

FACULDADE  
**SANTO AGOSTINHO**  
ITABUNA - BA



**COLEGIADO DO CURSO DE MEDICINA**

BERNARDO DE OLIVEIRA SALUME  
DAVID GOMES DO NASCIMENTO  
RAUL DIAS FITERMAN

**PREDIÇÃO DE ACIDENTE VASCULAR CEREBRAL POR  
MEIO DO APRENDIZADO DE MÁQUINA: UMA REVISÃO  
SISTEMÁTICA**

ITABUNA – BAHIA

DEZEMBRO/2021

BERNARDO DE OLIVEIRA SALUME  
DAVID GOMES DO NASCIMENTO  
RAÚL DIAS FITERMAN

**PREDIÇÃO DE ACIDENTE VASCULAR CEREBRAL POR  
MEIO DO APRENDIZADO DE MÁQUINA: UMA REVISÃO  
SISTEMÁTICA**

Trabalho de Conclusão de Curso, em formato de revisão sistemática, apresentado como requisito parcial à obtenção do título de bacharel em Medicina pela Faculdade Santo Agostinho de Itabuna, sob a orientação do Professor João Thadeu Santos Cerqueira.

ITABUNA - BAHIA  
DEZEMBRO/2021

S181p

SALUME, Bernardo de Oliveira

**Predição de Acidente Vascular Cerebral por meio do aprendizado de máquina:**  
uma revisão sistemática /Bernardo de Oliveira Salume [et al], --Itabuna: FASAI, 2021.  
25p.

Orientador:Prof.º João Thadeu Santos Cerqueira.

Trabalho de Conclusão de Curso em Bacharelado em Medicina pela  
Faculdades Santo Agostinho de Itabuna, BA – FASAI.

1.Aprendizado de Máquinas. 2.AVC( Acidente Vascular Cerebral).3.Probabilidade.  
I.Salume, Bernardo de Oliveira. II.Cerqueira, João Thadeu Santos.  
III. Faculdades Santo Agostinho de Itabuna. IV. Título.

CDU:616

Catálogo Biblioteca Dr. º Maria Odília Teixeira  
Aline Andrade Ferraz – Bibliotecária CRB 5/001881/O

SALUME, Bernardo; NASCIMENTO, David; FITERMAN, Raul. **Predição de Acidente Vascular Cerebral por meio do Aprendizado de Máquinas: uma Revisão Sistemática.** Trabalho de Conclusão de Curso, em formato de revisão sistemática, apresentado como requisito parcial a obtenção de título de bacharel em Medicina pela Faculdade Santo Agostinho de Itabuna, sob a orientação do Professor João Thadeu Santos Cerqueira. Itabuna: Faculdade Santo Agostinho de Itabuna, 2021.

### FOLHA DE APROVAÇÃO

**PREDIÇÃO DE ACIDENTE VASCULAR CEREBRAL POR MEIO DO  
APRENDIZADO DE MÁQUINAS: UMA REVISÃO SISTEMÁTICA.**

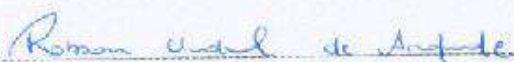
**SALUME, BERNARDO; NASCIMENTO, DAVID; FITERMAN, RAUL.**

Aprovado em 20 de dezembro de 2021.

### BANCA EXAMINADORA



Nome completo do orientador  
Faculdade Santo Agostinho de Itabuna



Nome completo do membro 01  
Faculdade Santo Agostinho de Itabuna



Nome completo do membro 02  
Faculdade Santo Agostinho de Itabuna

## AGRADECIMENTOS

A Deus, pela nossa vida, e por nos permitir ultrapassar todos os obstáculos encontrados ao longo da realização deste trabalho.

Aos nossos pais e familiares, que nos incentivaram nos momentos difíceis e compreenderam a nossa ausência enquanto nos dedicávamos à realização deste trabalho.

Ao professor João Thadeu Cerqueira, por ter sido nosso orientador e ter desempenhado tal função com tamanha dedicação e amizade.

Aos nossos colegas de turma, por compartilharem conosco tantos momentos de descobertas e aprendizado e por todo o companheirismo ao longo deste percurso.

Gostaríamos de enviar um singelo abraço ao querido amigo Vitor Mendes Cavaiheiro, o qual não mediu esforços para nos ajudar neste trabalho.

## LISTA DE FIGURAS

Figura 01 – Fluxograma PRISMA.

## LISTA DE TABELAS

**Tabela 01** – Um breve guia de introdução ao machine learning.

**Tabela 02** – Dados Gerais dos artigos de Aprendizado de Máquinas.

**Tabela 03** – Performance preditiva dos modelos de Aprendizado de Máquina.

## LISTA DE ABREVIATURAS E SIGLAS

AVC. – Acidente Vascular Cerebral.

OMS. – Organização Mundial da Saúde.

IAM. – Infarto Agudo do Miocárdio.

HAS. – Hipertensão Arterial Sistêmica.

SUS. – Sistema Único de Saúde.

SIM. – Sistema de Informação sobre Mortalidade.

FA. – Fibrilação Atrial.

PRISMA. – Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

PICO. – Problema, Intervenção, Controle e Outcome (desfecho).

AR. – Artrite reumatoide.

LDL. – Lipoproteína de Baixa Densidade.

IMC. – Índice de Massa Corporal.

AUC. – Area Under the Curve.

ACS. – American College of Surgeons.

ASA. – American Society of Anesthesiology.

DCV. – Doença Cardiovascular.

AVCI. – Acidente Vascular Cerebral Isquêmico.

AVCH. – Acidente Vascular Cerebral Hemorrágico.

## SUMÁRIO

|                              |    |
|------------------------------|----|
| 1 INTRODUÇÃO .....           | 5  |
| 2 METODOLOGIA .....          | 8  |
| 3 RESULTADOS .....           | 8  |
| 4 DISCUSSÃO .....            | 15 |
| 5 CONSIDERAÇÕES FINAIS ..... | 17 |
| REFERÊNCIAS .....            | 17 |
| ANEXOS .....                 | 19 |

Predição de Acidente Vascular Cerebral por meio do Aprendizado de Máquinas: uma  
Revisão Sistemática

*A Machine Learning approach on prediction of Stroke: a Systematic Review*

SALUME, Bernardo<sup>1</sup>

NASCIMENTO, David<sup>2</sup>

FITERMAN, Raul<sup>3</sup>

CERQUEIRA, Thadeu<sup>4</sup>

**RESUMO**

**Introdução:** De acordo com a Organização Mundial da Saúde, o Acidente Vascular Cerebral (AVC) é a segunda maior causa de morte no mundo. Cerca de 70% dos AVCs registrados são o primeiro evento, de forma que garantir a sua prevenção é essencial. Escores de riscos convencionais foram desenvolvidos para estimar o risco de doença cardiovascular/cerebrovascular, porém, por meio de técnicas computacionais de análise de dados, vislumbra-se a possibilidade de realizar pesquisas a níveis populacionais e que integram um maior leque de variáveis, de modo que a predição final do modelo seja superior aos métodos tradicionais. **Objetivo:** Investigar a aplicabilidade do aprendizado de máquinas na predição do primeiro episódio de AVC. **Metodologia:** Uma pesquisa foi conduzida nas bases de dados Pubmed, Web of Science e Embase em novembro de 2021. Os termos utilizados foram "machine learning" AND "stroke" AND "probability". **Resultados:** Apenas 6 artigos estiveram de acordo com os critérios de inclusão e exclusão, sendo incluídos na análise final. Os objetivos dos estudos variaram da predição de AVC em determinado tempo futuro, ao risco desse desfecho relacionado a complicações pós-cirúrgicas e doenças específicas, como Artrite Reumatoide e Fibrilação Atrial. O principal modelo encontrado foi a Árvore de Decisão. Nos estudos que houveram comparação com modelos estatísticos convencionais, houve uma performance similar ou melhor dos algoritmos de aprendizado de máquinas. **Considerações finais:** notou-se um crescente interesse no uso de aprendizado de máquinas, que pode alcançar altos níveis de acurácia, sensibilidade e especificidade. Entretanto, os estudos apresentam limitações, mas indicam que – num futuro próximo – os modelos de aprendizado de máquinas poderão substituir ou se aliar às análises estatísticas tradicionais.

**Palavras-chave:** Aprendizado de Máquinas, AVC (Acidente Vascular Cerebral), Probabilidade.

<sup>1</sup>Graduando em Medicina pela Faculdade Santo Agostinho de Itabuna.

<sup>2</sup>Graduando em Medicina pela Faculdade Santo Agostinho de Itabuna.

<sup>3</sup>Graduando em Medicina pela Faculdade Santo Agostinho de Itabuna.

<sup>4</sup>Titulação, Professor orientador. Docente do Curso de Medicina da Faculdade Santo Agostinho de Itabuna.

## ABSTRACT

**Background:** According to the World Health Organization, stroke is the second major cause of death in the world. About 70% of registered strokes are new-onset, so ensuring their prevention is essential. Conventional risk scores were developed to estimate the risk of cardiovascular/cerebrovascular disease, however, through computational data analysis techniques, it is possible to achieve surveys at populational levels that integrate a greater range of variables, resulting in a final prediction superior to traditional methods. **Objective:** To investigate the applicability of machine learning in predicting the first episode of stroke. **Methodology:** A search was conducted in the Pubmed, Web of Science and Embase databases in November 2021. The terms used were "machine learning" AND "stroke" AND "probability". **Results:** Only 6 articles met the inclusion and exclusion criteria, being included in the final analysis. The objectives of the studies ranged from predicting stroke at a certain future time to the risk of this outcome related to post-surgical complications and specific diseases, such as Rheumatoid Arthritis and Atrial Fibrillation. The most frequent model used was the Decision Tree. In studies that compare with conventional statistical models, there was a similar or better performance of machine learning algorithms. **Conclusions:** There was a growing interest in the use of machine learning, which can achieve high levels of accuracy, sensitivity and specificity. However, the studies have limitations, but they indicate that – in the near future – machine learning models will replace or combine with traditional statistical analyses.

**Keywords:** Machine Learning, Stroke, Probability

## 1 INTRODUÇÃO

AVC é a sigla para acidente vascular cerebral. Há diferentes tipos de AVC, mas geralmente ele acontece após a rápida morte de neurônios, causada por alguma interrupção (isquemia) ou rompimento de vasos sanguíneos (hemorragia) no cérebro. O AVC normalmente leva à perda brusca de algumas funções neurológicas (MALHEIROS e col., 2017).

De acordo com a OMS (2020), o AVC é a segunda maior causa de morte no mundo, resultando em 11% do total do número de óbitos, ficando atrás somente do IAM. No ano de 2019, as mortes por AVC – registradas – resultaram em mais de 6 milhões em todo o mundo. Conforme o Ministério da Saúde (2021), o AVC isquêmico pode ocorrer devido a um trombo (trombose) ou a um êmbolo (embolia) e é o tipo mais comum, representando cerca de 85% de todos os casos. Já o AVC hemorrágico ocorre nos outros 15% dos casos, mas pode causar

morte com mais frequência do que o isquêmico, tendo a HAS e a ruptura de aneurisma como as principais causas.

No Brasil, em 2016, foram notificadas 188.223 internações para o tratamento de AVC isquêmico e hemorrágico no SUS. Sobre as mortes, a pasta registrou em 2016, no Sistema de Informação sobre Mortalidade (SIM), 40.019 óbitos por AVC (BRASIL, 2018).

Os principais fatores de risco para desenvolvimento de AVC são: idade avançada (> 65 anos), sexo masculino, hipertensão arterial sistêmica, diabetes tipo 2, dislipidemia, sobrepeso/obesidade, tabagismo, etilismo, sedentarismo, histórico familiar e uso de drogas ilícitas (BRASIL, 2021).

De acordo com Diener (2020), a fibrilação atrial (FA) é um importante fator de risco para AVC.

Um cálculo tradicional de risco cardiovascular (ex. Escore de Framingham, CHADS<sub>2</sub> e CHA<sub>2</sub>DS<sub>2</sub>-VASc) pode auxiliar na estimativa de risco de evento cardiovascular/cerebrovascular em até 10 anos (MESCHIA et al., 2014).

A maior parte (> 70%) dos AVCs é o primeiro evento. A prevenção do primeiro AVC é de suma importância (PARMAR et al., 2014). Levando em consideração, o aprendizado de máquina oferece um resultado de previsão de forma rápida e precisa, se tornando uma ferramenta importante para a saúde.

A partir século XXI, várias áreas da Medicina, graças à evolução da ciência computacional, foram capazes de serem aprimoradas pela aprendizagem de máquina (HANDELMAN, 2018). Na tabela 1, encontram-se alguns termos comuns em estudos de aprendizado de máquinas.

Através da junção entre a "big data" e os prontuários eletrônicos de saúde, podemos usar os dados clínicos populacionais para conseguir uma visão mais fidedigna do que é o mundo real. Isso nos permite explorar soluções para dificuldades na saúde da população que antes eram inalcancáveis. Análises preditivas podem ser feitas através do aprendizado de máquinas, como por exemplo a identificação de subgrupos de pacientes com um risco maior de admissão hospitalar e que poderão ser responsáveis por custos aumentados em saúde. Sendo assim, essas previsões podem ajudar a desenvolver intervenções preventivas a fim de diminuir tais custos e também a morbidade dos pacientes (NGIAM, 2019).

Esse estudo tem como objetivo investigar a aplicabilidade do aprendizado de máquina na predição do primeiro episódio de AVC. Nesse sentido, não houve uma limitação dos métodos de aprendizado ou modelos de algoritmos buscados, muito menos do cenário em que o paciente se encontra, seja intra-hospitalar, ambulatorial ou decorrente de alguma complicação. A escolha por uma revisão sistemática ocorreu pela necessidade de avaliar o viés dos estudos

encontrados, haja que o uso do aprendizado de máquinas na área da saúde é um fato recente e ainda existe um desconhecimento das técnicas de análise de dados ligadas a inteligência artificial por parte dos profissionais da saúde. Desse modo, esse artigo busca compreender as experiências publicadas por outros pesquisadores e avaliar o nível de evidência delas, para julgar se a aprendizagem de máquina, de fato, pode substituir ou somar na análise estatística tradicional.

Tabela 1: Um breve guia de introdução ao aprendizado de máquina.

|                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Inteligência artificial         | Capacidade de máquinas de simular a inteligência humana. Inteligência pode ser definida como a habilidade de dar a solução mais adequada para determinado problema, de acordo com os dados disponíveis no momento da decisão, com aptidão para se adaptar a novas situações (GAMEIRO et al. 2019, apud CHIAVEGATTO, 2018).                                                                                                                       |
| Aprendizado de máquina          | Modelos algorítmicos e estatísticos programáveis que processam os dados inseridos para então executar a tarefa desejada, por exemplo, predição de um desfecho ou inferência (relação) entre variáveis e o desfecho (DU-HARPUR et al, 2020).                                                                                                                                                                                                      |
| Aprendizado supervisionado      | Os dados inseridos no programa para treiná-lo contêm o desfecho e o objetivo é identificar as relações entre ambos, de forma que o modelo consiga, por exemplo, prever um desfecho em dados novos apenas com o que foi inserido. Existem diversos modelos de algoritmo supervisionados, como: support vector machines, gradient boosted trees, deep learning (BADILLO et al, 2020).                                                              |
| Aprendizado não-supervisionado  | Os dados são explorados pelo algoritmo que reconhece padrões automaticamente e subdivide-os em grupos similares, sem a necessidade de um treinamento prévio. As principais técnicas utilizadas são o cluster e redução de dimensionalidade (ROOHI et al, 2020).                                                                                                                                                                                  |
| Aprendizado semi-supervisionado | Usa-se tanto de dados com e sem rótulos no algoritmo. Em geral, há poucos dados rotulados, mas sua presença permite otimizar a precisão do algoritmo. Um exemplo prático é a nomeação de fotos em redes sociais: o algoritmo identifica automaticamente a mesma pessoa em múltiplas fotos e, após o usuário nomeá-la em uma ou poucas das fotos, o programa aplica o rótulo (nome) em todas as fotos restantes (GUERREIRO, 2017 apud ZHU, 2005). |
| Aprendizado por reforço         | Dado um problema X, o algoritmo busca a melhor forma de resolvê-lo, por meio da tentativa e erro, sendo que para cada resultado ele pode ser recompensado ou punido (caso acerte ou não o desfecho esperado). O algoritmo busca maximizar as recompensas (JOHNSON et al. 2018).                                                                                                                                                                  |

## 2 METODOLOGIA

Seguimos o protocolo PRISMA para reportar os dados. As pesquisas foram realizadas nas bases de dados PubMed, Web of Science e Embase, em novembro de 2021. Os termos utilizados para a pesquisa foram: “*machine learning*” AND “*stroke*” AND “*probability*”, construídos a partir do PICO. Não houve limitação no período das publicações e nem na língua dos estudos incluídos.

Os artigos foram seleccionados conforme os seguintes critérios de inclusão: sem restrição para idade e que desenvolveram modelos com multivariáveis para predição de AVC. Já os critérios de exclusão foram: pacientes previamente diagnosticados com AVC e trauma.

Os dados extraídos foram: o objetivo e tipo do estudo, a amostra (participantes, base de dados e datas), as variáveis (selecção e dados desconhecidos), modelo de aprendizado de máquina (performance, calibração e validação externa) e comparação com outros modelos.

Foi utilizado um organizador de citações (EndNote) para compilar os artigos de todas as bases de dados e excluir – automaticamente – os duplicados, sendo que alguns foram excluídos manualmente.

Após a exclusão das duplicatas e dos artigos de revisão, os estudos foram analisados de forma independente, por três pesquisadores, por meio do título e resumo. Foram incluídos somente os artigos que, pelo menos, dois pesquisadores concordaram entre si. Um quarto pesquisador foi consultado em caso de discordância referente à inclusão, ou não, do estudo.

Os textos completos dos artigos foram examinados de forma independente pelos três pesquisadores conforme os critérios de elegibilidade.

## 3 RESULTADOS

Nós identificamos 230 estudos no PubMed, 114 no Web Of Science e 67 no Embase, totalizando 411 artigos. Depois de remover os duplicados e as revisões, sobraram 326. Após a análise dos títulos e resumos de cada, 49 artigos foram considerados potencialmente relevantes, sendo que 2 não foram encontrados na íntegra. Depois do rastreio do texto completo de cada artigo, 6 estudos foram seleccionados para extração dos dados (Figura 1).

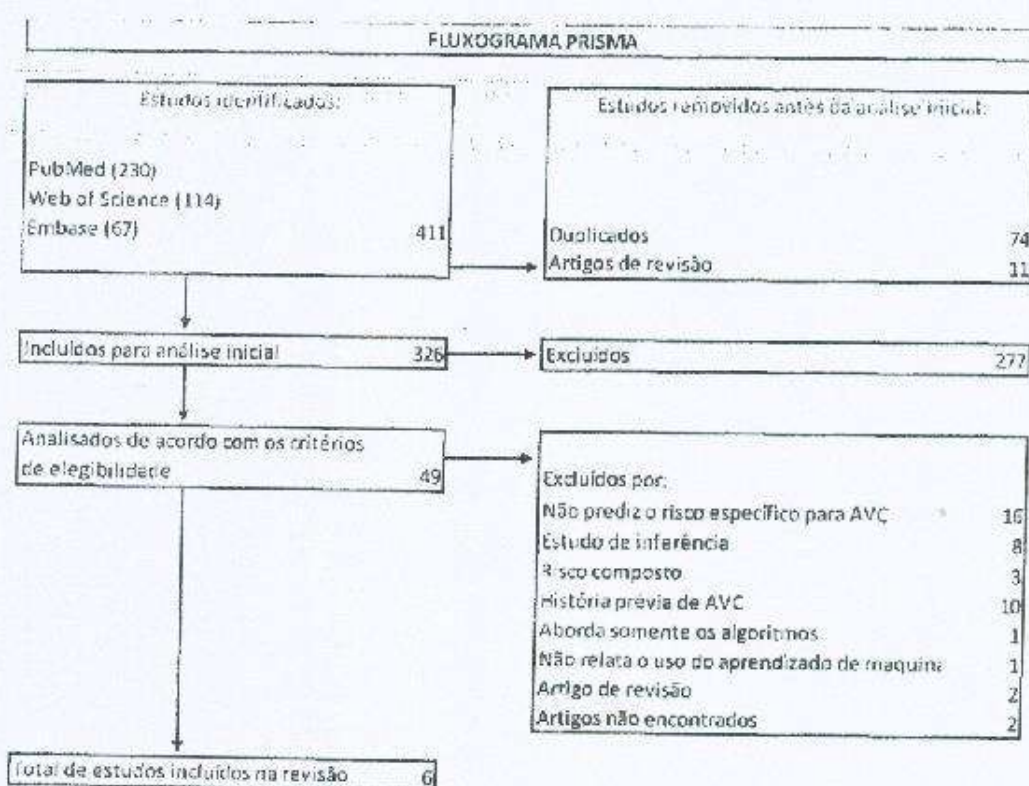


Figura 1: Fluxograma PRISMA.

Os artigos escolhidos foram publicados em 2019-2021. Dos artigos encontrados, o objetivo do estudo foi: predição de AVC relacionada a complicações pós-operatórias em cirurgias hepáticas, pancreáticas e colorretais; avaliação de risco de AVC em um determinado período futuro; predição de AVC relacionada a doenças específicas (FA e AR), e a estratificação de risco voltada para indivíduos com baixos níveis de LDL e a contribuição de outras anormalidades metabólicas para o risco.

Dos 6 artigos encontrados, os principais algoritmos utilizados foram: Árvore de Decisão (e suas variações), sendo o restante Ensemble (Rede Neural convolucional + Floresta Aleatória + CHA<sub>2</sub>DS<sub>2</sub>VASC), Nomograma de Regressão Logística, e Rede Neural profunda.

As variáveis mais comuns – presentes em três ou mais artigos – foram: idade, sexo, hipertensão arterial sistêmica, insuficiência cardíaca, tabagismo, glicemia em jejum, IMC, histórico de medicações (ex. anti-hipertensivos), diabetes mellitus e fibrilação atrial. Dois artigos relataram a inclusão de variáveis desconhecidas por meio de técnicas, como MedInput e o SPSS.

Cinco artigos relataram divisão da amostra total em uma parte para treinamento, testagem e ajustes, sendo que um não especificou como a amostra foi subdividida. Apenas dois artigos

utilizaram validação externa e quatro artigos relataram o método de calibração, sendo o mais comum Hosmer-Lemeshow.

Dentre os artigos focados em análise de predição de risco para AVC num tempo futuro, o modelo que utiliza Árvore de Decisão (ou variações) obteve uma AUC superior ao escore de risco de Framingham revisado (0,7403 e 0,6491, respectivamente). Já o modelo usando Rede Neural profunda, apresentou uma performance semelhante ao QRISK3 (0,896 e 0,880, em mulheres, e 0,870 e 0,858, em homens, respectivamente). Todavia, a Rede Neural profunda mostrou-se melhor que o escore de Framingham, QRISK1 e QRISK2.

A análise por meio do Ensemble, em pacientes com FA em uso de dispositivos cardíacos eletrônicos implantáveis, permitiu a predição de AVC para os próximos 30 dias. Em comparação com CHA<sub>2</sub>DS<sub>2</sub>-VASc, não houve alteração nesse período, mostrando a limitação do escore tradicional.

Em pacientes que fizeram procedimentos cirúrgicos hepáticos, pancreáticos ou colorretais, a análise por meio de um modelo de Árvore de Decisão permitiu avaliar desfechos que incluam desde qualquer complicação pós-operatória até as mais específicas, como o AVC. Nesse estudo, não foi possível fazer comparações com escores convencionais para o risco de AVC, pois o ACS Surgical Risk Calculator e a classificação da ASA atestam apenas qualquer complicação, não as mais específicas.

A análise feita por um Nomograma de Regressão Logística, em pacientes com artrite reumatoide (AR), demonstrou que os fatores de risco mais relevantes foram histórico de doença cardiovascular (DCV) e perfil lipídico. Esse modelo não atribuiu aos marcadores inflamatórios, um papel tão relevante na predição de AVC.

Entre os artigos, no que teve como base o uso do modelo criado Survival Conditional Inference Tree (SCTREE), pôde-se concluir que pressão arterial mal controlada, associada a uma concentração baixa de LDL-C e uma glicemia elevada, foram os principais preditores para desenvolvimento de AVCI e/ou AVCH.

Tabela 2. Dados Gerais dos artigos de Aprendizado de Máquinas

| 1°                                                                                                                                                |  |                                                                                                                                        |  |                                                                                                                                                                                                                                     |  |
|---------------------------------------------------------------------------------------------------------------------------------------------------|--|----------------------------------------------------------------------------------------------------------------------------------------|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| Autores                                                                                                                                           |  | 2°                                                                                                                                     |  | 3°                                                                                                                                                                                                                                  |  |
| Katuscha Merath et al. <sup>16</sup>                                                                                                              |  | Agni Orfanoudaki et al. <sup>17</sup>                                                                                                  |  | Chen-Ying Hung et al. <sup>18</sup>                                                                                                                                                                                                 |  |
| Use of Machine Learning for Prediction of Patient Risk of Postoperative Complications After Liver, Pancreatic, and Colorectal Surgery.            |  | Machine learning provides evidence that stroke risk is not linear. The non-linear Framingham stroke risk score                         |  | Development of an intelligent decision support system for ischemic stroke risk assessment in a population-based electronic health record database                                                                                   |  |
| Desenvolver um modelo de machine learning para prever o risco de complicações pós-operatórias em cirurgias hepáticas, pancreáticas e colorretais. |  | Desenvolver um algoritmo de machine learning a partir da base de dados do estudo de Framingham para estimar o risco de AVC em 10 anos. |  | Desenvolver um modelo de predição para estimar o risco de AVC isquêmico em até 3 anos na população geral.                                                                                                                           |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | 4°                                                                                                                                                                                                                                  |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Zhijun Wu et al. <sup>21</sup>                                                                                                                                                                                                      |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | The risk of ischemic stroke and hemorrhagic stroke in Chinese adults with low-density lipoprotein cholesterol concentrations < 70 mg/dL.                                                                                            |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Avaliar se modelos de machine learning podem prever o risco de AVC em 30 dias em pacientes com Fibrilação Atrial (FA) utilizando dispositivos eletrônicos implantáveis (DCEI) quando comparados com escore de riscos convencionais. |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | 5°                                                                                                                                                                                                                                  |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Lichy Han et al. <sup>20</sup>                                                                                                                                                                                                      |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Atrial Fibrillation Burden Signature and Near-Term Prediction of Stroke A Machine Learning Analysis                                                                                                                                 |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | 6°                                                                                                                                                                                                                                  |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Fangnan Xin et al. <sup>19</sup>                                                                                                                                                                                                    |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Development and validation of a nomogram for predicting stroke risk in rheumatoid arthritis patients.                                                                                                                               |  |
|                                                                                                                                                   |  |                                                                                                                                        |  | Desenvolver e validar um nomograma incorporando o perfil lipídico e marcadores inflamatórios como variáveis para avaliar a predição de risco de AVC em pacientes com artrite reumatoide (ARA).                                      |  |

Tabela 2. (Continuação)

| Amostra | 15.657<br>pacientes                        | FDI: 18.793 visitas<br>de 4.385 pacientes;<br>(treinamento<br>e<br>testagem)<br>FDI: 2.989 pacientes<br>(apenas testagem);<br>Validação externa:<br>17.527 pacientes. | 11.192.916<br>registros<br>de<br>840.487<br>pacientes | 9.327<br>participantes<br>para<br>treinamento,<br>1.753<br>participantes<br>para<br>validação<br>externa | 3.114 casos sem AVC e<br>71 casos com AVC (caso-<br>controle)                       | 1.935<br>pacientes<br>com<br>artrite<br>reumatoide |
|---------|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|----------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|----------------------------------------------------|
| Estudo  | Coorte<br>retrospectiva                    | Coorte retrospectiva                                                                                                                                                  | Coorte<br>retrospectiva                               | Coorte<br>retrospectiva                                                                                  | Coorte<br>retrospectiva                                                             | Coorte<br>retrospectiva                            |
| Modelo  | Árvore<br>de<br>Decisão<br>(Decision Tree) | Optimal<br>Classification Trees<br>(Árvore de Decisão)                                                                                                                | Rede Neural<br>Profunda (Deep<br>Neural<br>Network)   | Survival<br>Conditional<br>Inference<br>(Árvore<br>de<br>Decisão)                                        | Ensemble: (Convolutional<br>Neural Networks +<br>Random Forest + C11A2<br>DS2-VASc) | Modelo de<br>regressão<br>logística                |
| Ano     | 2019                                       | 2020                                                                                                                                                                  | 2019                                                  | 2021                                                                                                     | 2019                                                                                | 2021                                               |



