas de Depressão e Condição de Sedentário. O destaque se deu na baixa Importância de Sintomas de Ansiedade e Total de Pontos de Sofrimento Psíquico com 5,00% e 3,00%, respectivamente.

Então, ao grupo investigado àquelas variáveis de saúde e as demográficas não influenciariam a classificação, ficando essa dependente das três destacadas, anteriormente. Somente isso, suficiente seria à consideração multivariada da caracterização do transtorno de ansiedade (Silva, Sousa e Martuscello Neto, 2023; Da Fonseca, 2023), por conseguinte, a intervenção deveria abranger múltiplas variáveis (Leitão e Moura, 2023; Braitenbach de Oliveira e Trevisol, 2023; Santos, Almeida-Souza e Góes, 2023).

Tabela 1: Importância das Variáveis, y = Problema Clínico

| Variável | Importância | Variável | Importância |
|-----------------------------------|-------------|--|-------------|
| Tempo de Prática | 33,00 | Sintomas de Ansiedade | 5,00 |
| Sintomas de Depressão | 14,00 | Ocupação | 4,00 |
| Condição de Sedentário | 11,00 | Total de Pontos de Sofrimento Psíquico | 3,00 |
| Idade | 9,00 | Estado Civil | 2,00 |
| Tempo de Diagnóstico de Ansiedade | 9,00 | Número de Moradores | 1,00 |
| Escolaridade | 8,00 | Renda Mensal | 1,00 |

Fonte: Os autores (2024)

Ratificou tal constatação, o fato do Treino ter conseguido classificar corretamente 28 pessoas (de 30 indivíduos) com quadro de Ansiedade, o que representou 93,33% de acerto (Tabela 2), enquanto que dentre aqueles acometidos por Ambos transtornos o modelo identificou como tal, cerca de, 84,62% (11 pessoas) das ocorrências. À vista disso, possível foi inferir o desempenho geral do modelo, ou seja, estimar a proximidade ao Problema Clínico. Isso seria, essencialmente, a acurácia, pois indicaria a aproximação ao valor real (padrão, valor de referência ou tendência), ou, simplesmente, quantas classificações o modelo acertou em relação a todas realizadas, o que totalizou:

$$Acurácia\ Treino = \frac{11 + 28}{43} \cong 90,70\%$$

Marques *et al.* (2022) objetivaram avaliar quatro medidas lineares cranianas em 200 tomografias computadorizadas para classificar os indivíduos em razão da dicotomia sexual, identificando que metade delas detinha importância relevante, com isso, conquistando acurácia de 66,80%. A estimativa inferior pode ter ocorrido pela influência étnica, idade ou causa mortis.

Também foi inferior a acurácia (59,00%) do quadro de delírio, em Unidades de Terapia Intensiva, caracterizado por confusão psíquica associada ao risco de morbidade e mortalidade, obtida por Coombes, Coombes e Fareed (2021), talvez o aspecto dificultador residisse na confusão psíquica. Tanto que estudos com variável dependente objetiva tenderiam a apresentar elevada acurácia, conforme encontrado por Hashi, Zaman e Hasan (2017) para predição do diagnóstico de doenças, quando encontraram 90,43%, convergindo ao corrente estudo.

O desempenho no Teste (Tabela 2) foi adequado, pois somente houve dois erros (12,50%), quando indivíduos com Ansiedade foram classificados como Ambos, portanto os identificados como pertencentes ao grupo Ansiedade, realmente, o eram. Contudo, possivelmente, devido ao baixo número de pessoas (19 indivíduos), a acurácia sofreu atenuação de 1,23%:

$$Acurácia\ Teste = \frac{3 + 14}{19} \cong 89,47\%$$

Tabela 2: Comparação Predito x Observado, y = Problema Clínico

| | | Observado | | |
|---------------|-----------|--------------------|--------------------|-------|
| | | Ambos | Ansiedade | Total |
| Treino | | n; %Linha; %Coluna | n; %Linha; %Coluna | |
| | Ambos | 11; 84,62; 84,62 | 2; 15,38; 6,67 | 13 |
| | Ansiedade | 2; 6,67; 15,38 | 28; 93,33; 93,33 | 30 |
| | Total | 13 | 30 | 43 |
| Teste | | | | |
| Predito | Ambos | 3; 100,00; 60,00 | | 3 |
| | Ansiedade | 2; 12,50; 40,00 | 14; 87,50; 100,00 | 16 |
| | Total | 5 | 14 | 19 |

Fonte: Os autores (2024)

A melhor avaliação do modelo foi da impureza pelo Índice Gini = 0,46, resultado mediano, pois a Entropia = 0,85 significaria alto índice de impureza, o que poderia comprometer o Ganho de Informação (1,08), portanto, a princípio, a predição do Problema

Clínico pelas variáveis independentes não seria perfeita, mas poderia ser adequada (Simonetto, 2022), tal como demonstrado pela Tabela 2. Esses resultados convergiriam às expectativas da modelagem, especialmente na ciência de que o resultado dessa exporia uma possibilidade de representação do fenômeno estudado (Dutra, Firmino Júnior e Fernandes, 2023; Moretti *et al.*, 2023).

Modelos, em geral, proporcionariam melhores predições, à medida que variáveis intervenientes (aquelas que influenciariam a predição, mas não poderiam ser aferidas) fossem inexistentes. Flores *et al.* (2023) avaliaram a mortalidade de 318 idosos (Idade ≥ 65 anos) submetidos a medicamentos potencialmente inadequados. Os autores identificaram que 59 pessoas (37,30%) fizeram óbito, demonstrando que a prescrição anterior à internação seria um fator de risco, influenciado pela intervenção de variáveis dos domínios clínicos e funcionais.

Ratificando essas considerações, Nangino (2023) analisou o impacto da implantação de serviço de telessaúde sobre a taxa de reincidência de internação hospitalar em 30 dias para pacientes adultos de Programas de Atenção Domiciliar, não tendo constatado influência positiva, durante a pandemia de Covid-19. Talvez, em virtude de variáveis intervenientes como o comportamento das pessoas, resistindo em ir aos hospitais, associado à redução de acesso àquelas instituições; e desconhecimento da classificação de funcionalidade, condição socioeconômica e determinantes sociais de saúde dos pacientes.

No presente estudo, os Sintomas de Depressão poderiam ser intervenientes, pois não seriam passíveis de aferição objetiva, tanto que se caracterizariam por declaração própria, dada a dificuldade de aferição da intensidade deles, como no caso de ideação suicida (Freitas *et al.*, 2023), concentração (Silva R *et al.*, 2022) e autoestima (Lopes, Silva e Alves, 2023; Furtado *et al.*, 2023). Então, aquela variável recebeu a quantidade de sintomas.

A classificação dos voluntários pela Condição de Sedentário (Ativo ou Sedentário) homologou a relatividade das variáveis, pois, à primeira luz, o Tempo de Prática seria suficiente à separação dos indivíduos (Importância = 100,00%), porque os sedentários o teriam como nulo (Tabela 3). Contudo, isso não se concretizou, que pese, aquela variável ter conquistado 65,00% de Importância, indicando que os grupos teriam características sobrepostas, logo a menção multivariada foi corroborada.

No contexto da modelagem, a Condição Sedentário como variável independente demandou um modelo com menor número de variáveis, especialmente dada a ausência de influência das variáveis Sexo, Estado Civil, Ocupação, Tempo de Diagnóstico de Ansiedade, Sintomas de Ansiedade, Sintomas de Depressão, Total de Pontos de Sofrimento Psíquico e Problema Clínico.

Tabela 3: Importância das Variáveis, y = Condição de Sedentário

| Variável | Importância |
|---------------------|-------------|
| Tempo de Prática | 65,00 |
| Escolaridade | 12,00 |
| Idade | 12,00 |
| Número de Moradores | 6,00 |
| Renda Mensal | 5,00 |

Fonte: Os autores (2024)

Para além de tal consideração, conjuntamente Tempo de Prática, Escolaridade e Idade responderam por 89,00% da Importância do conjunto de 13 variáveis, potencialmente, explicativas. Possível foi, então, conjecturar que o estilo de vida, sedentarismo ou não, daqueles indivíduos dependeria, com alguma intensidade, da maturidade (representada na Idade) e educação formal (na Escolaridade). Essa poderia ser entendida como um princípio social (Pérez, 2023), dada a explícita contextualização demográfica.

Alguns princípios do treinamento desportivo, aparentemente, se apresentaram nesse modelo, pois a Idade lança foco sobre a relevância da individualidade biológica (Barbanti, Tricoli e Ugrinowitsch, 2004; Da Silva *et al.*, 2023). Lugar comum seria destacar que as idades cronológica e biológica, e a capacidade de desempenho, por exemplo seriam fatores intrínsecos e individuais impactantes na prática regular de exercícios, dada a maturidade mecânica e, particularmente, fisiológica (Lapin *et al.*, 2007). Porém, a percepção e colheitas dos benefícios inerentes exigiriam a continuidade, por isso o Tempo de Prática seria relevante. Essencialmente, os estímulos do treinamento originariam novas conexões neuronais no domínio do tempo, desde que houvesse adequação da 1) sobrecarga, a qual deveria ser crescente, 2) relação volume e intensidade, com comportamentos inversos; 3) variabilidade, os estímulos deveriam ser variados; e 4) adaptação, o organismo sempre buscaria a homeostase (Lussac, 2008; Roschel, Tricoli e Ugrinowitsch, 2011; Lima, Dos Reis Júnior e Bandeira, 2020; Medeiros, Silva e Machado Júnior, 2022; Vilela Júnior *et al.*, 2023; Cortez *et al.*, 2023), dentre outras.

Os desempenhos de Treino e Teste foram análogos no condizente à correta classificação dos indivíduos (Tabela 4), assim com as acurácias (100,00%). Na observação dos resultados de Treino possível seria supor a existência de sobreajuste (overfitting), pois o modelo se adaptou perfeitamente aqueles dados. Entretanto, ele foi capaz de identificar os padrões de classificação no Teste, portanto eliminando a suposição anterior (Silva e Silva Neto, 2022; Kanda e Carvalho, 2023; Lopes, Colombi e Mutz, 2023; Silva L *et al.*, 2023). Logo, a aprendizagem de máquina se substancializou adequadamente à predição proposta, o que foi confirmado pelas estimativas nulas de impureza, tanto pelo Índice Gini quanto pela Entropia, culminado no máximo Ganho de Informação. Explicitamente, significando que as variáveis independentes possibilitaram a exata predição da Condição de Sedentário.

Tabela 4: Comparação Predito x Observado, y = Condição de Sedentário

| | | Observado | | Total |
|---------------|--------------------|--------------------|--------------------|-------|
| | | Ativo | Sedentários | |
| Treino | | n; %Linha; %Coluna | n; %Linha; %Coluna | |
| | Predito Ativo | 32; 100,00; 100,00 | | 32 |
| | Predito Sedentário | | 11; 100,00; 100,00 | 11 |
| | Total | 32 | 11 | 43 |
| Teste | | n; %Linha; %Coluna | n; %Linha; %Coluna | |
| | Predito Ativo | 15; 100,00; 100,00 | | 15 |
| | Predito Sedentário | | 4; 100,00; 100,00 | 4 |
| | Total | 15 | 4 | 19 |

Fonte: Os autores (2024)

Considerações Finais

Objetivando estimar dois modelos de classificação com 62 indivíduos ansiosos, aplicou-se o algoritmo CART para árvore de decisão. O modelo para predição do Problema Clínico apresentou-se adequadamente com acurácia de 90,70%, enquanto que aquele para predizer a Condição de Sedentário deteve ajuste perfeito, acurácia de 100,00%. Então, concluiu-se que ambos modelos conseguiram classificações adequadas à realidade apresentada.

Aos futuros estudos, recomenda-se a comparação com modelos de floresta aleatória (randon forest) e máquina de vetores de suporte (support vector machine) pelas estimativas de acurácia. A investigação sistematizada da redução de dimensionalidade (quantidade de variáveis) pode fornecer orientação à ampliação do banco de dados, possibilitando a reavaliação dos modelos. As variáveis independentes de cada modelo podem ser utilizadas

na estimação de regressão logística, visando determinar a probabilidade de desenvolver Ansiedade e Ansiedade + Depressão, e, acometido pela ansiedade, a chance de adotar os estilos de vida Ativo e Sedentário.

Referências

ANDRADE, MS *et al.* Análise de métricas de desempenho no ambiente de intrusões em Redes IEEE 802.11 com aprendizado de máquina no Hospital NSC. **Research, Society and Development**, v. 4, p e22512441277, 2023.

ARIZA, VMP *et al.* Uso do algoritmo "Floresta Aleatória" na identificação do comportamento da população na busca por serviços de saúde após o início da pandemia do novo coronavírus. **AtoZ: novas práticas em informação e conhecimento**, v. 11, p. 1-15, 2022.

BARBANTI, VJ; TRICOLI, V; UGRINOWITSCH, C. Relevância do conhecimento científico na prática do treinamento físico. **Revista Paulista de Educação Física**, v. 18, n. 8, p. 101-109, 2004.

BARBOSA, CRAC. Transformações no ensino-aprendizagem com o uso da inteligência artificial: revisão sistemática da literatura. **RECIMA21 - Revista Científica Multidisciplinar**, v. 4, n. 5, e453103, 2023.

BARROS, PHB; FREITAS JÚNIOR, AM. Combinando inteligência artificial e imagens de satélite para a previsão de sinistros agrícolas: Uma nota. **Revista Brasileira de Economia**, v. 77, e012023, 2023.

BENEVENTO, M; MEIRELLES, FS. Prever e melhorar o desempenho dos alunos com o uso combinado de aprendizagem de máquina e GPT. **Revista de Gestão e Avaliação Educacional**, e74348, p. 1–22, 2023.

BRAITENBACH DE OLIVEIRA, R; TREVISOL, M. Transtorno de ansiedade (TA) em um viés psicanalítico. **Anuário Pesquisa E Extensão Unoesc São Miguel Do Oeste**, v. 8, e32904, 2023.

BREIMAN, L *et al.* **Classification and regression trees**. Abingdon: Routledge, 2017.

BREIMAN, L. Random forests. **Machine learning**, v. 45, n. 1, p. 5–32, 2001.

CARMO, CRS; SILVA, JRM. Aprendizado de máquina e prestação de serviços de armazenamento de dados: métricas para análise e validação de algoritmos previsores. **GETEC**, v. 12, n. 38, p. 123-144, 2023.

COOMBES, CE; COOMBES, KR; FAREED, N. A novel model to label delirium in an intensive care unit from clinician actions. **BMC Medical Informatics and Decision Making**, v. 21, n. 1, p. 1-12, 2021.

CORTEZ, ACL *et al.* Evidências científicas acerca dos efeitos do treinamento resistido, aeróbico e de flexibilidade e suas adaptações crônicas na saúde do idoso. **Retos**, n. 48, p. 978-987, 2023.

COSTA JUNIOR, S *et al.* Classificação da qualidade da biomassa de árvores de acácia-negra para fins energéticos. **Ciência Florestal**, v. 33, n. 2, e71436, 2023.

DA FONSECA, PAR. Transtorno de ansiedade: uma revisão bibliográfica sobre o mal do Século XXI. **Revista Contemporânea**, v. 3, n. 8, p. 12669–12677, 2023.

DA SILVA, TCF *et al.* Estrutura e viabilidade do treinamento físico através da consultoria online. **Revista Foco**, v. 16, n. 1, e756, 2023.

DIAS, JC; DIAS, JC. Análise preditiva de dados: Uma abordagem de conceitos e algoritmos. **Revista Processando o Saber**, v. 15, n. 01, p. 63-74, 2023.

DUTRA, JF; FIRMINO JÚNIOR, JB; FERNANDES, DYS. Fatores que podem interferir no desempenho de estudantes no ENEM: uma revisão sistemática da literatura. **Revista Brasileira de Informática na Educação**, v. 31, p. 323–351, 2023.

FLORES, TG *et al.* Sobrevida de pessoas idosas hospitalizadas com uso prévio de medicamentos potencialmente inapropriados. **Revista Brasileira de Geriatria e Gerontologia**, v. 26, e230017, 2023.

FREITAS, PHB *et al.* Symptoms of depression, anxiety and stress in health students and impact on quality of life. **Revista Latino-Americana de Enfermagem**, v. 31, e3885, 2023.

FURTADO, AR *et al.* Fatores associados ao nível de atividade física em adolescentes. **Texto & Contexto Enfermagem**, v. 32, e20220244, 2023

GIRAFFA, L; KHOLS-SANTOS, P. Inteligência artificial e educação: conceitos, aplicações e implicações no fazer docente. **Educação em Análise**, v. 8, n. 1, p. 116–134, 2023.

GOUVEIA, M *et al.* Avaliação econômica de ticagrelor em prevenção secundária pós Síndrome Coronária Aguda. **Revista Portuguesa de Cardiologia**, v. 34, n. 1, p. 17-25, 2015.

HASHI, EK; ZAMAN, MSU; HASAN, MR. An expert clinical decision support system to predict disease using classification techniques. **Anais International Conference on Electrical, Computer and Communication Engineering**. IEEE, 2017, p. 396-400.

KANDA, JY; CARVALHO, ACPLF. Previsão do preço do carbono por modelos de aprendizado de máquina. **Amazônia, Organizações e Sustentabilidade**, v. 12, n. 2, p. 158-176, 2023.

LAPIN, LP *et al.* Respostas metabólicas e hormonais ao treinamento físico. **Revista Brasileira de Educação Física, Esporte, Lazer e Dança**, v. 2, n. 4, p. 115-124, 2007.

LEITÃO, GJG; MOURA, LKS. Transtornos de ansiedade em estudantes de medicina no Brasil: uma revisão integrativa. **Brazilian Journal of Health Review**, v. 6, n. 3, p. 12011–12020, 2023.

LI, B. Design of early warning system for mental health problems based on data mining and database. **Revista Brasileira de Medicina do Esporte**, v. 29, e2022_0153, 2023.

LIMA, WP; DOS REIS JÚNIOR, J; BANDEIRA, JPB. Treinamento esportivo: um estudo introdutório sobre suas bases científicas. **Itinerarius Reflectionis**, v. 16, n. 3, p. 01-10, 2020.

LISBOA, GMJ. O consumidor de seguros na era do big data: desafios da atual regulamentação europeia frente às técnicas de definição de perfis, aprendizagem de máquina e decisões automatizadas. **Actualidad Jurídica Iberoamericana**, n. 18, p. 2134-2163, 2023.

LOPES, AS; SILVA, IVC; ALVES, PCBA. Psychology interventions with adolescents diagnosed with depression. **Research, Society and Development**, v. 12, n. 2, e7812239919, 2023.

LOPES, RS; COLOMBI, LR; MUTZ, F. Comparação de algoritmos de aprendizado de máquina para predição de pontuação de crédito. **Anais XIV Computer on the Beach**. Florianópolis (SC), p. 424-431, 2023.

LUSSAC, RMP. Os princípios do treinamento esportivo: conceitos, definições, possíveis aplicações e um possível novo olhar. **EFDeports**, v. 13, n. 121, 2008.

MARQUES, FL *et al.* Avaliação tomográfica de quatro medidas lineares cranianas para a estimativa do sexo. **Revista Brasileira de Odontologia Legal**, v. 9, n. 3, p. 25-34, 2022.

MEDEIROS, LF; SILVA, AC; MACHADO FILHO, R. A relação entre a sobrecarga de treinamento e hipertrofia muscular em praticantes de musculação. **Intercontinental Journal on Physical Education**, v. 4, n. 2, e2020033, 2022.

MEZÁVILA, SA; DIAS, NAA; FRANCO, ME. Aprendizagem de máquina aplicada à análise de batimentos cardíacos. **Revista Eixos Tech**, v. 8, n. 1, 2021.

MORETTI, E *et al.* Eficácia e custo-utilidade de intervenções para o tratamento da enurese em crianças e adolescentes sob a perspectiva do sistema único de saúde brasileiro. **Value in Health Regional Issues**, v. 37, p. 1-8, 2023.

NANGINO, GO. **A taxa de reinternação hospitalar após o uso da telessaúde em atenção domiciliar na pandemia de Covid-19**. Dissertação (Mestrado em Gestão para a Competitividade). Escola de Administração de Empresas de São Paulo. Fundação Getúlio Vargas. São Paulo, 2023.

PÉREZ, EVM. Fraternidad en la filosofía: de ausencias y rendimientos. **Veritas**, n. 55, p. 31-45, 2023.

ROSCHER, H; TRICOLI, V; UGRINOWITSCH, C. Treinamento físico: considerações práticas e científicas. **Revista Brasileira de Educação Física e Esporte**, v. 25, p. 53-65, 2011.

SANTOS, IM; ALMEIDA-SOUZA, TH; GOES, TC. Relação entre coping religioso/espiritual e níveis de ansiedade social em estudantes universitários com transtorno de ansiedade social. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 9, n. 6, p. 890–907, 2023.

SILVA, ABL; BARROS, KNNO; ALBUQUERQUE, MA. Modelagem via árvore de decisão para previsão de jogos de futebol. **Research, Society and Development**, v. 9, n. 9, e204996869, 2020.

SILVA, LL *et al.* dos. Análise das atividades interativas de um jogo sério usando aprendizagem de máquina: um estudo para pessoas com deficiência intelectual. **Revista Novas Tecnologias na Educação**, v. 21, n. 1, p. 160–169, 2023.

SILVA, LMO. **Uma aplicação de árvores de decisão, redes neurais e KNN para a identificação de modelos arma não-sazonais e sazonais**. Tese (Doutorado em Engenharia Elétrica). Programa de Pós-graduação em Engenharia Elétrica. Departamento de Engenharia Elétrica. Pontifícia Universidade Católica. Rio de Janeiro. 2005.

SILVA, PP *et al.* Application of data mining to extract knowledge about the occurrence of fistulas after palatoplasty. **Revista Brasileira de Cirurgia Plástica**, v. 38, n. 1, e0657, 2023.

SILVA, R; SILVA NETO, DR. Inteligência artificial e previsão de óbito por Covid-19 no Brasil: uma análise comparativa entre os algoritmos Logistic Regression, Decision Tree e Random Forest. **Saúde em Debate**, v. 46, n. (especial) 8, p. 118-129, 2022.

SILVA, RM *et al.* Avaliação do sono e fatores associados em trabalhadores de enfermagem hospitalar. **Texto & Contexto Enfermagem**, v. 31, e20220277, 2022.

SILVA, VC *et al.* Inteligência artificial na avaliação do consumo alimentar: avanços na análise da dieta e predição de desfechos cardiovasculares. **Revista da Sociedade de Cardiologia do Estado de São Paulo**, v. 32, n. supl. 1, p. 90-97, 2022.

SILVA, VM; SOUSA, MR; MARTUSCELLO NETO, C. Uma análise dos transtornos de ansiedade: ansiedade generalizada, pânico e ansiedade social. **Revista Eletrônica Acervo Médico**, v. 23, n. 7, e13531, 2023.

SIMONETTO, GF. **Estudo de possíveis discriminantes de TEAF utilizando dados do DATASUS**. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação). Universidade Federal de Santa Catarina. Florianópolis (SC), 2022.

SOARES JUNIOR, NJS. **Obtenção de aplicativo para uso na predição de risco hipertensivo e nível de atividade física em adolescentes**. Tese (Doutorado em Biotecnologia em Saúde). Programa de Pós-Graduação em Biotecnologia - RENORBIO/CCBS. Universidade Federal do Maranhão. São Luís (MA), 2023.

TONIDANDEL, PH; SILVA, RA; BRUN, AL. Meta aprendizado aplicado ao câncer de mama. **Anais Escola Regional de Banco de Dados (ERBD)**, 18, 2023, Palmas/PR. Porto Alegre: Sociedade Brasileira de Computação, 2023. p. 90-99.

VILELA JUNIOR, GB. Consumo de bebidas alcoólicas e hipertrofia muscular. **Revista CPAQV – Centro de Pesquisas Avançadas em Qualidade de Vida**, v. 15, n. 2, 2023.