

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA (NEUROLOGIA)

MANOELA GOMES DOMINGOS DA SILVA

Desenvolvimento e avaliação de um aplicativo de celular com uso de inteligência artificial para diagnóstico da etiologia da hemorragia intraparenquimatosa cerebral

Ribeirão Preto

2022

MANOELA GOMES DOMINGOS DA SILVA

Desenvolvimento e avaliação de um aplicativo de celular com uso de inteligência artificial para diagnóstico da etiologia da hemorragia intraparenquimatosa cerebral.

Versão corrigida.

A versão original encontra-se disponível tanto na Biblioteca da Unidade que aloja o Programa, quanto na Biblioteca Digital de Teses e Dissertações da USP (BDTD)

Dissertação apresentada ao Programa de Pós-Graduação em Medicina (Neurologia), da Faculdade de Medicina de Ribeirão Preto para obtenção do título de Mestre em Ciências.

Área de concentração : Neurociências

Orientador:

Prof. Dr. Octávio Marques Pontes-Neto

Ribeirão Preto

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Silva, Manoela Gomes Domingos da
Desenvolvimento e avaliação de um aplicativo de
celular com uso de inteligência artificial para diagnóstico da
etiologia da hemorragia intraparenquimatosa cerebral, 2022.
86 p. : il. ;

Dissertação de Mestrado, apresentada à Faculdade de Medicina de
Ribeirão Preto/USP. Área de concentração: Neurologia.
Orientador: Marques Pontes Neto, Octávio.

1. Hemorragia intraparenquimatosa cerebral.
2. Investigação etiológica.
3. Aplicativos móveis.
4. Inteligência artificial.
5. Aprendizado semi-supervisionado.

Nome: SILVA, Manoela Gomes Domingos da

Título: Desenvolvimento e avaliação de um aplicativo de celular com uso de inteligência artificial para diagnóstico da etiologia da hemorragia intraparenquimatosa cerebral

Aprovado em:

Banca Examinadora

Prof. Dr (a).

Instituição:

Julgamento:

Profa. Dr (a).

Instituição:

Julgamento:

Prof. Dr (a).

Instituição:

Julgamento:

AGRADECIMENTOS

À minha família, em especial, meus pais, Alex e Tânia e minha avó, Marlene, por sempre me mostrarem a importância do conhecimento para a formação de uma pessoa.

Ao meu namorado, Gustavo, por todo o apoio e paciência durante a pós-graduação.

Ao meu orientador Prof. Dr. Octávio, pelos ensinamentos e orientações e pela enriquecedora oportunidade de participar do seu laboratório de pesquisa.

Ao Dr. Otávio Vicenzi, pela ajuda essencial na organização e classificação dos dados.

Aos amigos do laboratório, Thiago, Pamella, Flávia, Karina, Luan, Diandra e Eduardo, pelos preciosos momentos de aprendizagem.

À Dra. Millene Camilo, pelas sugestões importantes para o trabalho.

À equipe do REAVER, por todo o trabalho e dedicação na coleta, registro e organização de dados para o laboratório.

Aos amigos da empresa Zeus e do Projeto Levacc, por me compreenderem e apoiarem durante todo o processo, em especial, Thiago, pelas dicas de escrita da tese.

À instituição cujo fomento foi imprescindível para o presente trabalho. Este projeto foi viabilizado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

A todos que, de alguma forma, me motivaram a superar os desafios da pós-graduação.

"Não sei se cada um tem um destino ou se só flutuamos sem rumo, como numa brisa... Mas acho que talvez sejam ambas as coisas. Talvez as duas coisas aconteçam ao mesmo tempo."
Forrest Gump

SILVA, M. G. D. Desenvolvimento e avaliação de um aplicativo de celular com uso de inteligência artificial para diagnóstico da etiologia da hemorragia intraparenquimatosa cerebral. 2022. Dissertação (Mestrado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2022.

RESUMO

Introdução: O acidente vascular cerebral (AVC) tem mais de 150 causas conhecidas, e por isso é considerada uma doença heterogênea. Dentre os subtipos de AVC, a hemorragia intraparenquimatosa cerebral é o de pior prognóstico. Embora investigar a etiologia do AVC seja fundamental para a prevenção de um novo episódio, estima-se que esta investigação ocorre de forma adequada no nosso meio em menos de 30% dos casos. No caso da hemorragia intraparenquimatosa cerebral (HIC), diversos sistemas de classificação foram propostos (como o sistema SMASH-U), mas ainda não existe um sistema ideal para este fim. Neste cenário, o uso de um aplicativo para classificar de forma automática a etiologia de um episódio de hemorragia intraparenquimatosa cerebral (HIC) é uma possível forma de otimizar a investigação etiológica na prática clínica.

Objetivo: Este estudo visa desenvolver e validar um aplicativo de celular que faz uso de métodos de Aprendizado de Máquina para prever a etiologia da hemorragia intraparenquimatosa cerebral de pacientes usando como base a classificação etiológica SMASH-U.

Metodologia: Para este estudo foram selecionados dados de forma retrospectiva de 66 pacientes com HIC consecutivos admitidos no HCFMRP-USP entre 01/01/2018 e 31/12/2019 e incluídos de forma prospectiva no Registro de Acidente Vascular Cerebral de Ribeirão Preto (REAVR). Dos dados coletados, 30% foi separado para a validação externa e 70% foi utilizado para a comparação dos modelos. Foram testados 18 pipelines, compostos por três fases: rotulação de instâncias não rotuladas (Máquina de Vetor de Suporte, Florestas Aleatórias, Redes Neurais Artificiais), seleção de atributos (Regressão Logística, Qui-Quadrado), e classificação supervisionada da etiologia SMASH-U (Máquina de Vetor de Suporte, Florestas Aleatórias, Redes Neurais Artificiais). O pipeline com melhor desempenho foi escolhido para servir ao aplicativo. As classes foram balanceadas com o método *Random Oversampling* e a AUC - do inglês “*Area Under the Curve*”, que corresponde à área sob a curva ROC - foi obtida por cross-validation (k=5). Os dados de validação externa foram

classificados pelo pipeline escolhido para servir ao aplicativo e a acurácia inter-classe foi obtida. O aplicativo para celulares foi desenvolvido através do *framework* React Native e se comunica com o servidor por meio de requisições. O estudo foi aprovado no comitê de ética em pesquisa do HCFMRP-USP e obtivemos dispensa do termo de consentimento.

Resultados: A amostra foi composta de 66 pacientes com HIC. Do total de pacientes, a maioria eram sexo masculino (60,6 %), declarados brancos (71,2%), e com idade média de 60,8 e desvio padrão de 13,12. Quanto à classificação SMASH-U dos episódios, a etiologia primária hipertensiva foi a mais comum (46 pacientes - 69,7%), seguida pelas etiologias indefinida (9 pacientes - 13,6%), sistêmica (3 pacientes - 4,6%), angiopatia amilóide (3 pacientes - 4,6%), estrutural (3 pacientes - 4,6%) e medicação (2 pacientes - 3,0%).

O pipeline escolhido obteve a AUC de 0,99 e na validação externa o modelo acertou 80% das etiologias HAS, 20% das U e 0% das demais. O aplicativo móvel desenvolvido possui interface de fácil usabilidade e apresenta duas funcionalidades principais: a classificação da etiologia de dados de um único paciente preenchidos por um formulário, e a classificação em lote para uma lista de muitos pacientes. O aplicativo obteve desempenho satisfatório, boa responsividade nos aparelhos testados, e funciona nas plataformas iOS e Android.

Conclusões: Desenvolvemos e validamos um aplicativo de celular que faz uso de métodos de Aprendizado de Máquina para prever a etiologia da hemorragia intraparenquimatosa cerebral de pacientes usando como base a classificação etiológica SMASH-U com boa acurácia e usabilidade. O aplicativo possui funcionalidades que podem auxiliar a investigação etiológica na prática clínica, assim como facilitar a rotulação de grande quantidade de dados para pesquisas.

PALAVRAS-CHAVE: Hemorragia intraparenquimatosa cerebral. Investigação etiológica. Aplicativos móveis. Inteligência artificial. Aprendizado semi-supervisionado.

SILVA, M. G. D. Development and evaluation of a mobile application using artificial intelligence to diagnose the etiology of cerebral intraparenchymal hemorrhage. 2022. Dissertação (Mestrado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2022.

ABSTRACT

Introduction: Stroke has more than 150 known causes, and therefore is considered a heterogeneous disease. Among hemorrhagic stroke subtypes, intraparenchymal hemorrhage is the one with the worst prognostic. It's caused by the rupture of a vessel and its consequent extravasation of blood into the brain. Although investigating the etiology of the stroke is essential for the prevention of a new episode, it occurs in only 30% of the cases. Machine In The use of an application to automatically classify the etiology of an episode of cerebral intraparenchymal hemorrhage (ICH) is a possible way to optimize the etiological investigation in clinical practice.

Objective: This study aims to develop and validate a mobile application that uses Machine learning methods to predict a SMASH-U of an episode of intraparenchymal hemorrhage. **Methods:** For this study, data were retrospectively selected from 66 consecutive ICH patients admitted to the HCFMRP-USP between 01/01/2018 and 12/31/2019 and prospectively included in the Ribeirão Preto Cerebrovascular Accident Registry (REAVAR). Of the collected data, 30% was set aside for external validation and 70% was used for model comparison. Eighteen pipelines were tested, consisting of three phases: labeling of unlabeled instances (Support Vector Machine, Random Forests, Artificial Neural Networks), attribute selection (Logistic Regression, Chi-Square), and supervised classification of the SMASH-U etiology (Support Vector Machine, Random Forests, Artificial Neural Networks). The best performing pipeline was chosen to serve the application. The classes were balanced using the Random Oversampling method and the AUC (“Area Under the ROC Curve”) was obtained by cross-validation (k=5). The external validation data was sorted by the chosen pipeline to serve the application and the inter-class accuracy was obtained. The mobile application was developed using the React Native framework and communicates with the server through

requests. The study was approved by the Research Ethics Committee of the HCFMRP-USP and we obtained a waiver of the consent form.

Results: The sample consisted of 66 patients with ICH. Of the total number of patients, most were male (60.6%), declared white (71.2%), with a mean age of 60.8 and standard deviation of 13.12. As for the SMASH-U classification of the episodes, the primary hypertensive etiology was the most common (46 patients - 69.7%), followed by undefined (9 patients - 13.6%), systemic (3 patients - 4.6%), amyloid angiopathy (3 patients - 4.6%), structural (3 patients - 4.6%) and medication (2 patients - 3.0%). The chosen pipeline obtained an AUC of 0.99 and in the external validation the model was correct in 80% of the SAH etiologies, 20% of the U and 0% of the others. The developed mobile application has an easy-to-use interface and features two main functionalities: classification of the etiology of data from a single patient filled in by a form, and batch classification for a list of many patients. The application achieved satisfactory performance, good responsiveness on the tested devices, and works on the iOS and Android platforms.

Conclusions: We developed and validated a mobile application that makes use of Machine Learning methods to predict the etiology of cerebral intraparenchymal hemorrhage in patients using the SMASH-U etiological classification with good accuracy and usability.

The application has features that can help etiological investigation in clinical practice, as well as facilitate the labeling of large amounts of data for research.

KEYWORDS: Intraparenchymal hemorrhage. Etiological investigation. Mobile apps. Artificial intelligence. Semi-supervised learning.

LISTA DE ABREVIACOES

AM - Aprendizado de Mquina

AUC - rea sob a curva ROC (do ingls “*area under the ROC curve*”)

AVC - Acidente vascular cerebral

AVCI - Acidente vascular cerebral isqumico

AVCH - Acidente vascular cerebral hemorrgico

HIC - Hemorragia intraparenquimatosa cerebral

IA - Inteligncia Artificial

FA - Florestas Aleatrias

MVS - Mquina de Vetor de Suporte

REAVR - Registro de Acidente Vascular Enceflico

RNR - Redes Neurais Recorrentes

TC - Tomografia Computadorizada

TP - Tempo de Protrombina

TTPA - Tempo de Tromboplastina Parcial Ativada

LISTA DE FIGURAS

Figura 1 - Esquema de classificação SMASH-U utilizado para classificar a HIC nas seguintes etiologias: lesão estrutural, medicação, angiopatia amilóide, doença sistêmica/outras, hipertensão e indeterminado.

Figura 2 - Esquema de uma validação cruzada *k-fold* com $k=5$. Os quadrados azuis representam as subdivisões (*fold*) usadas para teste e os cinzas, as subdivisões usadas para treino.

Figura 3 - Arquitetura de uma rede com 1 camadas de entrada (com 5 nós), 1 camada oculta (com 3 nós) e uma 1 camada de saída (com 2 nós)

Figura 4 - Processo de inclusão de artigos na revisão bibliográfica.

Figura 5 - Fluxo de informações entre o aplicativo e o servidor.

Figura 6 - Os 18 pipelines construídos pela combinação de 3 modelos de rotulação, 2 algoritmos de seleção de variáveis preditoras, e 3 modelos de classificação.

Figura 7 - Preparação e divisão de dados e construção e comparação dos pipelines para obtenção do melhor modelo.

Figura 8 - Interface gráfica do aplicativo em diferentes tamanhos de telas.

Figura 10 - Telas para a predição de dados de um único paciente.

Figura 11 - Tela de resultados da predição de dados de um único paciente.

Figura 12 - Tela da predição de grande quantidade de dados (*Batch Prediction*).

Figura 13- Planilha com resultados para a predição de grande quantidade de dados.

Figura 14 - Instruções para utilização da funcionalidade de predição de grande quantidade de dados.

Figura 15 - Tela de Cadastro de Usuário, de Login e tela Inicial (*Home*).

LISTA DE TABELAS

Tabela 1 - Esquema comparativo entre as etiologias SMASH-U e H-TOMIC.

Tabela 2 - Quantidade de dados faltantes em cada variável estratificada pela variável etiologia.

Tabela 3 - Frequência das diferentes localizações onde a hemorragia foi identificada.

Tabela 4 - Frequência dos valores da escala neurológica ICH.

Tabela 5 - Frequência dos valores da escala neurológica SICH.

Tabela 6 - Atributos selecionados pelos algoritmos χ^2 e RL.

Tabela 7 - Atributos com seus valores respectivos de χ^2 e o *ranking* (para o algoritmo de RL).

Tabela 8 - Resultados dos pipelines que usaram MVS na primeira etapa.

Tabela 9 - Resultados dos pipelines que usaram RNA na primeira etapa.

Tabela 10 - Resultados dos pipelines que usaram RF na primeira etapa.

Tabela 11 - Emuladores e aparelhos usados para teste, com as respectivas versões do sistema operacional neles instalados no sistema operacional Android.

Tabela 12 - Emuladores e aparelhos usados para teste, com as respectivas versões do sistema operacional neles instalados no sistema operacional iOS.

Tabela 13 - Tabela comparativa com Meretoja et al (2012) entre as estatísticas descritivas agrupadas por etiologia para os atributos em comum.

SUMÁRIO

1 - Introdução	16
1.1 - Contextualização	16
1.2 - Motivação	17
1.5 - Conceitos Fundamentais: Acidente Vascular Cerebral	18
1.5.1 - Acidente Vascular Cerebral	18
1.5.2 - Hemorragia Intraparenquimatosa Cerebral	18
1.5.3 - Classificação Etiológica da Hemorragia Intraparenquimatosa Cerebral	20
1.5.4 - Variáveis associadas à Hemorragia Intraparenquimatosa Cerebral	22
1.6 - Conceitos Fundamentais: Aplicações Móveis	24
1.6.1 - Aplicações Móveis	24
1.6.2 - Aplicações Móveis no setor da Saúde	25
1.6.3 - Desenvolvimento de Aplicações Móveis	26
1.7 - Conceitos Fundamentais: Inteligência Artificial	27
1.7.1 - Inteligência Artificial e Aprendizado de Máquina (AM)	27
1.7.2 - Tipos de Aprendizado de Máquina	28
1.7.3 - Métricas de Avaliação	30
1.7.4 - Balanceamento	31
1.7.5 - Seleção das variáveis preditoras	31
1.7.6 - Data Splitting e Validação Cruzada	32
1.7.7 - Algoritmos de Aprendizado de Máquina	33
1.7.8 - Pipelines de Aprendizado de Máquina	38
2 - Revisão Bibliográfica	39
2.1 - Estudos na literatura que utilizam métodos de AM em dados clínicos na predição de diagnósticos/condições médicas	39
3 - Objetivos e Hipótese	41
3.1 - Objetivos	41
3.2 - Hipótese	41
4 - Materiais e Métodos	42
4.1 - Delineamento do estudo	42
4.2 - Desenvolvimento do aplicativo	42
4.2.1 - Ferramentas de desenvolvimento	42

4.2.2 - A arquitetura do sistema	42
4.3 - O modelo de predição	43
4.3.1 - Data-Splitting, Pré-processamento, Tratamento de Dados Faltantes e Balanceamento de Classes	43
4.3.2 - Rotulação de Classes Faltantes	44
4.3.3 - Seleção de Variáveis Predictoras	45
4.3.4 - Classificação de Etiologias	45
4.5 - Validação Externa	47
4.6 - Variáveis predictoras	47
4.6.1 - Variáveis demográficas	47
4.6.2 - Variáveis clínicas	48
4.6.3 - Variáveis neuroimagem	48
4.6.4 - Variável de interesse - etiologia segundo a classificação SMASH-U	49
4.7 - Dados Utilizados : Registro de Acidente Vascular Encefálico (REAYER)	49
4.7.1 - Amostra	50
4.7.2 - Critérios de Inclusão e Exclusão	50
Foram utilizados os seguintes critérios de inclusão dos pacientes no REAYER:	50
5 - Resultados e Discussão	51
5.1 - Resultados Descritivos	51
5.1.1 - Estatísticas descritivas	51
5.2 - Comparação de pipelines	55
5.2.1 - Rotulação de etiologias faltantes com método semi supervisionado	55
5.2.2 - Seleção de atributos	56
5.2.3 - Classificação de etiologias	58
5.3 - Resultados da validação externa	60
5.4 - Resultados da criação do aplicativo	60
5.4.1 - Interface Gráfica	60
5.4.2 - Funcionalidades	62
5.5 - Discussão	69
5.6 - Sugestões para trabalhos futuros	73
6 - Conclusões	74
7 - Referências	75
8 - Anexos	82

1 - Introdução

1.1 - Contextualização

Em 2021, a Organização Mundial de Saúde publicou um relatório com orientações para o uso de técnicas de Inteligência Artificial (IA) na Saúde, no qual se afirma que a IA é uma ferramenta importante para melhorar a velocidade e a precisão do diagnóstico e da triagem de doenças, e para dar suporte ao atendimento clínico.

"Use of AI technologies for health holds great promise and has already contributed to important advances in fields such as drug discovery, genomics, radiology, pathology and prevention. AI could assist health-care providers in avoiding errors and allow clinicians to focus on providing care and solving complex cases. The potential benefits of these technologies and the economic and commercial potential of AI for health care presage ever greater use of AI worldwide.

...

While AI may not replace clinical decision-making, it could improve decisions made by clinicians. In settings with limited resources, AI could be used to conduct screening and evaluation if insufficient medical expertise is available, a common challenge in many resource-poor settings."

"O uso de tecnologias de IA para a saúde é uma grande promessa e já contribuiu para avanços importantes em áreas como descoberta de medicamentos, genômica, radiologia, patologia e prevenção. A IA pode ajudar os profissionais de saúde a evitar erros e permitir que os médicos se concentrem na prestação de cuidados e na resolução de casos complexos. Os benefícios potenciais dessas tecnologias e o potencial econômico e comercial da IA para os cuidados de saúde pressupõem um uso cada vez maior da IA em todo o mundo.

...

Embora a IA não substitua a tomada de decisões clínicas, ela pode melhorar as decisões tomadas pelos médicos. Em ambientes com recursos limitados, a IA pode ser usada para

realizar triagem e avaliação se houver conhecimento médico insuficiente, um desafio comum em muitos ambientes com poucos recursos.”

(Orientações da OMS Para a Ética e Governança de Inteligência Artificial para Saúde)

O aumento do uso de aparelhos celulares favoreceu a integração de modelos de IA às aplicações móveis no contexto da saúde. Unir a praticidade e portabilidade de tablets e celulares à capacidade preditiva da IA na área médica é uma forma de aumentar a acessibilidade à informação clínica e a serviços médicos de qualidade.

Dessa forma, este projeto de pesquisa propõe o desenvolvimento de um aplicativo de celular que auxilie na investigação etiológica de um episódio de hemorragia intraparenquimatosa cerebral (HIC), através da predição da etiologia por um algoritmo de Aprendizado de Máquina (AM).

1.2 - Motivação

Existem mais de 150 causas conhecidas para um episódio de acidente vascular cerebral (AVC). Identificar a etiologia, isto é, a causa de um episódio de AVC é fundamental como tratamento secundário, podendo reduzir em até 90% o risco de recorrência. Além disso, auxilia a estabelecer grupos de risco, escolher testes diagnósticos e melhorar o manejo do paciente. [1]

Determinar a etiologia também é importante para estratificar grupos de interesse em pesquisas clínicas e epidemiológicas.

Em 25-39% dos pacientes com AVC, os médicos falharam em identificar uma causa definida, dado às variações de qualidade, integridade e rapidez da investigação etiológica [1]. Neste contexto, uma ferramenta de classificação automatizada poderia ser uma forma de otimizar a investigação etiológica.

Considerando a importância de determinar a etiologia de um episódio de AVC, a classificação automática da etiologia da HIC, o subtipo do AVC de pior prognóstico, é uma maneira de otimizar a investigação etiológica desta doença. Nesse contexto, um aplicativo móvel de predição da etiologia da HIC é uma forma de agilizar o processamento de grandes quantidades de dados, além de oferecer uma maneira acessível de apoio à investigação etiológica.

1.5 - Conceitos Fundamentais: Acidente Vascular Cerebral

1.5.1 - Acidente Vascular Cerebral

Segundo a Organização Mundial de Saúde (OMS), o AVC pode ser descrito como o desenvolvimento rápido de sintomas de um distúrbio focal da(s) função(ões) cerebral(is), com duração superior a 24 horas ou óbito, sem outro motivo aparente além de fatores vasculares. Em 2005, o AVC foi responsável por quase 90.006 mortes no Brasil [2,3] .

O AVC pode ser do subtipo isquêmico (AVCI) ou hemorrágico (AVCH). O AVCH ocorre quando há ruptura espontânea (não traumática) de um vaso com conseqüente extravasamento de sangue para o interior do cérebro (hemorragia intraparenquimatosa cerebral), para o sistema ventricular (hemorragia intraventricular) e/ou espaço subaracnóideo (hemorragia subaracnóide).

Quanto à etiologia, o AVC pode ter mais de 150 causas conhecidas, sendo considerada uma doença heterogênea [1]. O AVCI, por exemplo, é causado em até 50% dos casos por doença aterotrombótica de grandes artérias, cerca de 25% por doença das pequenas artérias intracranianas (resultando em AVC lacunar), até 20% por cardioembolismo e 5% por causas raras diversas [4]. Resta ainda uma porcentagem de 10-40%, em diferentes casuísticas, de AVCI de etiologia indeterminada.

1.5.2 - Hemorragia Intraparenquimatosa Cerebral

A hemorragia intraparenquimatosa cerebral (HIC) é causada pela ruptura de um vaso sanguíneo na intimidade do encéfalo e é a de pior prognóstico dentre os subtipos de AVCH, sendo responsável por 10–20% dos episódios dessa doença. A letalidade é de 40% em 1 mês, e de 54% em um ano [5,6]. A manifestação clínica da HIC inclui o surgimento abrupto de um déficit neurológico focal (hemiparesia, hipoestesia unilateral, hemianopsia, afasia, etc.) de rápido início e progressão. Embora os sintomas dependam da região cerebral acometida, em geral os pacientes com HIC podem apresentar aumento da pressão intracraniana, vômito e cefaléia (principalmente quando há hemorragia subaracnóide associada).

O exame de neuroimagem, geralmente a tomografia computadorizada (TC) de crânio, é fundamental para a confirmação do diagnóstico de HIC [7,8]. A TC de crânio permite a

diferenciação entre HIC lobares, profundas e de fossa posterior. A etiologia do sangramento também pode estar relacionada a achados radiográficos na TC. Por exemplo: quando a hemorragia é nos núcleos da base, a etiologia mais comum é a hipertensiva; quando existe um ou mais hematomas lobares associados à leucoaraiose, a etiologia mais provável é a angiopatia amilóide; quando há sangue no espaço subaracnóideo, a etiologia mais associada é a ruptura de um aneurisma; quando há fluido dentro do hematoma, pode haver uma coagulopatia. [9].

O principal fator de prognóstico de pacientes com HIC é o volume da hemorragia, geralmente estimado pelo método ABC/2 [10,11]. De forma geral, os pacientes com HIC devem ser avaliados através da escala de coma de Glasgow e a escala de AVC do *National Institutes of Health* (NIH) [12,13].

Episódios de HIC possuem elevadas taxas de mortalidade e incapacidade. Do total de pacientes com HIC, é estimado que 35 a 52% evoluam para o óbito ao final do primeiro mês. Dentre os fatores atribuídos ao prognóstico ruim para os pacientes com HIC estão sangramento intraventricular, o volume inicial da hemorragia maior que 30 cm³, idade avançada, localização primária infratentorial, rebaixamento do nível de consciência na admissão [14,15], crescimento precoce do hematoma [16] e uso de trombolíticos [17]. A escala ICH, estratifica o risco de mortalidade em 30 dias, é a mais utilizada para predizer o prognóstico de um episódio de HIC.

Quanto ao prognóstico a longo prazo, a mortalidade da HIC está entre 60 a 80% em até dois anos após o episódio. Apenas 20% dos pacientes recuperam a independência funcional em seis meses [18]. Os preditores de mortalidade a longo prazo são: idade avançada, rebaixamento do nível de consciência na admissão, localização e volume do sangramento e gravidade do déficit neurológico [19].

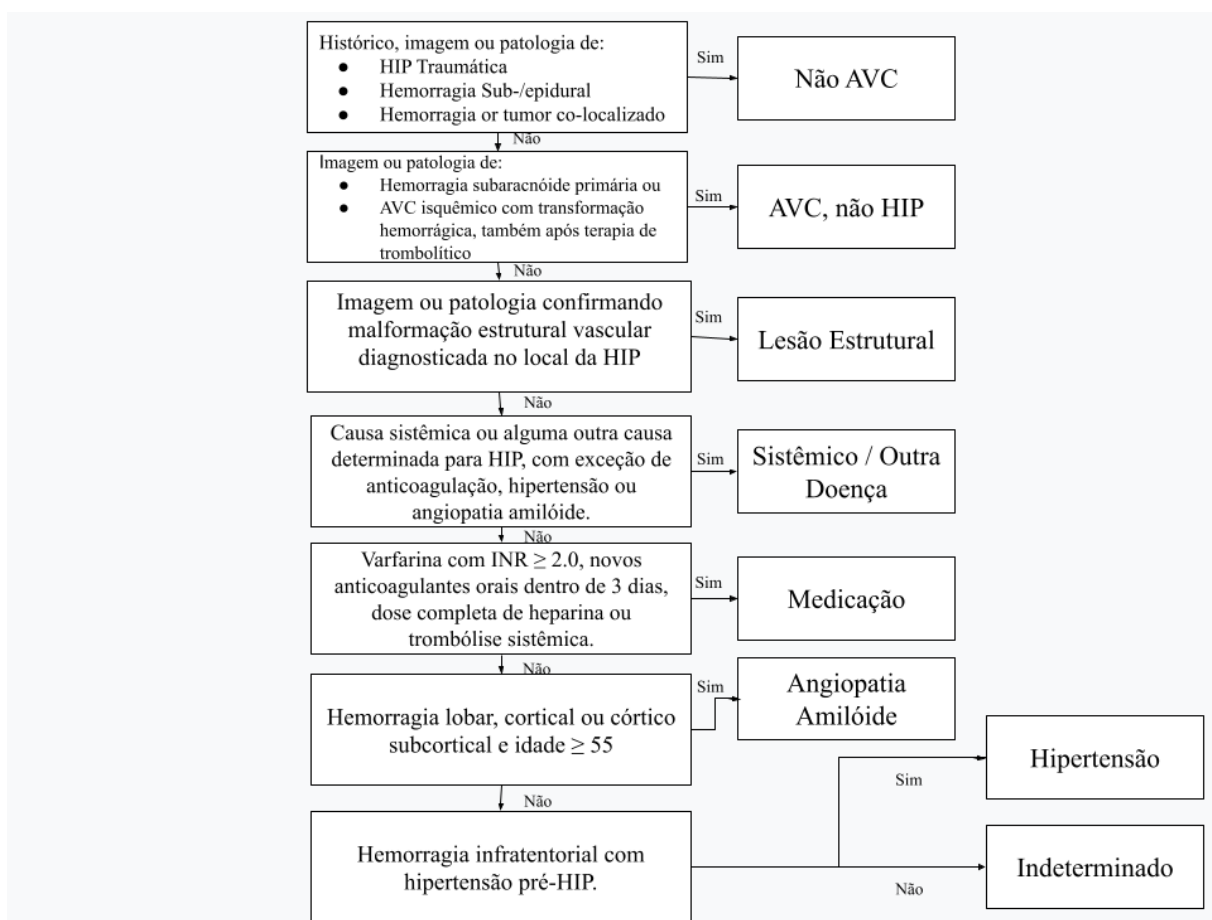
Os tratamentos específicos para HIC são ainda bastante limitados, e o manejo guarda ainda similaridades aos demais subtipos de AVC. Contudo, os pacientes com HIC devem ser rapidamente encaminhados para leitos monitorizados em uma unidade de AVC ou em leitos de terapia intensiva dada a gravidade da doença. Como existem poucas intervenções eficazes para a HIC, o correto manejo de pacientes é essencial. Neste contexto, destacam-se as intervenções para controle pressórico rigoroso, a reversão de qualquer coagulopatia e a cirurgia minimamente invasiva.

1.5.3 - Classificação Etiológica da Hemorragia Intraparenquimatosa Cerebral

A etiologia da HIC pode ser classificada em primária ou secundária. A HIC primária, que engloba de 80 a 85% dos casos, está associada à ruptura de pequenos vasos cronicamente danificados pela hipertensão ou à angiopatia amilóide [20]. Por outro lado, a HIC secundária representa de 15 a 25% dos casos e é ocasionada pela ruptura de aneurismas ou malformações arteriovenosas cerebrais, pela anticoagulação oral e coagulopatias, drogas antiplaquetárias, cirrose hepática, vasculites, neoplasias, eventos de trauma, doença de Moya-Moya, trombose venosa cerebral, eclâmpsia, entre demais causas. Identificar a etiologia é fundamental para manejo do paciente, a escolha de testes diagnósticos, para a previsão do prognóstico e para a prevenção de um novo episódio (medidas secundárias), além de facilitar a comparação de populações de pacientes de diferentes grupos e a padronização na pesquisa [1].

As possíveis etiologias para o AVC hemorrágico são diversas, o que pode tornar o processo de investigação etiológica ainda mais trabalhoso. Meretoja et al (2012) desenvolveram uma classificação clínica para a etiologia da HIC denominada SMASH-U, que se mostrou viável e obteve concordância entre avaliadores de 92 dos 100 pacientes, avaliados de forma idêntica por 2 neurologistas. Nesse estudo retrospectivo, que utilizou dados de 1013 pacientes do Helsinki ICH Study, foi possível determinar uma etiologia plausível na maioria dos pacientes, embora a etiologia permaneça indefinida em 21% dos casos. Lesões estruturais (n = 50), doença sistêmica (n = 48) e anticoagulação (n = 143) somaram 25% dos casos. Os episódios de HIC devido à angiopatia amilóide (n = 205) e à hipertensão (n = 354) somaram 20% e 30%. A classificação baseada no SMASH-U estava fortemente associada à sobrevida, o que ressalta a importância de determinar a etiologia para estabelecer grupos de risco. A SMASH-U classifica um episódio de HIC nas seguintes etiologias: lesão estrutural, medicação, angiopatia amilóide, doença sistêmica / outra, hipertensão e indeterminado [21]. Na Figura 1, encontra-se o esquema de classificação usado por SMASH-U.

Figura 1 - Esquema de classificação SMASH-U utilizado para classificar a HIC nas seguintes etiologias: lesão estrutural, medicação, angiopatia amilóide, doença sistêmica/outras, hipertensão e indeterminado.



Fonte: Elaborada pelo autor com base em Meretoja et al, 2012

Outra classificação proposta para determinar a etiologia de um episódio de HIC é a classificação H-ATOMIC, que de forma simplificada define 8 categorias: hipertensão, angiopatia amilóide, tumor, anticoagulantes orais, malformação, infrequente, criptogênica e combinação.

Em uma comparação direta e aleatória entre H-ATOMIC e SMASH-U, Martí-Fàbregas et al (2016) obtiveram boa confiabilidade para ambos os sistemas de classificação. Neste estudo, neurologistas aplicaram ambos os sistemas a uma lista de pacientes consecutivos de um banco de dados prospectivo. Os pacientes a serem avaliados foram designados aleatoriamente. Foram calculados o percentual de concordância e a confiabilidade inter-examinador (coeficiente kappa). O estudo, que incluiu 156 pacientes (idade 72.3 ± 13.5 anos), obteve a porcentagem de concordância entre os investigadores e a confiabilidade entre avaliadores alta para ambos os sistemas de classificação. Contudo, discrepâncias entre os 2 sistemas ocorrem em cerca de um terço dos pacientes por causa de

diferentes categorias e definição [22]. Na Tabela 1, encontra-se um esquema comparativo entre as classificações.

Tabela 1- Esquema comparativo entre as etiologias SMASH-U e H-TOMIC.

	SMASH-U	H-ATOMIC	Comentários
Categorias iguais	Hipertensão Angiopatia Amilóide Estrutural Medicamento Indeterminado	Hipertensão Amilóide Malformação Anticoagulação Oral Criptogênico	SMASH-U requer que a hipertensão seja conhecida antes da HIC, mas H-ATOMIC permite o diagnóstico de hipertensão após HIC.
Categorias diferentes	Sem AVC AVC não ICH Sistêmico/Outra doença	Tumor Combinação Pouco frequente	SMASH-U inclui como etiologia a trombólise e a heparina, que são considerados na categoria infrequente em H-ATOMIC.
Graus de certeza	Não	Sim	As hemorragias cerebelares primárias são consideradas hipertensivas na SMASH-U, mas pode ser hipertenso ou relacionado à amilóide em H-ATÔMICO.
Permite mais de uma categoria	Não	Sim	O tumor não é considerado um acidente vascular cerebral em SMASH-U
Plano diagnóstico pré-determinado	Não	Sim	O AVC isquêmico com transformação hemorrágica é um AVC-não-HIC no SMASH-U e não existe no H-ATOMIC. A categoria de doença sistêmica ou outra não está definida no SMASH-U mas mas é detalhada em H-ATOMIC (categoria pouco frequente) A categoria de combinação não é permitida em SMASH-U

Fonte: Elaborada pelo autor com base em Martí-Fàbregas et al , 2016.

Neste contexto, o desenvolvimento de uma aplicação móvel para classificação da etiologia da HIC é uma forma de otimizar a investigação etiológica.

1.5.4 - Variáveis associadas à Hemorragia Intraparenquimatosa Cerebral

Idade avançada, raças negra e orientais e gênero masculino são os fatores de risco não modificáveis mais conhecidos para HIC [23]. O uso do cigarro também é um fator de risco modificável. Em fumantes, o risco é cerca de duas e meia vezes maior do que em indivíduos

que não fumam. Quanto ao álcool, diversos estudos indicam que o seu consumo é um dos principais fatores de risco para HIC [24], sendo que um episódio de sangramento intracraniano espontâneo também parece ser precipitado pelo consumo de quantidades moderadas ou grandes de álcool nas 24 horas que antecedem ao ictus [25]. Além dos fatores já citados, o uso de fármacos ou drogas com atividade simpaticomimética, tais como fenilpropanolamina, cocaína, anfetaminas e/ou efedrina, aumentam o risco de HIC [26].

A hipertensão arterial sistêmica (HAS) é o fator de risco mais comumente associado à HIC, estando presente em 70 a 80% dos pacientes com este tipo de AVC [23]. A ruptura de pequenos vasos cronicamente danificados pela HAS está intrinsecamente associada à etiologia hipertensiva de causa primária da HIC.

Outro fator relacionado à etiologia de causa primária é a angiopatia amilóide, condição que ocorre devido ao depósito de proteína beta-amilóide na parede das artérias cerebrais de pequeno e médio calibre localizadas principalmente na superfície cortical e leptomeníngea. A angiopatia amilóide é um fator de risco para HIC lobar, especialmente nos lobos parietal e occipital, ocorrendo particularmente nos pacientes com idade superior a setenta anos.

As coagulopatias estão associadas à etiologia secundária da HIC. A necessidade crescente do uso de antiagregantes e anticoagulantes orais para prevenir eventos aterotrombóticos e tromboembólicos, tem aumentado o número de casos de HIC secundárias.

O uso de anticoagulantes orais aumenta o risco de HIC em torno de 8 a 10 vezes em relação a pacientes de mesma idade não submetidos à anticoagulação [27]. O uso prévio de drogas antiplaquetárias também está associado, de forma independente, ao aumento do risco de HIC e ao crescimento do hematoma intracerebral [28].

Quanto à diabetes, ela afeta múltiplos sistemas orgânicos, podendo levar à complicações macrovasculares, que incluem acidente vascular cerebral. As consequências da hiperglicemia nas complicações cardiovasculares em pessoas com diabetes ainda não está totalmente claro, mas diversos estudos apontam que existe uma relação.

As vasculopatias, entre elas as vasculites sistêmica, vasculite isolada do sistema nervoso central e demais vasculites (sarcoidose, doença de Behçet, doença de Moya-Moya, dissecação arterial, vasculite infecciosa e anemia falciforme) são associadas à etiologia secundária da HIC sendo inclusive algumas das possíveis causas desse evento hemorrágico.

A presença de aneurismas também foi incluída como um preditor porque a ruptura de aneurismas são causas secundárias para a ocorrência de HIC.

A investigação etiológica da HIC deve ser orientada pela localização da lesão no exame de imagem. A TC de crânio permite a diferenciação entre três grandes grupos de HIC: lobares, profundas e de fossa posterior. A localização de alguns achados tomográficos indicam a etiologia do sangramento: hematoma hipertensivo, por exemplo, é o diagnóstico mais provável quando a hemorragia está localizada nos núcleos da base, a angiopatia amilóide é mais provável quando existe um ou mais hematomas lobares associados à leucoaraiose, ruptura de aneurisma é sugerida pela presença de sangue no espaço subaracnóideo, uma HIC por coagulopatia pode se acompanhar de níveis de fluidos dentro do hematoma e em hemorragias traumáticas costumam ser encontradas contusões e fraturas ósseas associadas [8].

As escalas SICH e ICH têm mostrado sucesso ao predizer o risco de um determinado paciente com HIC de possuir uma determinada etiologia vascular, podendo orientar a seleção de pacientes para a avaliação neurovascular e para excluir a presença de uma anormalidade vascular. Os volumes de sangramento dos episódios de HIC são normalmente utilizados como estratificação de risco e prognóstico dos pacientes.

1.6 - Conceitos Fundamentais: Aplicações Móveis

1.6.1 - Aplicações Móveis

Aplicações móveis são softwares desenvolvidos para operarem em dispositivos eletrônicos móveis, como os smartphones. Elas surgiram em resposta ao mercado emergente de comercialização de smartphones e ganharam popularidade no ano de 2008 através do lançamento pela Apple de uma loja online de venda de aplicativos (App Store). Após isso, outras lojas foram lançadas com o mesmo princípio de atender seus usuários e oferecer diversas funcionalidades pelos aplicativos móveis, tais como Google Play e BlackBerry World.

A ideia de diversificar as funções de celulares já estava presente na década de 80. Em 1983 aconteceu um evento de tecnologia em Aspen, Colorado (EUA), no qual Steve Jobs discursou sua intenção de desenvolver softwares que seriam baixados nos celulares via ondas de rádio. Antes do sucesso atual dos aplicativos, muitas empresas apresentaram tentativas de aumentar o poder computacional dos dispositivos móveis e de oferecer mais funcionalidades para esses dispositivos. Por exemplo, entre 1996 e 2009, o Palm OS foi o sistema operacional

presente em palmtops. Ele oferecia aplicativos básicos, inclusive reproduzidor de áudio e vídeo do famoso RealPlayer. A empresa que o desenvolveu, contudo, falhou ao ingressar no mercado de smartphones e faliu em 2011. Apesar do sistema operacional não ser mais utilizado atualmente, a interface utilizada inspirou empresas atuais a modificarem a interface de seus *softwares*. Outro exemplo de precursor dos atuais smartphones é o aparelho Nokia N95. Possuía recursos considerados avançados para a época, como a navegação em HTML, reproduzidor de áudio e uma rede social chamada "Nokia Life Blog".

Atualmente, os sistemas operacionais mais comuns nos dispositivos móveis são o Android, da empresa Google e iOS, da empresa Apple. A grande maioria dos smartphones que não utilizam iOS, possuem o sistema Android instalado.

A facilidade de publicar e utilizar os aplicativos nestas plataformas é um incentivo para o seu desenvolvimento e uso. A diversidade e a usabilidade destes aplicativos os tornaram populares entre todos os públicos, inclusive entre aqueles que não tinham o hábito de utilizar smartphones. Dessa forma, os aplicativos alteraram a forma das pessoas trabalharem e se comunicarem e se tornaram de grande importância no cotidiano [29].

1.6.2 - Aplicações Móveis no setor da Saúde

O advento de dispositivos de computação móvel e a sua popularização afetaram a forma como as pessoas realizam suas atividades no trabalho, inclusive no campo da medicina. Smartphones e tablets combinam recursos de computação e comunicação em um único dispositivo portátil, permitindo fácil acesso e uso no ponto de atendimento. Devido a sua facilidade e praticidade, os profissionais de saúde agora usam smartphones ou tablets para funções que anteriormente eram realizadas com uso de pagers e telefones [30]. Os dispositivos móveis tornaram-se comuns em ambientes de cuidados de saúde, levando a um rápido crescimento no desenvolvimento de aplicativos de software médicos para essas plataformas.

Melhorar a qualidade dos cuidados de saúde, reduzir os custos e facilitar o trabalho dos profissionais da saúde são motivos pelos quais os aplicativos de dispositivos móveis têm sido introduzidos em clínicas e hospitais.

O uso de apps por profissionais de saúde transformou muitos aspectos da prática clínica [31]. Vários aplicativos estão disponíveis para ajudar esses profissionais no exercício de sua profissão, como por exemplo nas tarefas de gerenciamento de informações ao longo do

tempo, manutenção e acesso a registros de saúde, monitoramento de pacientes, consultoria, tomada de decisão clínica, treinamentos e educação médica [32].

1.6.3 - Desenvolvimento de Aplicações Móveis

Um grande problema no desenvolvimento de aplicativos móveis é que o mercado móvel é dividido entre várias plataformas. Atualmente, são duas as mais utilizadas: Android, pertencente à empresa Google e iOS, pertencente à empresa Apple. O desenvolvimento nativo para Android ocorre na linguagem programação Java enquanto o desenvolvimento para iOS é Objective-C ou Swift [33,34].

Uma alternativa para o desenvolvimento nativo, que demanda conhecimento em diferentes linguagens de programação e possui diferentes ambientes de compilação, é o desenvolvimento de plataforma cruzada. Este permite desenvolver um aplicativo para várias plataformas ao mesmo tempo.

Alguns dos *frameworks* usados para o desenvolvimento de plataforma cruzada são [35]:

- a) Ionic (IONIC, 2019);
- b) React Native (FACEBOOK INC., 2019);
- c) Xamarin (MICROSOFT, 2019);
- d) Adobe PhoneGap (ADOBE SYSTEMS INC, 2016);
- e) Flutter (GOOGLE, 2019).

Dentre eles, o React Native (RN) tem conquistado espaço em empresas e startups de tecnologia. RN é uma estrutura de *script* nativa usada para criar aplicativos móveis de plataforma cruzada, introduzidos pela primeira vez durante a conferência React.js em 2015. O RN baseia-se nos princípios e conceitos do ReactJS [36], que é uma estrutura Javascript de código aberto. RN proporciona maior facilidade de desenvolvimento, e aplicativos desenvolvidos com esse framework oferecem experiências de usuário semelhantes às das versões nativas, sem afetar significativamente o desempenho [37].

A intensificação do desenvolvimento e grande difusão de aplicativos de hardware e software (incluindo os aplicativos móveis) facilitou a digitalização de dados, o que favoreceu a utilização de modelos computacionais para o uso de sistemas de IA. A IA viabiliza o processamento e a análise de grandes quantidades de dados, sendo uma ferramenta de grande potencial para a área da saúde.

1.7 - Conceitos Fundamentais: Inteligência Artificial

1.7.1 - Inteligência Artificial e Aprendizado de Máquina (AM)

De forma simplificada, o termo Inteligência Artificial (IA) refere-se a sistemas ou máquinas que imitam a inteligência humana para realizar tarefas e que se aprimoram de forma iterativa a partir de dados fornecidos. [38]

A primeira vez em que um artigo foi publicado em uma área correlata à IA foi em 1943. Neste ano, Warren McCulloch e Walter Pitts escreveram o artigo que deu início ao conceito de Redes Neurais. O conceito de IA também foi mencionado posteriormente por Alan Turing no seu artigo de 1950 “Computing Machinery and Intelligency”, no qual a ideia do Teste de Turing é introduzida. Esse teste sugere a possível dificuldade de distinguir entre um ser humano e uma máquina inteligente. O computador passaria no teste caso uma pessoa não conseguisse distingui-lo de um humano após interrogá-lo.

O período entre 1952 e 1969 foi marcado pelo entusiasmo em relação à área. Em 1956, John McCarthy, Hyman Minsky, Claude Shannon e Nathaniel Rochester organizaram um seminário em Dartmouth, no qual foi apresentado o programa de raciocínio Logic Theorist (LT), por Allen Newell e Herbert Simon. O termo “Inteligência Artificial” foi utilizado para se referir a um novo campo do conhecimento pela primeira vez neste seminário.

Em 1969, o programa DENDRAL foi desenvolvido pela Universidade de Stanford para fornecer soluções capazes de encontrar as estruturas moleculares orgânicas a partir da espectrometria de massa de ligações químicas de uma molécula, fazendo uso de métodos de IA.

A partir da década de 80, a IA começa a ser vista como uma indústria. O R1, primeiro sistema de IA bem-sucedido comercialmente, começou a operar na Digital Equipment Corporation (DEC), auxiliando na configuração de pedidos de novos sistemas de computador. Em 1988, operavam 40 sistemas especialistas do grupo de IA da DEC (CHARNIAK; MCDERMOTT, 1985). A empresa Du Pont já possuía 100 desses sistemas em uso e estava desenvolvendo mais 500. Praticamente todos os grandes grupos empresariais dos Estados Unidos possuíam ou estavam desenvolvendo tecnologias em IA. O Japão iniciou o Fifth Generation em 1981, uma meta de 10 anos para montar computadores inteligentes com a linguagem Prolog. Também nesta década, os Estados Unidos fundaram a Microelectronics

and Computer Technology Corporation (MCC), que também fazia uso de abordagens de IA [39].

No século 21, a IA permanece a ser utilizada em tarefas de diversas áreas devido à capacidade de automatizar tarefas que atualmente requerem intervenção humana. Métodos de IA são utilizados, por exemplo, para auxiliar a visão computacional, reconhecer fala e processar linguagem natural.

As tecnologias usadas na IA podem simular inteligência humana em vários níveis. Aprendizado de máquina (AM) e aprendizado profundo são subconjuntos de IA. Métodos de AM permitem aos sistemas aprender com os dados [40].

O problema de aprendizagem por um algoritmo supervisionado de AM pode ser definido como a sua melhora (medida por alguma métrica de desempenho) ao executar alguma tarefa, através do treino. Por exemplo, em um problema de classificação supervisionada de tumores malignos e benignos, o modelo deve atribuir um rótulo de “maligno” ou “benigno” aos tumores. Uma possível métrica de desempenho a ser utilizada para mostrar o desempenho do modelo seria a acurácia, ou seja, o número de rótulos (maligno ou benigno) dados pelo modelo que estão corretos. O treinamento consistiria em um conjunto de dados rotulados como maligno ou benigno, sendo cada instância referente a um tumor. O aprendizado do modelo, neste caso, é observado pelo aumento da acurácia, à medida que os dados de treino são fornecidos [41].

AM encontra-se na interseção das áreas de ciência da computação e estatística. O atual crescimento foi impulsionado principalmente pela contínua disponibilidade de dados online e o aumento da computação de baixo custo [41]. Atualmente, os algoritmos de AM são empregados com sucesso em tarefas de classificação, regressão, agrupamento e redução de dimensionalidade de grandes conjuntos de dados de entrada especialmente de alta dimensão [42].

1.7.2 - Tipos de Aprendizado de Máquina

AM pode ser dividido em quatro categorias principais: aprendizado supervisionado, aprendizado semi-supervisionado, aprendizado não supervisionado e aprendizado por reforço [43].

Segundo Friedman et al. (2001), no aprendizado não supervisionado a resposta não é uma saída mensurável, mas associações ou agrupamentos dos dados de entrada. Um modelo

de aprendizado não supervisionado não classifica as instâncias, mas os agrupa segundo um critério pré-estabelecido ou faz uma redução de dimensionalidade [44].

Já os métodos supervisionados tentam encontrar a função desconhecida que conecta entradas conhecidas a saídas desconhecidas. Este resultado desejado para domínios desconhecidos é estimado com base na extrapolação de padrões encontrados nos dados de treinamento rotulados.

Na classificação supervisionada separa-se um subconjunto da população relevante para treinamento para a qual os valores da propriedade de destino (rótulos) são conhecidos. Em seguida, é necessário decidir como adequar as propriedades do sistema para o algoritmo escolhido (traduzir as informações brutas em certos recursos que serão usados como entradas para o algoritmo). Terminado este processo, o modelo é treinado otimizando seu desempenho, geralmente medido através de algum tipo de função de custo. Isso envolve o ajuste de hiperparâmetros que controlam o processo de treinamento, a estrutura e as propriedades do modelo. Os dados são divididos em um conjunto de dados de validação, separado dos conjuntos de teste e treinamento. A otimização de hiperparâmetros ajuda a controlar a performance do modelo e evitar *over/underfitting*. *Underfitting* ocorre quando o modelo não tem a capacidade de expressar a complexidade dos dados, recursos, ou não descreve adequadamente os dados. Isso inevitavelmente leva a um alto erro de treinamento.

Por outro lado, um *overfitting* interpreta parte do ruído nos dados de treinamento como informação relevante, não conseguindo prever com confiabilidade novos dados. Geralmente, um modelo de *overfitting* contém mais parâmetros livres do que o número necessário. Para evitar o *overfitting*, é essencial monitorar durante o treinamento, o erro de treinamento e o erro do conjunto de validação [42].

O aprendizado semi-supervisionado faz uso tanto de dados rotulados quanto de dados não rotulados. Na aprendizagem semi-supervisionada, parte dos dados que contém rótulos é usada para treinar um algoritmo supervisionado. Para isso, temos que alavancar instâncias não rotuladas para reestruturar hipóteses durante o processo de aprendizagem.

Um algoritmo de maximização de expectativa é usado para automatizar o processo de anotação para todo o conjunto de dados não rotulados. Existe vantagem no uso de modelos semi-supervisionados porque o processo de rotulação pelo ser humano demanda muito recursos e necessita muitas vezes de um profissional especializado para determinar os rótulos, o que torna o processo mais custoso e demorado [43].

Pesquisadores verificaram que dados não rotulados usados em conjunto com uma pequena quantidade de dados rotulados, podem resultar em melhorias significativas na acurácia de modelos de predições [44]. O motivo principal para utilizar a abordagem semi-supervisionada é usar dados não rotulados para aumentar o conjunto de dados e conseqüentemente, o desempenho da classificação supervisionada.

Finalmente, métodos de aprendizagem por reforço consistem em modelos que aprendem por tentativa e erro ao interagir com um ambiente dinâmico [45]. Estes métodos são muito utilizados em jogos e sistemas de navegação dinâmicos, pois não é necessário fornecer dados como exemplo. Os modelos procuram maximizar seu desempenho geral a partir das suas experiências [43].

1.7.3 - Métricas de Avaliação

Para medir a performance de um classificador e comparar classificadores entre si, são usadas métricas de avaliação. As mais comumente usadas são acurácia, precisão, recall, F1 e AUC - do inglês “*Area Under the Curve*”, que corresponde à área sob a curva ROC .

A acurácia representa uma performance geral do modelo, indicando quantas classificações realizadas foram corretas. A precisão indica o número de classificações corretas dentre as classes positivas do modelo e a sensibilidade, o número de classificações corretas dentre as classes positivas esperadas. F1 representa a média harmônica entre precisão e sensibilidade. A curva ROC (do inglês, *Receiver operating characteristic*), representa graficamente como a sensibilidade varia com a especificidade [46]. A métrica conhecida como AUC é utilizada como uma propriedade que resume a curva ROC para que se obtenha uma medida do desempenho de um classificador em distinguir entre as classes [47].

No caso de uma classificação multiclases, a qualidade da classificação geral pode ser avaliada por meio da macro-média (média das medidas para cada classe) ou da micro-média (média de cada medida de verdadeiro positivo, falso positivo, verdadeiro negativo, e falso negativo). Enquanto a macro-média trata todas as classes igualmente, a micro-média favorece as classes com maior número de elementos [48].

A escolha da métrica depende do contexto da classificação [49]. Na hora da escolha das métricas de acurácia, deve-se considerar o contexto do problema e usar-se sempre mais de uma métrica para garantir que uma única métrica escolhida não seja limitada em descrever o desempenho da classificação.

1.7.4 - Balanceamento

Um aspecto que deve ser considerado no aprendizado de máquina é o balanceamento de dados. Esse processo deve ser realizado quando um conjunto de dados possui uma ou mais classes com incidência muito maior que as outras. Se não tratada, a incidência discrepante de classes do conjunto de dados faz com que o modelo seja inacurado, embora as métricas de acurácias sejam altas. Nesse caso, os parâmetros do algoritmo não diferenciam as classes minoritárias. Para tratar classes desbalanceadas, usamos o *undersampling* ou *oversampling*, que são técnicas de balanceamento. No *undersampling*, reduzimos o número de observações da classe majoritária. Existem duas formas de *undersampling*: o *Random Undersampling* consiste na retirada aleatória de elementos da classe majoritária, enquanto o agrupamento de classes, junta as classes minoritárias em classes mais numerosas com algoritmos de clusterização, ou métodos como *K-Nearest Neighbors*. O *oversampling* funciona por meio de cópias de dados existentes na classe minoritária. No caso do *Random Oversampling*, exemplos de classes minoritárias são duplicados aleatoriamente. O *oversampling* oferece resultados satisfatórios quando a classe minoritária não tem uma variação quantitativa muito grande nos seus parâmetros [50].

1.7.5 - Seleção das variáveis preditoras

O foco da seleção de preditores é obter um subconjunto de variáveis da entrada que possa descrever eficientemente os dados de entrada, reduzindo os efeitos de ruído ou variáveis irrelevantes e ainda assim fornecer bons resultados de previsão [51]. Um preditor consiste em uma variável do processo que está sendo observado. Nos últimos anos, nas aplicações de aprendizado de máquina ou reconhecimento de padrões, o domínio de preditores aumentou de dezenas para centenas usados nessas aplicações. Várias técnicas são desenvolvidas para resolver o problema de reduzir variáveis irrelevantes e redundantes. A seleção de variáveis preditoras (eliminação de variáveis) ajuda a entender os dados, reduzindo os recursos computacionais e o efeito do *overfitting* e melhorando o desempenho do preditor [45]. Existem duas principais classes de métodos para seleção de preditores: *wrapping* e *filtering*. Os métodos de *wrapper* (*wrapping*) usam o preditor como uma caixa preta e o desempenho do preditor como a função objetivo para avaliar o subconjunto de variáveis. Por outro lado, os

métodos de filtro (*filtering*) usam técnicas de classificação de variáveis como critério principal para a seleção de variáveis. Um critério de classificação adequado é usado para pontuar as variáveis e um limite é usado para remover variáveis abaixo do limite. [52,53,54].

1.7.6 - *Data Splitting* e Validação Cruzada

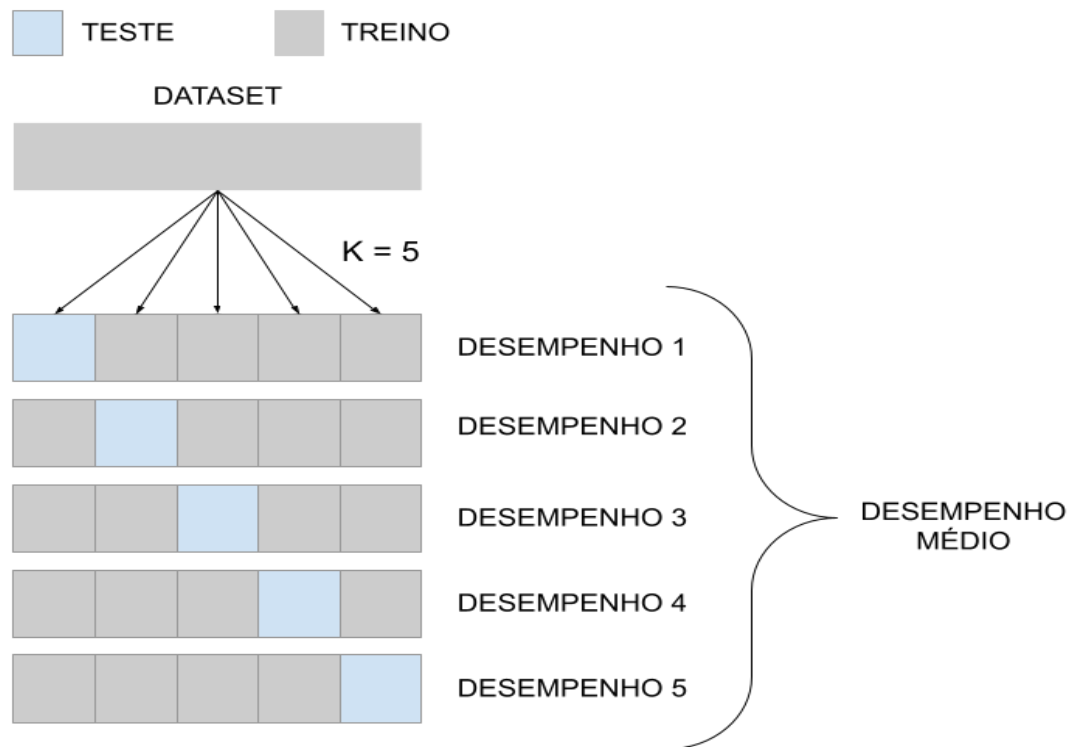
Para separar o conjunto de dados em conjunto de treinamento e conjunto de validação, é possível escolher os componentes de cada conjunto aleatoriamente, utilizando-se um gerador de números aleatórios. Esse procedimento previne vieses, como o ocasionado pela ordenação dos valores das variáveis no banco de dados. A separação em conjunto de treinamento e conjunto de validação garante que o conjunto de teste não seja usado para estimar os parâmetros do estimador, o que o torna consistente pela lei dos grandes números [55].

Data splitting é o nome dado ao processo de dividir os dados conjunto treino e conjunto teste para diferenciar os dados usados para estimar o risco do modelo.

A validação cruzada é uma forma de *data splitting* que faz uso de toda a amostra. Na validação cruzada *leave-one-out*, todas as observações, exceto a *i*-ésima delas, são alocadas para o treino do modelo [56].

Diferentemente, na validação cruzada *k-fold*, o processo consiste em dividir o conjunto de dados aleatoriamente em *k* conjuntos. São gerados *k* estimadores com *k*-1 subconjuntos de treino. O conjunto que não foi utilizado em cada estimador é usado para teste. O desempenho de cada estimador é obtido e a média destes desempenhos é a métrica final. Uma esquematização para a validação cruzada *k-fold*, com *k*=5 encontra-se na Figura 2.

Figura 2 - Esquema de uma validação cruzada *k-fold* com *k*=5. Os quadrados azuis representam as subdivisões (*fold*) usadas para teste e os cinzas, as subdivisões usadas para treino.



Fonte: Autor.

1.7.7 - Algoritmos de Aprendizado de Máquina

No que se refere aos algoritmos de AM, existe uma grande diversidade de métodos que podem ser usados para classificação. Alguns dos métodos mais tradicionalmente usados são: Redes Neurais Artificiais, K-Nearest Neighbours, Árvores de Decisão e Florestas Aleatórias. A escolha de um método depende de vários fatores. Trabalhos que buscam comparar o desempenho de classificadores têm ganho relevância ao longo dos anos. No início da década de 90, o projeto StatLog realizou uma comparação de algoritmos de classificação em grandes conjuntos de dados do mundo real e concluiu que o melhor algoritmo para um determinado conjunto de dados depende muito das características dos dados. Em outro artigo que compara a acurácia de classificadores em larga escala, Caruana e Niculescu-Mizil comparam dez algoritmos de aprendizagem supervisionada usando oito critérios de desempenho em onze problemas de classificação binária. Os resultados mostraram que Florestas Aleatórias, Máquina de Vetor de Suporte e Redes Neurais Artificiais estão entre os classificadores mais acurados. [48].

Mais recentemente, Fernández-Delgado et al testaram a acurácia de 179 classificadores (desenvolvidos em ambientes diversos) em 121 datasets, resultando em 21,659 combinações de modelos. Os resultados mostraram que as três melhores famílias de classificadores foram: Florestas Aleatórias, Máquina de Vetor de Suporte e Redes Neurais Artificiais [57]. Segundo R.D. King, C. Feng e A. Sutherlandz, os desenvolvedores do Statlog Project citado anteriormente, não existe um único algoritmo mais acurado. Entretanto, os estudos mencionados acima mostraram que certos classificadores obtiveram desempenho geral acima da média quando aplicados a conjuntos de dados diversos.

Os algoritmos Florestas Aleatórias, Máquina de Vetor de Suporte e Redes Neurais Artificiais se apresentaram como alguns desses algoritmos e por isso, foram utilizados neste projeto.

A seguir, o funcionamento destes algoritmos será explicado.

- Florestas Aleatórias

Árvores de classificação são algoritmos não paramétricos com resultados geralmente fáceis de interpretar. São construídas por meio de particionamentos recursivos no espaço das variáveis preditoras. Os particionamentos são denominados nós e os resultados, folhas. Uma nova predição da árvore é feita por meio do seguinte processo: verifica-se se a condição descrita no primeiro nó é satisfeita. Em caso afirmativo, segue-se pelo nó da esquerda e em caso contrário, pelo nó da direita. O processo é repetido até uma folha ser encontrada. A predição para a resposta Y de uma observação com variáveis preditoras x que estão em uma região R é dada pela moda, representada na Equação 1.1:

$$(1.1) \quad g(x) = \text{moda} \{ y_i : x_i \in R_k \}$$

A criação da estrutura de uma árvore de classificação consiste em duas etapas: a criação de uma árvore completa e complexa e a poda dessa árvore para evitar o *overfitting*. Na primeira parte, procura-se construir uma árvore com partições nas quais os valores de Y nas observações do conjunto de treinamento em cada uma das folhas sejam homogêneos. Para tanto, avalia-se o quão razoável uma dada árvore T é através de alguma métrica, como o

índice de Gini. O índice de Gini se torna mínimo quando cada folha contém somente observações de uma única classe, por isso tenta-se minimizá-lo. Apesar de serem altamente interpretáveis, as árvores de classificação são comumente menos acuradas do que os outros algoritmos.

O estimador conhecido como Florestas Aleatórias (Breiman, 2001) procura contornar essa limitação através da combinação de diversas árvores. A abordagem desse estimador consiste em criar B árvores distintas e combinar seus resultados para melhorar o poder preditivo em relação a uma árvore individual. Para diminuir a correlação entre as predições das B árvores, o algoritmo de Florestas Aleatórias modifica o mecanismo de criação das árvores de maneira que estas se tornem diferentes umas das outras. Dessa forma, ao invés de escolher qual das d variáveis preditoras será utilizada em cada um dos nós da árvore, só é permitido que sejam escolhidas aleatoriamente m variáveis preditoras, sendo $m < d$.

Resultados empíricos sugerem que $m \approx d/3$ leva a uma boa performance [31]. A cada nó criado, um novo subconjunto de variáveis preditoras é sorteado. A agregação das diferentes árvores de classificação é realizada através da Equação 1.2:

$$(1.2) \quad g(x) = \text{moda} \{g^b(x), b = 1, \dots, B\}$$

Essa equação significa que a predição final é obtida pela categoria predita com maior frequência dentre as árvores construídas [31].

- Máquina de Vetor de Suporte

Máquina de Vetor de Suporte (MVS) é uma técnica de aprendizagem supervisionada proposta por Cortes e Vapnik (1995) na qual o espaço do problema deve ser linearmente separável. Um hiperplano é traçado a partir do modelo, maximizando a margem de classificação. O hiperplano possui $N-1$ dimensões, sendo N o número de recursos presentes. Os nós de fronteira são chamados de vetores de suporte. Para encontrar a saída, um hiperplano ótimo é desenhado no ponto médio. Quando há mais de um hiperplano que separa os dados

perfeitamente, o estimador busca por aquele que tem maior margem M , ou seja, aquele que tem maior distância de todos os pontos observados. Dessa forma, o MVS busca pela solução da Equação 1.3:

$$(1.3) \quad \text{argmax } M$$

A solução ótima $f(x)$ do MVS pode ser escrita como na Equação 1.4 :

$$(1.4) \quad f(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_d x_d = \beta_0 + \sum_{k=1}^n \alpha_k \langle x, x_k \rangle$$

em que $\langle x, x_k \rangle = \sum_{i=1}^d x_i x_{k,i}$. Pode-se trocar $\langle x, x_k \rangle$ por um kernel genérico $K(x, x_k)$ e dessa forma, buscar por divisões mais complexas que hiperplanos.

Para calcular os coeficientes α , só é necessário o produto interno entre todas as observações. Adicionalmente, todos os pontos que não são vetores de suporte, terão valores de α_i iguais a zero. Dessa forma, a equação anterior pode ser escrita como na Equação 1.5:

$$(1.5) \quad f(x) = \beta_0 + \sum_{k \in S} \alpha_k \langle x, x_k \rangle$$

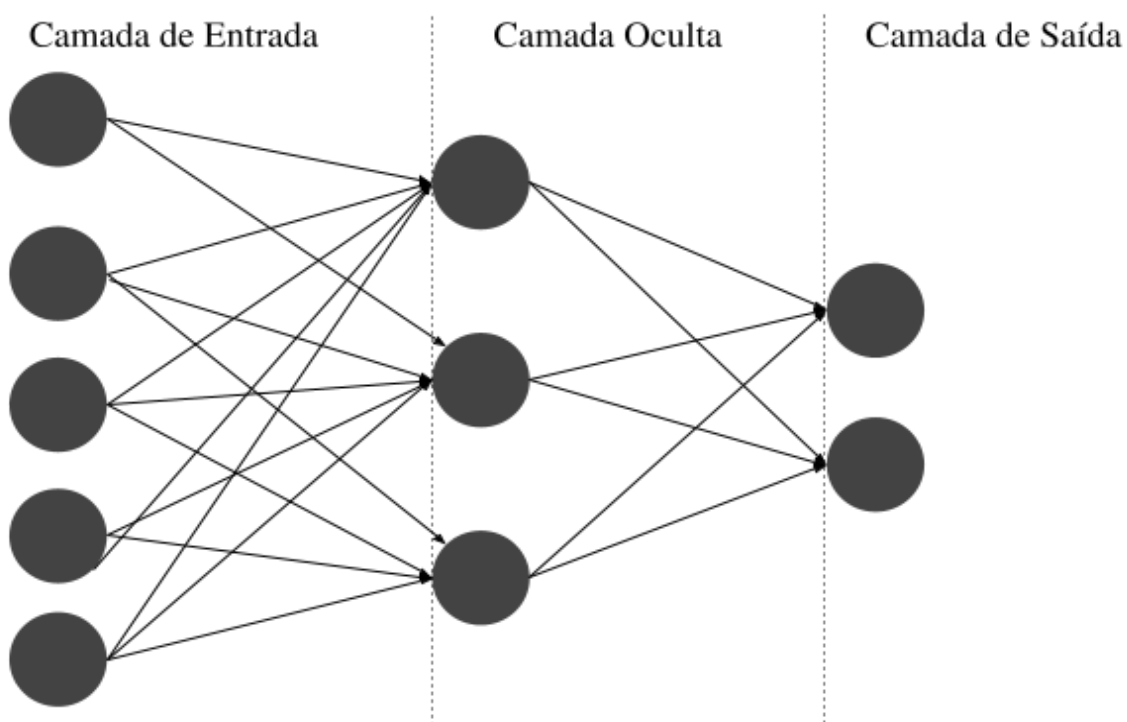
em que S é o conjunto de todos os vetores de suporte [35].

- Redes Neurais Artificiais

Redes neurais artificiais (RNA) são estimadores constituídos por camadas de nós, sendo a primeira, a camada de entrada. As camadas intermediárias são as camadas ocultas e a última camada é a camada de saída. A cada neurônio artificial (nó), que conecta-se aos demais, é associado um peso e um limite. O limite especificado tem a função de limitar as

saídas que estiverem acima de seu valor, ativando o nó. Caso um nó for ativado, ele enviará dados para a próxima camada da rede. A Figura 3 representa uma rede neural com 1 camada de entrada (com 5 nós), 1 camada oculta (com 3 nós) e uma 1 camada de saída (com 2 nós).

Figura 3 - Arquitetura de uma rede com 1 camadas de entrada (com 5 nós), 1 camada oculta (com 3 nós) e uma 1 camada de saída (com 2 nós)



Fonte: Autor.

Cada nó individual possui dados de entrada, pesos, um viés (ou limite) e uma saída. Os pesos ajudam a determinar a importância das variáveis fornecidas. Cada entrada deverá ser multiplicada por seu respectivo peso e em seguida, somadas. A saída do nó é transmitida por meio de uma função de ativação, que dispara o nó caso a saída exceda um determinado limiar, permitindo que os dados passem para a camada seguinte (rede *feedforward*). A precisão do modelo é avaliada por meio de uma função de custo (ou perda), por exemplo o erro médio quadrático (EMQ), ou no caso da classificação, entropia cruzada [40]. O modelo procura minimizar a função de custo utilizada para atingir o ponto de convergência (mínimo local) por meio da descida de gradiente.

As RNA podem ser classificadas em diferentes tipos: com o perceptron, as RNA *feedforward*, e as RNA recorrentes. O perceptron foi criado em 1958 por Frank Rosenblatt. É

a rede neural mais antiga e mais simples, possuindo apenas um neurônio. As redes neurais feedforward, ou perceptrons multicamadas (MLPs) são compostas por uma camada de entrada, uma camada ou mais camadas ocultas, e uma camada de saída. Apesar de serem denominados perceptrons, geralmente não são compostas por perceptrons, mas neurônios sigmóides devido ao fato de a maioria dos problemas do mundo real não serem lineares. Finalmente, as redes neurais recorrentes (RNR) são compostas por loops de feedback. Esses estimadores são utilizados principalmente com dados de séries temporais [58].

1.7.8 - Pipelines de Aprendizado de Máquina

Pipeline são fluxos de trabalho executáveis no qual cada tarefa é realizada de forma independente. A maioria dos pipelines de Aprendizado de Máquina incluem as seguintes tarefas:

- pré-processamento dos dados: abrangem as subtarefas de importação, limpeza e normalização dos dados;
- particionamento de dados: separação do conjunto de dados em teste, treinamento e validação;
- treinamento: conjunto de treino é aplicado ao modelo, o que significa que dados rotulados serão fornecidos para o modelo realizar seu processo de aprendizagem;
- validação do modelo: a partição dos dados separados para teste são rotulados pelo modelo e os resultados são contabilizados por meio de uma métrica de acurácia;

Entre as vantagens de usar pipelines de aprendizado de máquina estão a melhora da produtividade, a otimização do uso de técnicas de validação e seleção e a menor ocorrência de erros de manipulação e particionamento dos dados [59].

2 - Revisão Bibliográfica

2.1 - Estudos na literatura que utilizam métodos de AM em dados clínicos na predição de diagnósticos/condições médicas

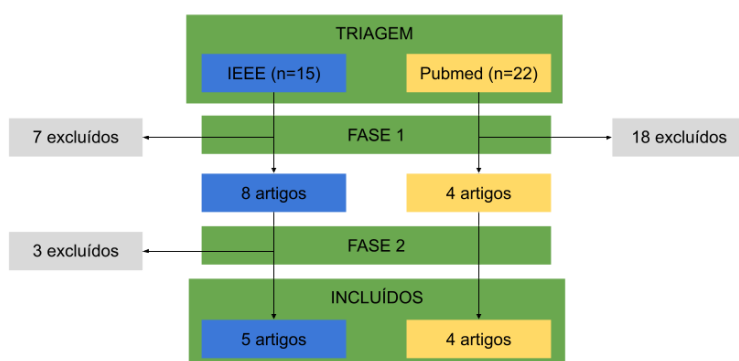
Muitos trabalhos científicos que utilizam métodos de AM aplicados a problemas da área da saúde foram realizados ao longo dos anos. Para avaliar a viabilidade do uso de modelos de AM aplicados a dados da área da saúde e obter seu desempenho geral para fim de comparação com os resultados do presente estudo, realizou-se uma revisão bibliográfica.

Foram avaliados estudos que aplicaram métodos de AM em dados clínicos na predição de diagnósticos/condições médicas (de maneira similar ao que foi realizado neste estudo).

No que se refere à busca pelos artigos, foram utilizadas as palavras chave “*machine learning*” (aprendizado de máquina), “*clinical data*” (dados clínicos) e “*medical diagnosis prediction*” (predição de diagnóstico médico) para as buscas da base IEEE e PubMed.

Foram encontrados 15 e 22 artigos nas bases do IEEE e Pubmed, respectivamente. Dos 15 artigos encontrados no IEEE, 7 foram excluídos na primeira fase, e 3 na segunda, por não utilizarem dados clínicos ou por não tentarem predizer um diagnóstico médico. Dos 22 artigos encontrados no Pubmed, 18 foram excluídos na primeira fase por serem estudos de revisões sistemáticas, por não utilizarem dados clínicos ou não procurarem predizer um diagnóstico médico. Nenhum estudo obtido no Pubmed foi excluído na segunda fase. A Figura 4 esquematiza o processo de seleção de artigos.

Figura 4 - Processo de inclusão de artigos na revisão bibliográfica



Fonte: Autor. IEE: *Institute of Electrical and Electronics Engineers*.

No total foram incluídos 9 artigos, cujos resultados serão apresentados a seguir.

Zhang et al (2020) compararam a acurácia de um modelo integrado com modelos tradicionais ao prever uma condição médica dentre 187 doenças conhecidas. O modelo integrado teve uma melhor performance (acurácia média = 0,913) quando comparado com os modelos RNA (acurácia média = 0,822), MVS (acurácia média = 0,872), FA (acurácia média = 0,868) e *Extreme Gradient Boosting* (acurácia média = 0,886) [60].

Ni et al (2018) utilizaram 8131 casos do conjunto de dados Greater Cincinnati/Northern Kentucky Stroke Study de 2005 a 2010 para prever o diagnóstico de AVC. Como resultados, obtiveram as seguintes acurácias 61,68%, 85,85%, 87,21%, 87,89%, 88,38%, 88,57% e 86,90% no grupo teste para os algoritmos ICD9, CLIN, RL, MVS-P, MVS-R, FA e RNA respectivamente [61].

Zeiberg et al (2019) procuraram desenvolver uma abordagem de aprendizado de máquina para a previsão de Síndrome Respiratória Aguda com uso de dados de registros eletrônicos de saúde. O modelo que fazia uso do algoritmo *Extreme Gradient Boosting* obteve AUC igual a 0,75 e os modelos que fizeram uso de algoritmos de RL (L1 e L2), obtiveram AUC igual a 0,76 (L1) e 0,81 (L2) [62].

Coombes, Coombes e Fareed (2021) utilizaram os algoritmos de RL, Árvores de Classificação e Regressão, FA, Naïve Bayes e MVS para rotular o quadro de delírio (estado agudo e confusional associado a alta risco de morbidade e mortalidade a curto e longo prazo) em UTIs. O algoritmo de RL obteve AUC igual a 0,83, Árvores de Classificação e Regressão obteve AUC igual a 0,59, Florestas Aleatórias obteve AUC igual a 0,58, Naïve Bayes obteve AUC igual a 0,66 e MVS obteve AUC igual a 0,61 [63].

Anooj (2013) comparou modelos de lógica fuzzy com modelos de RNA utilizando dois conjuntos de dados públicos com dados de doenças cardiovasculares. O modelo de RNA obteve acurácia igual a 53,862% para o primeiro conjunto e 46,417%, para o segundo. Os modelos de lógica fuzzy obtiveram acurácias iguais a 51,793% e 36% (fuzzy-based CDSS), 51,793% e 36% (algoritmo de Tsipouras M.G.et al.'s) e 57,851 % e 50,583% (modelo Fuzzy desenvolvido no artigo) [64].

Hashi, Zaman e Hasan (2017) obtiveram acurácias iguais a 97,73 %, 76,96% e 90,43% para os algoritmos de RNA, MVS, KNN e Árvores de Classificação e Regressão, respectivamente, ao prever o diagnóstico de diversas doenças [65].

Zhang e Sun (2018) utilizaram 1031 pacientes para prever risco de câncer de mama em mulheres na pré-menopausa. Os resultados apresentaram que os algoritmos CD5 e

Adaptive Boosting C5.0 obtiveram acurácia de 97,73% e o algoritmo *Model Costmatrix* C5.0, de 85,39% [66]

Zhang et al (2021) compararam FA, Gradient Boosting, Extreme Gradient Boosting, MVS e ICD9 com algoritmos de *Ensemble Learning* para prever o diagnóstico de hemorragia pós-parto. Os resultados mostraram acurácias iguais a 97,8% (FA), 94,9 (Gradient Boosting), 96,6% (Extreme Gradient Boosting), 97,3% e 83,87% (dois conjuntos de dados diferentes - MVS), 96,7%, 96,2% e 90,3% para modelos diferentes de *Ensemble Learning* [67].

Thallam et al (2020) encontraram acurácias de 95,99% (RNA), 97,5% (FA), 95% (MVS) 97%, (RNA) e 99,5% (classificador por votos), ao prever o diagnóstico de câncer de pulmão [68].

3 - Objetivos e Hipótese

3.1 - Objetivos

O objetivo principal deste projeto de pesquisa é construir e validar um aplicativo de celular que determine automaticamente a etiologia do HIC de um paciente utilizando fatores de risco, variáveis clínicas, demográficas, e de neuroimagem como variáveis preditoras da etiologia.

Os objetivos secundários consistem em :

- Construir pipelines compostos pelas tarefas de rotulação por métodos semi supervisionados, seleção de atributos e classificação.
- Comparar a acurácia dos pipelines de modelos semi-supervisionados com estudos similares encontrados na literatura.
- Selecionar o melhor pipeline em prever a etiologia.
- Projetar aplicativo de celular com interface amigável para entrada e saída de dados.

3.2 - Hipótese

A ferramenta proposta obterá acurácia, medida por AUC, semelhante a modelos de classificação de dados médicos que fazem uso de métodos de AM para prever um diagnóstico/condição médica encontrados na literatura.

4 - Materiais e Métodos

4.1 - Delineamento do estudo

Esse projeto de pesquisa constitui um estudo analítico observacional transversal. O projeto foi aprovado pelo Comitê de Ética (CAAE: 52502721.9.0000.5440 - Número do Parecer: 5.043.454).

4.2 - Desenvolvimento do aplicativo

4.2.1 - Ferramentas de desenvolvimento

Foi utilizado o framework React-Native para desenvolvimento do aplicativo neste projeto de pesquisa. O código em Javascript foi compilado com auxílio de ferramentas do ambiente nativo (Android Studio e Xcode). O aplicativo foi testado em emuladores (ambientes virtuais que simulam o sistema operacional do aparelho celular) para ambas as plataformas, além de um aparelho físico Iphone 6S para a versão iOS. O design e fluxo das telas foi idealizado com auxílio de um designer.

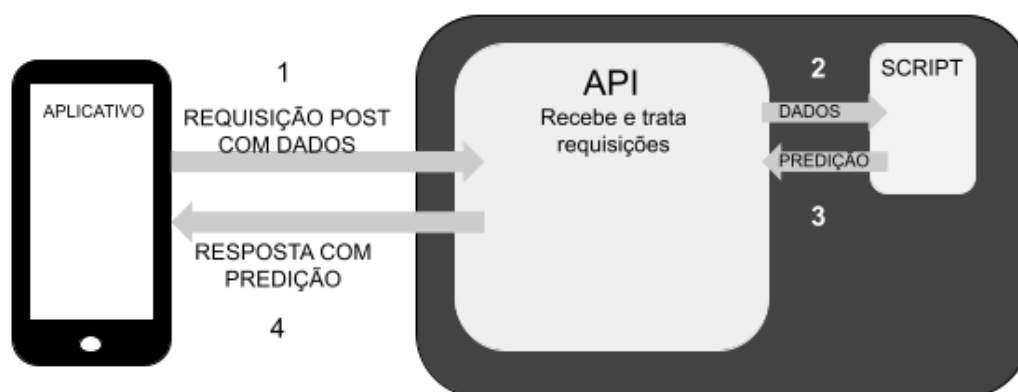
As funcionalidades planejadas para existirem no aplicativo foram cadastro, login e logout de usuários, classificação da etiologia por meio de um formulário de dados de um paciente, além da classificação de grande quantidade de dados por meio do download de um template no formato .csv e do upload desse template preenchido com os dados a serem classificados.

O modelo de predição foi implementado em um *script* Python, que é executado quando o aplicativo faz uma requisição a uma API (Interface de Programação de Aplicações, do inglês *Application Programming Interface*). Esta API foi desenvolvida em Node.js, um framework JavaScript assíncrono e se comunica com um banco de dados orientado a documentos (MongoDB) para armazenar dados de acesso dos usuários do aplicativo.

4.2.2 - A arquitetura do sistema

A arquitetura do sistema consiste principalmente em dois componentes: uma API, hospedada em um servidor em nuvem, e no aplicativo. Quando o aplicativo faz uma requisição POST (requisição que envia os parâmetros no corpo da requisição) enviando os dados a serem preditos à API, esta executa o *script* para que o modelo classifique esses novos dados e envia as etiologias preditas na resposta para o aplicativo. A API também lida com as requisições de login e de criação de novos usuários do aplicativo, armazenando os dados no banco de dados. Por questões de privacidade, os dados médicos (variáveis preditoras e etiologias) não serão armazenados no banco de dados. A Figura 5 apresenta uma representação do sistema:

Figura 5 - Fluxo de informações entre o aplicativo e o servidor.



Fonte: Autor. API: Interface de programação de aplicativos (do inglês, *Application Programming Interface*). 1 - O aplicativo faz uma requisição POST enviando os dados a serem preditos à API. 2 - API passa os dados e executa o *script* de classificação. 3 - *Script* de classificação envia as etiologias preditas para API. 4 - API envia resposta com etiologias preditas para o aplicativo.

4.3 - O modelo de predição

4.3.1 - *Data-Splitting*, Pré-processamento, Tratamento de Dados Faltantes e Balanceamento de Classes

O primeiro passo para a construção dos modelos de predição foi separar os dados (*data-splitting*) de forma aleatória e estratificada em dois grupos: 70% dos dados foram

destinados ao treino e teste dos modelos e 30% dos dados foram destinados à validação externa.

Na etapa seguinte, a etapa do pré-processamento, os dados foram convertidos para os tipos necessários, isto é, os dados que representam variáveis numéricas foram convertidos em números, e os que representam variáveis categóricas foram convertidos em texto. Isso é necessário, porque o computador nem sempre usa o tipo esperado quando faz a leitura de uma planilha ou banco de dados, o que influencia posteriormente nas análises e no desempenho dos modelos.

Depois do pré-processamento, os dados faltantes foram preenchidos com a média (variáveis numéricas) ou com classe majoritária (variáveis categóricas) da respectiva variável. Em seguida, o processo de balanceamento de classes foi executado para garantir que as classes de etiologias não estivessem em quantidades muito diferentes, o que influencia negativamente no desempenho do modelo. Foi utilizado o método de balanceamento conhecido como *Random Over Sampling*.

4.3.2 - Rotulação de Classes Faltantes

O primeiro passo dos pipelines desenvolvidos foi rotular as instâncias de dados que não possuíam a respectiva etiologia do episódio de HIC. Pois para criar os modelos de predição, é necessário que cada instância de dados possua o respectivo rótulo, tanto para treinar quanto para validar o modelo.

Como muitas vezes o processo de classificação de dados por médico especialista é demorado e a quantidade de dados é extensa, optou-se por utilizar métodos semi-supervisionados para validá-los. Para isso, foram retirados os rótulos de 5% das instâncias dos dados destinados ao treino e teste (70% dos dados iniciais) dos modelos, de forma aleatória para classificá-los por meio de um processo semi-supervisionado.

Neste processo, os dados rotulados e não rotulados foram separados. Foram criados três diferentes modelos treinados com os dados rotulados, utilizando-se os algoritmos Florestas Aleatórias (FA), Máquina de Vetor de Suporte (MVS) e Redes Neurais Artificiais (RNA).

Os dados não rotulados foram preditos por cada um destes modelos e as instâncias que tiveram probabilidade de estar correta acima de 80% são incorporados ao conjunto de treinamento. Esse processo é repetido até não sobraem dados não rotulados.

4.3.3 - Seleção de Variáveis Predictoras

A próxima etapa de construção dos pipelines consistiu em selecionar as variáveis predictoras com maior relevância preditiva. Esse passo é importante para diminuir ruídos e evitar o *overfitting*.

Foram utilizados dois métodos de seleção de variáveis predictoras: um filtro com uso do algoritmo estatístico χ^2 e o algoritmo de Regressão Logística (RL) por meio de uma eliminação recursiva de atributos.

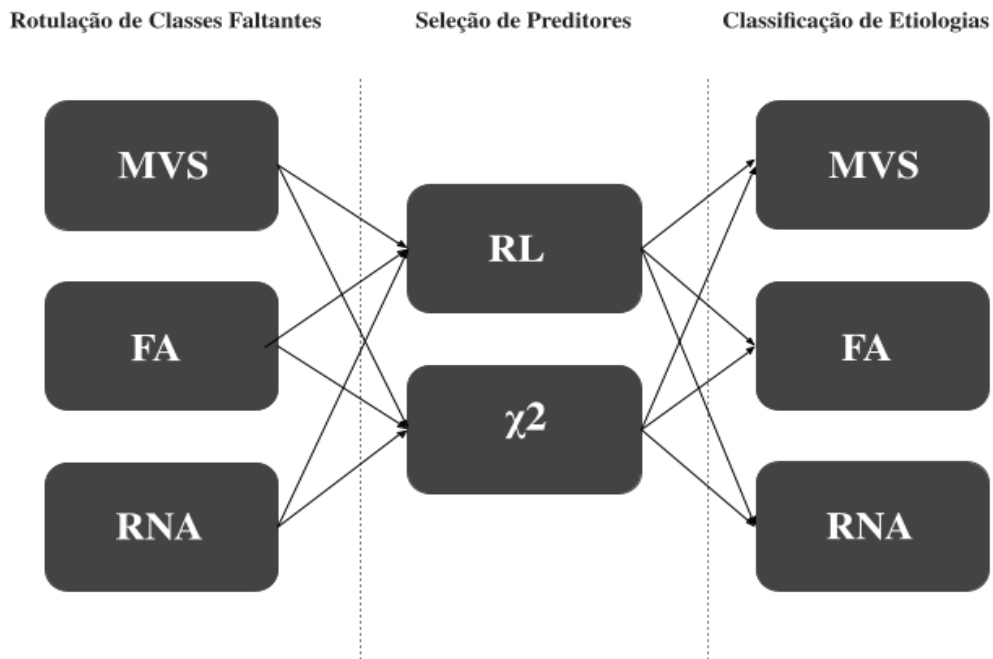
Ao final de ambos os processos, foram selecionadas as 18 variáveis mais relevantes. Esse número foi uma escolha arbitrária, dado que os parâmetros e os resultados da seleção de variáveis predictoras são muito específicos do conjunto de dados e não há um valor global [69].

4.3.4 - Classificação de Etiologias

A etapa final de construção dos pipelines é o treino dos modelos que fazem a classificação das etiologias. Os grupos de treino e teste foram separados por *cross-validation* (k=5). Com isso, foi possível obter mais grupos de treino e teste. Os algoritmos usados para construir os modelos foram MVS, RNA, e FA.

Ao final do processo, obtivemos 18 pipelines: 3 diferentes modelos de rotulação seguidos de 2 algoritmos de seleção de preditores, e finalizados por 3 diferentes modelos de classificação (Figura 6).

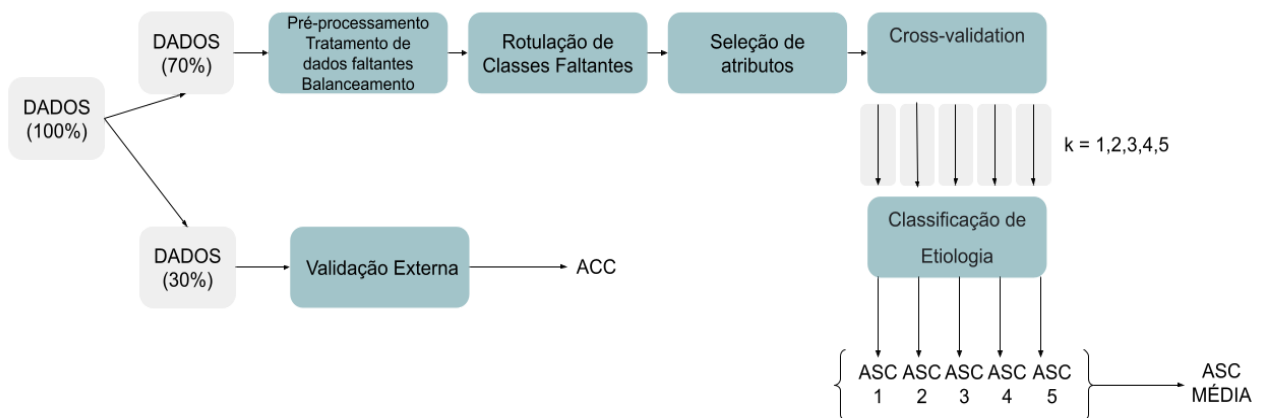
Figura 6 - Os 18 pipelines construídos pela combinação de 3 modelos de rotulação, 2 algoritmos de seleção de variáveis predictoras, e 3 modelos de classificação.



Fonte: Autor.

O esquema completo de como ocorreu a preparação e a divisão de dados e a construção e a comparação dos pipelines para a obtenção do melhor modelo a ser utilizado no aplicativo encontra-se na Figura 7.

Figura 7 - Preparação e divisão de dados e construção e comparação dos pipelines para obtenção do melhor modelo.



Fonte: Autor.

4.4 - Avaliação do modelo de predição

Estes pipelines tiveram suas acurácias comparadas e o que obteve melhor performance foi persistido para servir ao aplicativo. A métrica utilizada foi a AUC.

4.5 - Validação Externa

Dos episódios de HIC coletados, 30% foi separado para utilização na validação externa. Esses episódios tiveram suas etiologias classificadas pelo aplicativo e a acurácia inter-classe foi obtida, utilizando-se a classificação do especialista como padrão ouro.

4.6 - Variáveis preditoras

As variáveis a serem selecionadas como preditores da etiologia foram escolhidas devido a sua relação com certas classes de etiologia da HIC. Além disso, alguns dos fatores de risco mais conhecidos para HIC foram incluídos. Todos os dados foram coletados de forma retrospectiva no sistema do prontuário médico do HCFMRP, com exceção do volume e a localização de sangramento dos episódios, que foram verificados no exame de neuroimagem de admissão (TC). A seguir, encontra-se a descrição de como as variáveis foram definidas no presente estudo.

4.6.1 - Variáveis demográficas

As variáveis demográficas coletadas foram sexo, idade, raça, tabagismo, etilismo e uso de drogas, que foram definidas da seguinte forma:

- sexo: sexo de nascimento do paciente;
- idade: idade do paciente quando ocorreu o respectivo episódio de HIC;
- raça: raça declarada presente no prontuário médico;
- tabagismo: tabagismo atual (pelo menos 1 cigarro/dia nos últimos 30 dias);
- etilismo: alcoolismo atual ou no último ano (maior que 2 dose-padrão);
- uso de drogas: paciente faz uso de qualquer tipo de droga de forma regular;

4.6.2 - Variáveis clínicas

As variáveis demográficas coletadas foram pressão sistólica, pressão diastólica, glicemia, TTPA, TP, diagnóstico de diabetes, hipertensão, angiopatia, coagulopatia, vasculopatia, angiopatia amilóide e aneurisma, ingestão de varfarina e de AAS. Estas variáveis foram definidas da seguinte forma:

- pressão sistólica: pressão sistólica em mmHG coletada na admissão;
- pressão diastólica: pressão diastólica em mmHG coletada na admissão;
- glicemia: exame de glicemia em mg/dL realizado em até 1 dia após a admissão;
- TTPA: valor de tempo do paciente do teste de TTPA realizado até 1 dia após a admissão.
- TP: valor de tempo do paciente do teste de TP realizado até 1 dia após a admissão;
- diagnóstico de diabetes: diagnóstico prévio de diabetes mellitus identificado por testes de tolerância à glicose oral e em jejum.
- diagnóstico de hipertensão: história relatada de hipertensão arterial ou uso prévio de anti-hipertensivos;
- diagnóstico de angiopatia: diagnóstico prévio de algum tipo de angiopatia, como angiopatia diabética ou amilóide;
- diagnóstico de coagulopatia: diagnóstico prévio de coagulopatia detectada por alterações de fatores de coagulação em exame médico;
- diagnóstico de angiopatia amilóide: diagnóstico prévio de angiopatia amilóide através de exame do enquadramento nos Critérios de Boston [70];
- diagnóstico de vasculopatia: diagnóstico prévio de doenças ateroscleróticas como doença arterial obstrutiva periférica, doenças coronarianas/infarto, aterosclerose de carótidas.
- diagnóstico de aneurisma: diagnóstico prévio de aneurisma através de exame de Doppler Vascular;
- ingestão de varfarina: uso prévio de varfarina;
- ingestão de AAS: uso prévio de AAS;

4.6.3 - Variáveis neuroimagem

As variáveis de neuroimagem coletadas foram volume de sangramento, calculados por um especialista segundo o método ABC/2 [10,11] pelo exame de TC de admissão; lado e localização do sangramento, verificados também por um especialista nos exames de TC de admissão; escalas ICH e SICH, obtidas do prontuário médico.

4.6.4 - Variável de interesse - etiologia segundo a classificação SMASH-U

A classificação SMASH-U foi realizada também por um especialista, segundo o esquema representado na Figura 1.

Como ambos os sistemas normalmente utilizados (SMASH-U e H-ATOMIC) têm boa confiabilidade, foi escolhido o sistema SMASH-U para a classificação etiológica neste projeto devido à maior familiaridade com esse sistema da equipe de especialistas que rotularam as etiologias manualmente. Além disso, o sistema SMASH-U não utiliza sobreposição de categorias, o que facilita a construção do modelo de AM. Caso fosse utilizada uma classificação que permitisse sobreposição de classes, como a H-ATOMIC, seria necessário construir um modelo de classificação multirrótulo. Na classificação multirrótulo, cada instância de dados pode estar relacionada a uma ou mais classes preditas.

Adicionalmente, a classificação pelo sistema H-ATOMIC demandaria a coleta de dados prospectivamente, o que seria inviável neste estudo no momento.

4.7 - Dados Utilizados : Registro de Acidente Vascular Encefálico (REAVÉR)

As variáveis clínicas usadas como variáveis preditoras e os rótulos para o treinamento dos métodos de classificação supervisionada serão retirados do Registro de Acidente Vascular Encefálico (REAVÉR) referentes aos anos de 2018 e 2019 (01/01/2018 - 31/12/2019). O REAVÉR faz parte do esforço para coletar grande quantidade de informações relevantes sobre os pacientes e ao mesmo tempo, avaliar prospectivamente a sua evolução clínica e funcional. Sendo assim, consiste em registros eletrônicos prospectivos, unificados, digitais e de caráter científico, que incluem o perfil clínico e epidemiológico de todos os doentes acima de 18 anos com diagnósticos de AVC atendidos na Unidade de Emergência (UE) do HCFMRP. Contém informações abrangentes sobre características clínicas e complementares sobre estes pacientes, incluindo avaliação de desfechos de longo prazo. A utilização do REAVÉR neste trabalho tem como motivação o fato da implementação de novas terapias e estratégias de

saúde populacional voltadas para o combate do AVC ter fundamentado-se na instalação e manutenção de registros prospectivos. A manutenção de registros prospectivos tem servido no cenário global para fomentar o desenvolvimento de pesquisa clínica, sabendo-se das vantagens metodológicas da coleta prospectiva de dados. Como exemplo da relevância de registros clínicos na pesquisa clínica do AVC pode-se citar a iniciativa de aderência às diretrizes de tratamento do AVC da American Heart Association (Get With the Guidelines-Stroke) e o registro Safe Implementation of Thrombolysis in Stroke, que serviu de base na Europa para a aplicação segura da terapia trombolítica para o AVC isquêmico agudo [71]. Apesar disso, existem poucos dados de registros prospectivos de atendimento agudo para AVC no país e a maioria destes são descontínuos e restritos a subgrupos de doenças. A seguir, encontra-se a descrição da amostra de dados do REAVER, bem como os critérios de inclusão e exclusão.

O REAVER engloba todos os pacientes com AVC admitidos a partir da data de sua aprovação pelo Comitê de Ética na Unidade de Emergência do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (UE-HC-FMRP-USP). A UE-HC é o principal pólo de atendimento a pacientes politraumatizados de maior complexidade de toda região Nordeste do estado de São Paulo que abrange cerca de 4 milhões de pessoas

4.7.1 - Amostra

Foram utilizados dados de todos os pacientes com diagnóstico de HIC (validado por especialista) inclusos no REAVER nos anos de 2018 e 2019.

4.7.2 - Critérios de Inclusão e Exclusão

Foram utilizados os seguintes critérios de inclusão dos pacientes no REAVER:

- idade igual ou superior a 18 anos;
- diagnóstico de AVC confirmado por exame de neuroimagem, com início dos sintomas até 24 horas antes da admissão;
- assinatura de termo de consentimento livre e esclarecido pelo paciente ou representante legal.

5 - Resultados e Discussão

A seção a seguir apresenta os resultados obtidos neste estudo e faz uma comparação com resultados de trabalhos similares na literatura.

5.1 - Resultados Descritivos

5.1.1 - Estatísticas descritivas

Foram incluídos 66 pacientes no estudo e coletados 25 atributos (categóricos e numéricos). A quantidade de dados faltantes em cada variável está representada de forma estratificada pela etiologia na Tabela 2.

Tabela 2 : Quantidade de dados faltantes em cada variável estratificada pela variável etiologia.

	SISTÊMICO	MEDICAÇÃO	HAS	AAC	ESTRUTURAL	U
sexo	0	0	3	0	0	0
idade	0	0	3	0	0	0
raça	0	0	3	0	0	0
tabagismo	0	0	3	0	0	0
alcoolismo	0	0	3	0	0	0
drogas	0	0	5	0	0	0
pressão sistólica	1	0	7	0	1	1
pressão diastólica	1	0	6	0	1	1
glicemia	0	0	24	2	1	6
diabetes	0	0	3	0	0	0
hipertensão	0	0	3	0	0	0
angiopatia	0	0	3	0	0	0
coagulopatia	0	0	3	0	0	0
vasculopatia	0	0	3	0	0	0
angiopatia	0	0	4	0	0	2

amilóide						
aneurisma	0	0	13	0	0	3
volume	0	0	2	0	0	0
sich	2	3	24	2	1	4
ich	1	2	13	2	1	2
lado	0	0	1	0	0	0
localização	0	0	0	0	0	0
TTPA	0	0	4	1	0	1
TP	0	0	5	1	0	1
anticoagulação (varfarina)	0	0	3	0	0	0
anticoagulação (AAS)	0	0	6	0	0	0

Fonte: Autor. AAC: angiopatía amilóide, HAS: hipertensão, U: indefinido, AAS: ácido acetilsalicílico.

Os atributos numéricos são: idade, pressão sistólica, pressão diastólica, glicemia, volume da hemorragia, TTPA e TP. Os atributos categóricos são: sexo, raça, tabagismo, etilismo, uso de drogas, diagnósticos de diabetes, hipertensão, angiopatias, coagulopatias, vasculopatias, angiopatía amilóide, aneurisma, SICH, ICH, lado e localização da hemorragia, anticoagulação (varfarina) e anticoagulação (AAS). As variáveis contínuas normais foram descritas pela média e desvio padrão e as variáveis contínuas não-normais foram descritas pela mediana e variação interquartil. As variáveis categóricas foram descritas pela contagem e pela porcentagem.

Do total de pacientes, a maioria eram sexo masculino (60,6 %), declarados brancos (71,2%), e com idade média de 60,8 e desvio padrão de 13,12. Os fumantes somaram um total de 16 (24,24 %) pacientes, assim como os etilistas, e apenas 2 (3,0 %) usavam drogas. A mediana da glicemia foi de 148 (122-190) ml/dm³. As medianas de TTPA e TP foram de 29,25 (25,7-34,45) e 12,3 (10,55-14,45). Os que tomam varfarina e AAS contabilizaram 5 (7,57 %) e 14 (21,21%), respectivamente.

Pacientes hipertensos constituíram 71,2 % do total (47 pacientes), valo3a da pressão sistólica foi de 170 (153-200) e a da pressão diastólica 96 (85-110). Quanto às demais comorbidades, 12 (18,18%) possuem diabetes, 3 (4,54%) possuem um ou mais aneurismas, 1

(1,51%) possui angiopatia amilóide e nenhum possui coagulopatias, vasculopatias ou angiopatias. Em relação às variáveis de neuroimagem, a mediana do volume da hemorragia foi de 9,0 (2,6-22,65) cm³. Houve mais pacientes com hemorragia no lado esquerdo (30; 45,45%) do que aqueles do lado direito (27; 40,9%), ou em ambos os lados (8; 12,12%). O valor 1 foi o mais frequente nas escalas ICH e SICH. A localização da hemorragia mais encontrada foi a nucleocapsular, seguida da talamocapsular e da talâmica. A frequência de pacientes segundo a localização da hemorragia pode ser vista na Tabela 3 e a de pacientes segundo os valores das escalas SICH e ICH, nas Tabelas 4 e 5. Os dados que faltam para somar a porcentagem de 100% são faltantes.

Tabela 3 : Frequência das diferentes localizações onde a hemorragia foi identificada

Localização da Hemorragia	Número de pacientes (66)
NUCLEOCAPSULAR	28 (42,42%)
TALAMOCAPSULAR	9 (13,63%)
TALÂMICO	5 (7,6%)
TRONCO	5 (7,6%)
FRONTAL	4 (6,0%)
OCCIPITAL	4 (6,0%)
PONTE	3 (4,5%)
NUCLEOCAPSULAR/TRONCO	1 (1,5%)
PARIETOOCCIPITAL	1 (1,5%)
TEMPORAL	1 (1,5%)
FRONTOTEMPORAL	1 (1,5%)
F-P-T	1 (1,5%)
PARIETAL	1 (1,5%)
TEMPOROPARIETAL	1 (1,5%)
FRONTOPARIETAL	1 (1,5%)

Fonte: Autor. FPT: frontal-parietal-temporal

Tabela 4 - Frequência dos valores da escala neurológica ICH.

ICH	Número de pacientes (45/66)
0.0	13 (19,7%)
1.0	17 (25,7%)
2.0	7 (10,6%)

3.0	5 (7,6%)
4.0	2 (3,0%)
5.0	0 (0%)
6.0	1 (1,5%)

Fonte: Autor. ICH: Escore de Hemorragia Intracerebral (do inglês, *Intracerebral Hemorrhage Score*).

Tabela 5 - Frequência dos valores da escala neurológica SICH.

SICH	Número de pacientes (30/66)
0.0	1 (1,5%)
1.0	14 (21,21%)
2.0	11 (16,7%)
3.0	1 (1,5%)
4.0	2 (3,0%)
5.0	1 (1,5%)
6.0	0 (0%)

Fonte: Autor. SICH: Escore de Hemorragia Intracerebral Secundária (do inglês, *Secondary Intracerebral Hemorrhage Score*).

No que se refere à etiologia (variável dependente), a etiologia primária hipertensiva foi a mais comum (46 pacientes - 69,7%), seguida pelas etiologias indefinida (9 pacientes - 13,6%), sistêmica (3 pacientes - 4,6%), angiopatia amilóide (3 pacientes - 4,6%), estrutural (3 pacientes - 4,6%) e medicação (2 pacientes - 3,0%).

Ao agrupar os pacientes por etiologias e calcular estatísticas descritivas para cada grupo, foi observado que a média de idades dos subgrupo de etiologia indefinida foi a maior (74,8; com desvio padrão de 14,1), seguida pelas médias de idades dos subgrupos das etiologias angiopatia amilóide (68,0; com desvio padrão de 5,2), medicação (61,0 com desvio padrão de 4,2), sistêmico (53,3; com desvio padrão de 17,7) e estrutural (52,0; com desvio padrão de 14,2). O subgrupo com as maiores medianas de pressão sistólica e diastólica foi o subgrupo de causa hipertensiva 180 (161,50-200) para a sistólica e 102,5 (90-120), para a diastólica. As medianas da glicemia de admissão foram maiores nos subgrupos sistêmico - 304 (196-412) ml/dm³ e indefinido - 240 (199-268) ml/dm³.

O tabagismo esteve proporcionalmente mais presente na subgrupos de etiologia estrutural e sistêmico e etilismo, nas etiologias estrutural e hipertensiva. Todos os pacientes

que usavam drogas estavam no subgrupo hipertensivo. Os subgrupos indefinidos e hipertensivos, nesta ordem, foram os que obtiveram maiores porcentagens de indivíduos com o diagnóstico de hipertensão (cálculo de porcentagem dentro de um mesmo grupo).

O diagnóstico de diabetes esteve proporcionalmente mais presente no subgrupo sistêmico. Todos os pacientes com o diagnóstico de angiopatia amilóide estavam no subgrupo de etiologia por angiopatia amilóide, mas curiosamente eles somam apenas 33% dos pacientes deste grupo. A maior parte dos pacientes com aneurisma estavam no subgrupo indefinido e o restante no subgrupo hipertensivo.

Quanto aos volumes da hemorragia, o subgrupo com maior mediana foi o subgrupo por angiopatia amilóide - 63 (31,8-96,50) cm³ seguido pelo subgrupo com causa indefinida - 41 (9-81) cm³. O lado da hemorragia mais encontrado em todos os subgrupos foi o esquerdo.

Já as localizações da hemorragia mais encontradas, foram respectivamente: angiopatia amilóide (Frontal/Frontotemporal/Occipital), estrutural (Frontoparietal/Talâmico/Temporal), hipertensiva (Nucleocapsular), medicação (Nucleocapsular), sistêmica (Frontal) e indefinida (Nucleocapsular).

Pacientes em anticoagulação (varfarina) estiveram proporcionalmente mais presentes no subgrupo de medicação, já os em anticoagulação (AAS), nos subgrupos estrutural, sistêmico e indefinido.

As estatísticas descritivas completas por subgrupo de etiologia, inclusive os referentes às variáveis que não foram discutidas acima, podem ser encontradas no Anexo 1.

5.2 - Comparação de pipelines

5.2.1 - Rotulação de etiologias faltantes com método semi supervisionado

Um dos objetivos do presente estudo é validar a aplicação de métodos semi-supervisionados no contexto de dados demográficos e clínicos de episódios de HIC. Dessa forma, foram retiradas 5% das etiologias dos dados destinados à comparação de modelos para que os modelos semi-supervisionados os rotulassem. A ideia inicial era fazer essa remoção de etiologias de forma estratificada, ou seja, retirar a mesma proporção dentro de cada classe. Mas como a quantidade de pacientes incluídos de algumas classes foi pequena (inferior a 3), optou-se por não seguir essa abordagem. Como resultado, obteve-se que todas as etiologias retiradas aleatoriamente foram da classe de etiologia hipertensiva. Todos os três

algoritmos utilizados (MVS, RNA, FA) acertaram a rotulação, isto, atribuíram a classe de etiologia hipertensiva às etiologias faltantes.

5.2.2 - Seleção de atributos

A seleção de atributos resultou em 18 atributos selecionados por cada algoritmo. Em ambos os algoritmos (χ^2 e RL), foram selecionados os seguintes atributos em comum: sexo, idade, raça, etilismo, pressão sistólica, pressão diastólica, glicemia, diabetes, hipertensão, volume da hemorragia, ICH, localização da hemorragia, TTPA, TP, anticoagulação (AAS) e anticoagulação (varfarina).

Adicionalmente, no algoritmo χ^2 foram selecionados “tabagismo” e “angiopatia amilóide”, e no algoritmo χ^2 , foram selecionados “SICH” e “lado da hemorragia”. Os atributos selecionados pelos algoritmos podem ser encontrados na Tabela 6. Os atributos selecionados apenas por um algoritmo estão em negrito. O valor de χ^2 e o ranking (para o algoritmo de RL) para cada atributo encontram-se na Tabela 7.

Tabela 6 - Atributos selecionados pelos algoritmos χ^2 e RL.

χ^2	RL
sexo	sexo
idade	idade
raça	raça
tabagismo	etilismo
etilismo	pressão sistólica
pressão sistólica	pressão diastólica
pressão diastólica	glicemia
glicemia	diabetes
diabetes	hipertensão
hipertensão	volume da hemorragia
angiopatia amilóide	ICH
volume da hemorragia	SICH
ICH	lado da hemorragia

localização da hemorragia	localização da hemorragia
TTPA	TTPA
TP	TP
anticoagulação (varfarina)	anticoagulação (varfarina)
anticoagulação (AAS)	anticoagulação (AAS)

Fonte: Autor. χ^2 : Qui-Quadrado, RL: Regressão Logística, AAS: ácido acetilsalicílico. Os atributos em negrito são os selecionados apenas por um algoritmo.

Tabela 7 - Atributos com seus valores respectivos de χ^2 e o *ranking* (para o algoritmo de RL).

Atributo	χ^2	<i>Ranking</i>
sexo	29,573770	1
idade	64,236237	1
raça	51,393939	1
tabagismo	17,000000	2
alcoolismo	40,000000	1
drogas	NaN	5
pressão sistólica	490,408274	1
pressão diastólica	144,402707	1
glicemia	276,530347	1
diabetes	44,000000	1
hipertensão	8,696970	1
angiopatia	NaN	7
coagulopatia	NaN	8
vasculopatia	NaN	6
angiopatia amilóide	48,000000	4
aneurisma	4,000000	3

volume	831,196102	1
sich	2,086957	1
ich	22,762712	1
lado	8,254777	1
localização	118,155172	1
TTPA	190,655108	1
TP	1096,882298	1
anticoagulação (varfarina)	33,454545	1
anticoagulação (AAS)	40,484848	1

Fonte: Autor. χ^2 : Qui-Quadrado, RL: Regressão Logística, AAS: ácido acetilsalicílico. O valor χ^2 é o valor calculado e usado para selecionar variáveis no método de *filtering*. O *ranking* representa a posição em que o atributo se encontra no ranqueamento de atributos para a eliminação recursiva. Os atributos com *ranking* 1 foram os selecionados.

O pipeline que obteve melhor desempenho foi o que utilizou o algoritmo de seleção de features RL, portanto os atributos utilizados para classificar os dados no modelo final foram os selecionados por este algoritmo. Isto significa que estes atributos são os necessários para classificar os dados pelo aplicativo desenvolvido, que utiliza o modelo final para classificar as etiologias.

5.2.3 - Classificação de etiologias

A classificação, que constituiu a última etapa dos pipelines construídos, resultou em AUC média para os grupos de teste acima de 0,90 na maior parte dos pipelines.

Como optou-se por separar os grupos de teste e treino por *cross-validation*, foram obtidas 5 métricas de desempenho para cada pipeline, uma para cada dobra do processo de *cross-validation*. Os dados referentes à AUC para cada dobra de cada pipeline podem ser encontrados nas Tabela 8, 9 e 10. A Tabela 8 representa os dados obtidos nos pipelines cujo algoritmo utilizado na primeira etapa (rotulação de dados faltantes) foi MVS. A Tabela 9, representa os resultados obtidos nos pipelines cujo algoritmo utilizado na primeira etapa foi RNA, e a Tabela 10, RF. Em cada uma das tabelas, a segunda (seleção de atributos) e a última

(classificação) etapas dos pipelines se encontram no cabeçalho, separados por “/”. As linhas são referentes a cada dobra do *cross-validation*. Por exemplo, o dado encontrado na linha 1 e coluna 1 da Tabela 8 representa a AUC da primeira dobra do pipeline composto pelos algoritmos: MVS- χ^2 -MVS, nessa ordem. É possível reparar que em todas as tabelas os resultados são idênticos. Isso ocorre porque os resultados dos diferentes algoritmos de rotulação foram iguais, resultando no mesmo conjunto de dados rotulados.

Tabela 8 - Resultados dos pipelines que usaram MVS na primeira etapa.

	χ^2 /MVS	χ^2 /RNA	χ^2 /FA	RL/MVS	RL/RNA	RL/FA
0	0,95	0,509375	0,975	0,95	0,475	0,975
1	0,9375	0,5375	0,971875	0,9375	0,5	0,9875
2	0,978125	0,5	0,96875	0,978125	0,5	0,975
3	0,9625	0,5375	1	0,9625	0,5	1
4	0,921875	0,5	0,9875	0,921875	0,5	0,990625

Fonte: Autor. χ^2 : Qui-Quadrado, MVS: Máquina de Vetor de Suporte, RNA: Redes Neurais Artificiais, FA: Florestas Aleatórias, RL: Regressão Logística.

Tabela 9 - Resultados dos pipelines que usaram RNA na primeira etapa.

	χ^2 /MVS	χ^2 /RNA	χ^2 /FA	RL/MVS	RL/RNA	RL/FA
0	0,95	0,509375	0,975	0,95	0,475	0,975
1	0,9375	0,5375	0,971875	0,9375	0,5	0,9875
2	0,978125	0,5	0,96875	0,978125	0,5	0,975
3	0,9625	0,5375	1	0,9625	0,5	1
4	0,921875	0,5	0,9875	0,921875	0,5	0,990625

Fonte: Autor. χ^2 : Qui-Quadrado, MVS: Máquina de Vetor de Suporte, RNA: Redes Neurais Artificiais, FA: Florestas Aleatórias, RL: Regressão Logística.

Tabela 10 - Resultados dos pipelines que usaram RF na primeira etapa.

	χ^2 /MVS	χ^2 /RNA	χ^2 /FA	RL/MVS	RL/RNA	RL/FA
--	---------------	---------------	--------------	--------	--------	-------

0	0,95	0,509375	0,975	0,95	0,475	0,975
1	0,9375	0,5375	0,971875	0,9375	0,5	0,9875
2	0,978125	0,5	0,96875	0,978125	0,5	0,975
3	0,9625	0,5375	1	0,9625	0,5	1
4	0,921875	0,5	0,9875	0,921875	0,5	0,990625

Fonte: Autor. χ^2 : Qui-Quadrado, MVS: Máquina de Vetor de Suporte, RNA: Redes Neurais Artificiais, FA: Florestas Aleatórias, RL: Regressão Logística.

Os resultados variaram entre AUC=0,475 e AUC=0,99. Os pipelines que obtiveram os piores resultados foram: MVS-RL-RNA, RNA-RL-RNA, FA-RL-RNA, com AUC média igual a 0.495. Os que obtiveram os melhores desempenhos foram: MVS- χ^2 -FA, RNA- χ^2 -FA e FA- χ^2 -FA, com AUC média igual a 0,986. Como estes três obtiveram a maior média e apenas um deveria ser escolhido para executar as predições do aplicativo, foi escolhido o pipeline MVS- χ^2 -FA de forma arbitrária.

5.3 - Resultados da validação externa

Na validação externa, modelo escolhido para executar as classificações do aplicativo, obteve acurácia de 90%, porém as classes de etiologia atribuídas para os dados foram apenas U e HAS. O modelo acertou 80% das etiologias HAS, 20% das U e 0% das demais.

5.4 - Resultados da criação do aplicativo

A seguir, os resultados da criação do aplicativo ICH-EPS (*ICH Etiology Prediction System*) serão apresentados.

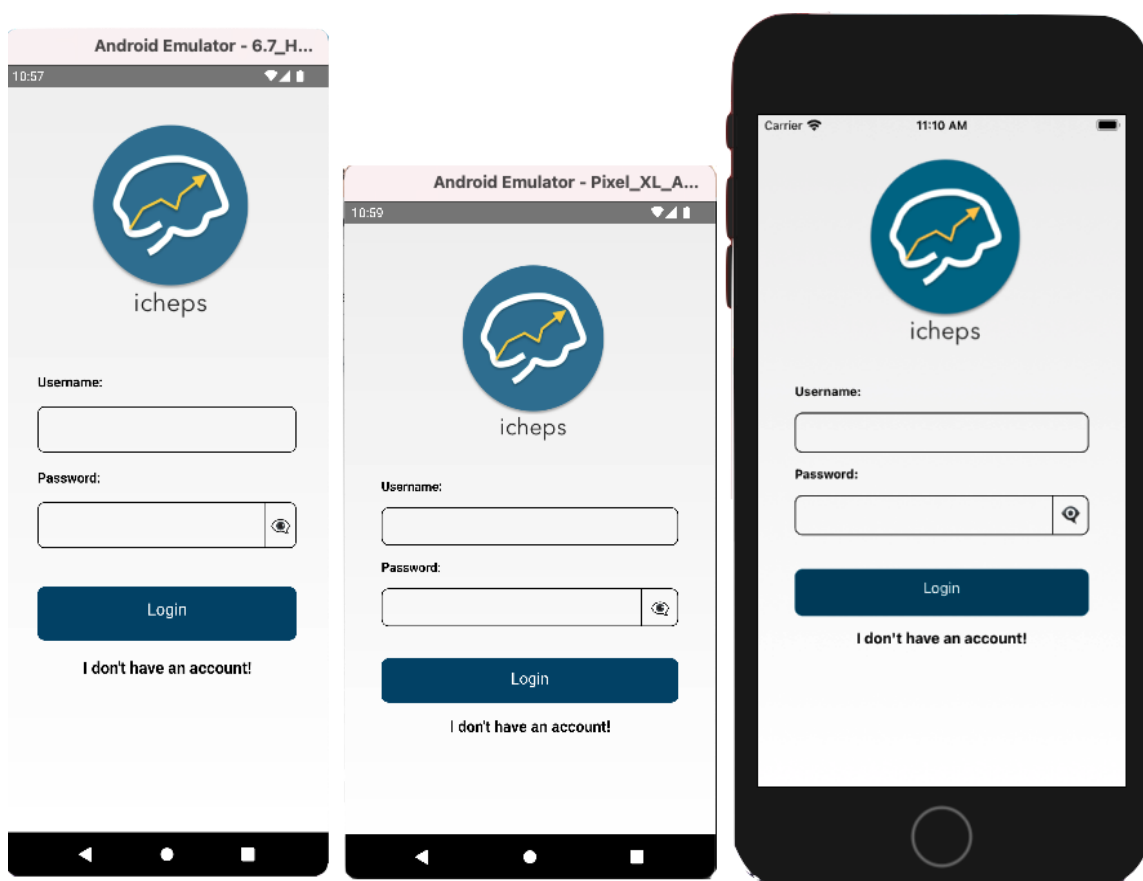
5.4.1 - Interface Gráfica

A interface gráfica do aplicativo foi implementada para garantir a correta comunicação com o usuário, facilitar o manuseio e apresentar uma estética agradável. O aplicativo foi desenvolvido para a língua inglesa, devido ao maior uso desta língua no meio acadêmico e

para facilitar a validação por escalas de usabilidade, que geralmente são desenvolvidas e validadas primeiramente para o inglês.

O aplicativo foi desenvolvido para ser responsivo em diversos aparelhos, ou seja, para que a interface gráfica seja adaptada em telas de tamanhos diferentes. Embora não tenha apresentado erros de responsividade nos aparelhos testados, mais testes devem ser realizados para garantir a correta responsividade no maior número de aparelhos. Na Figura 8, encontra-se um exemplo de como a interface gráfica do aplicativo se comporta em diferentes telas.

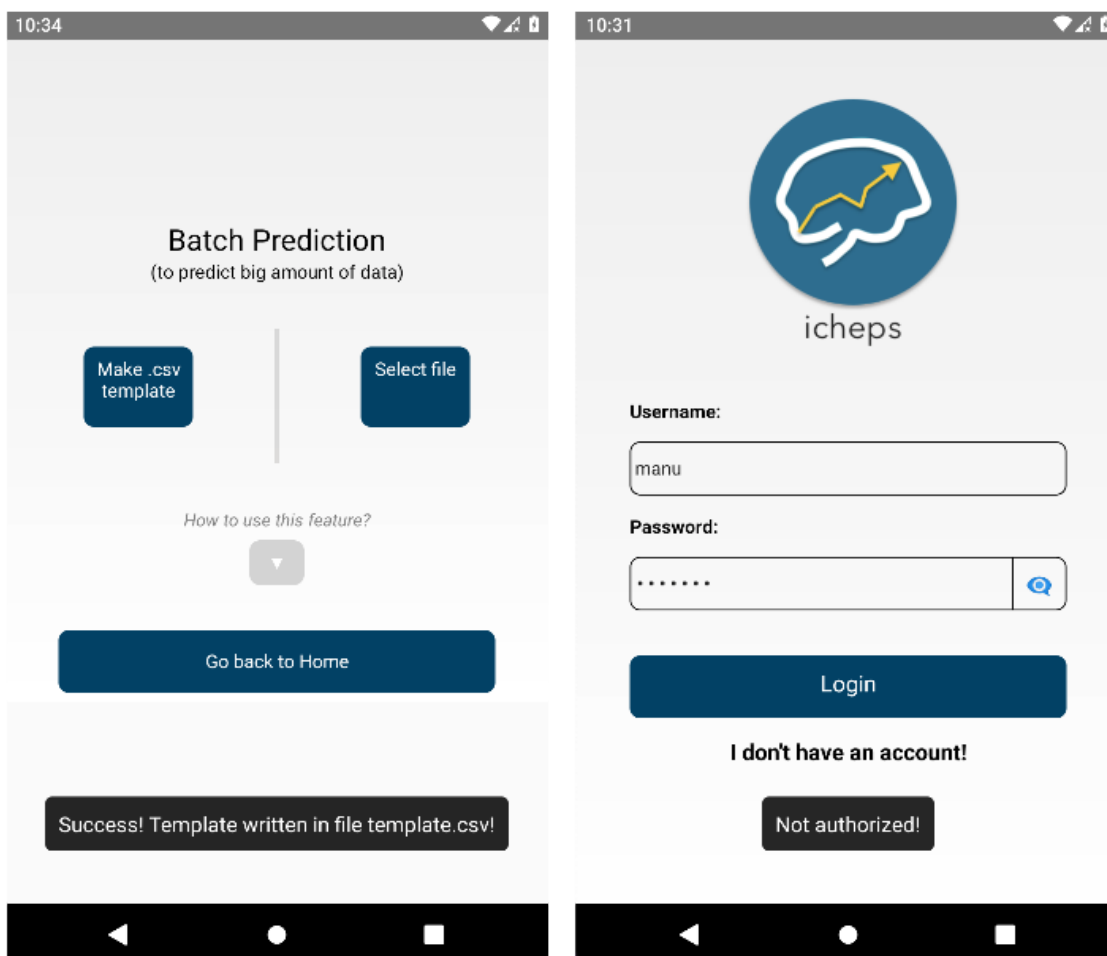
Figura 8 - Interface gráfica do aplicativo em diferentes tamanhos de telas.



Fonte: Autor. A primeira tela (esquerda) é de um simulador do aparelho 6.7 Horizontal FoldIn (Android), a segunda, do Pixel XL (Android) e a terceira, do Iphone 8 (iOS).

Foram utilizadas alertas de mensagens para facilitar a comunicação do aplicativo com o usuário, sendo estas apresentadas na tela com mensagens de sucesso e erro. Exemplos de alertas de erro e sucesso encontram-se na Figura 8.

Figura 8 - Telas com exemplos de alertas de sucesso e erro.



Fonte: Autor.

5.4.2 - Funcionalidades

O aplicativo desenvolvido apresenta duas funcionalidades principais: a classificação da etiologia de dados de um único paciente preenchidos por um formulário, e a classificação de grande quantidade de dados. A primeira funcionalidade tem como objetivo principal fornecer apoio ao diagnóstico e por isso foi idealizada para ter resultados facilmente interpretáveis.

Dessa forma, após o médico preencher o formulário com dados de um paciente, ele pode visualizar as probabilidades do episódio da HIC ser de cada classe SMASH-U. Os dados a serem preenchidos são os dados utilizados pelo modelo de AM selecionado. São eles: sexo, idade, raça, tabagismo, etilismo, pressão sistólica, pressão diastólica, glicemia, diabetes, hipertensão, angiopatia amilóide, volume de hemorragia, localização, ICH, TTPA, TP, anticoagulação (varfarina) e anticoagulação (AAS). Caso haja um problema de rede, os dados

não poderão ser transmitidos ao modelo, e uma mensagem de erro é mostrada ao usuário. As telas do aplicativo referentes a essa funcionalidade podem ser encontradas na Figura 10. A tela que apresenta os resultados (probabilidades do episódio da HIC ser de cada classe SMASH-U) pode encontrar-se amplificada na Figura 11.

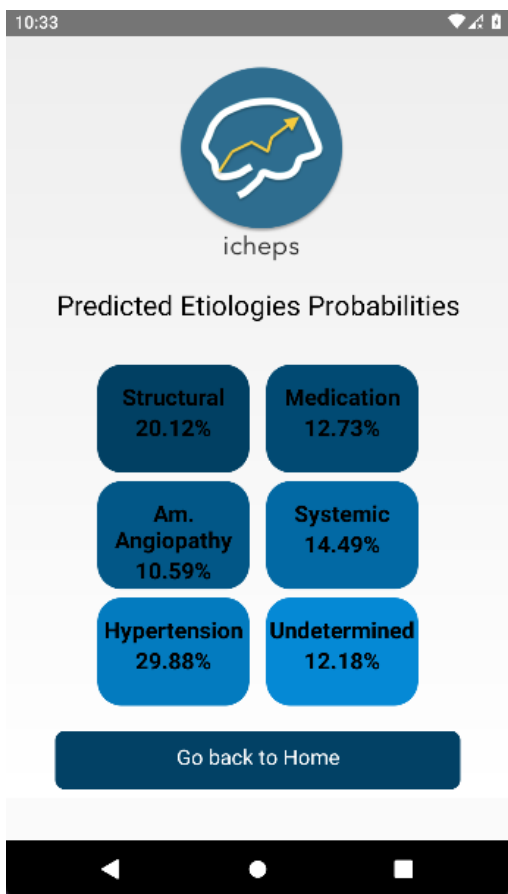
Figura 10 - Telas para a predição de dados de um único paciente.

The figure displays three sequential screenshots of a mobile application interface for data entry. Each screen features a progress indicator at the top and a set of input fields for patient information.

- Screen 1 (Progress: 1/3): Patient's Demographic Data**
 - Age:
 - Gender: Female, Male
 - Ethnicity: White, African American, Indian, Asian, Unknown
 - Does patient drink alcohol?: No, Yes
 - Navigation: Back, Next
- Screen 2 (Progress: 2/3): Patient's Clinical Data**
 - Systolic BP on Admission:
 - Diastolic BP on Admission:
 - Blood Sugar on Admission:
 - APTT -Patient Time:
 - PT -Patient Time:
 - Is the patient hipertense?: No, Yes
 - Does the patient have diabetes?: No, Yes
 - Does the patient take ASA?: No, Yes
- Screen 3 (Progress: 3/3): Patient's Neuroimage Data**
 - Size of Bleeding (ml):
 - ICH:
 - SICH:
 - Local of Bleeding: Frontal, Nucleocapsular, Nucleocapsular/Stem, Occipital, Pons, Thalamus, Parieto-occipital, Temporal, Fronto-temporal, F-P-T, Stem, Capsular-thalamus

Fonte: Autor.

Figura 11 - Tela de resultados da predição de dados de um único paciente.



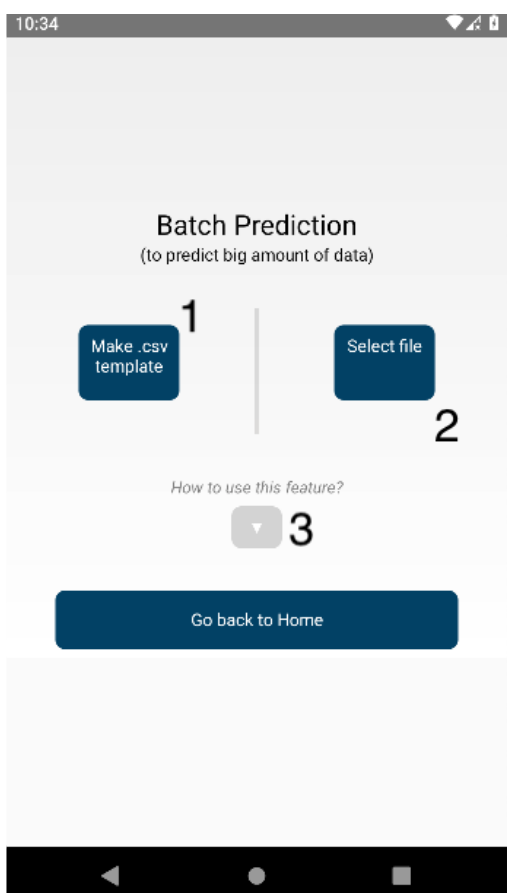
Fonte: Autor.

A segunda funcionalidade desenvolvida foi a classificação de grande quantidade de dados (chamada no app, de *batch prediction*). Essa funcionalidade tem o objetivo de facilitar a predição de uma grande quantidade de dados para ser utilizada majoritariamente no campo da pesquisa. Ela deve ser utilizada da seguinte forma: o médico ou pesquisador faz um *download* de um *template* (tabela vazia, apenas com o cabeçalho), preenche o dado de vários pacientes no *template*, e faz o upload da tabela preenchida para o aplicativo predizer as etiologias. Os dados a serem preenchidos são os mesmos encontrados no formulário da funcionalidade de classificação da etiologia de um único paciente.

Caso os dados estiverem em um formato errado ou incompatível com o template, uma mensagem de erro é apresentada na tela. Para evitar isso, o aplicativo contém instruções para o preenchimento dos dados na mesma tela desta funcionalidade. O usuário precisa clicar na seta para baixo para essas informações aparecerem na tela. Caso a rede de internet esteja lenta ou desconectada, o aplicativo também apresenta uma mensagem de erro.

A tela do aplicativo referente a essa funcionalidade pode ser encontrada na Figura 12. Uma planilha com o resultado de exemplo encontra-se amplificada na Figura 13. A tela com as instruções de apoio para seu uso encontra-se na Figura 14.

Figura 12 - Tela da predição de grande quantidade de dados (*Batch Prediction*).



Fonte: Autor 1 - *Make .csv template*: botão que inicia download do template usado para preencher os dados no formato aceito. 2 - *Select file*: botão que permite ao usuário escolher o arquivo do template preenchido no sistema de arquivos do celular. 3 - *How to use this feature*: botão que abre a visualização da descrição das variáveis e dos tipos de dados aceitos.

Figura 13 - Planilha com resultados para a predição de grande quantidade de dados.

Probabilities	Amyloid Angiopathy	Structural	Hypertensive Angiopathy	Medication	Systemic	Undetermined
patient_0	0.07456034	0.14120163	0.28941693	0.11368588	0.1137612	0.26737402
patient_1	0.0748745	0.18435036	0.26767581	0.11406172	0.13542884	0.22360877
patient_2	0.10628227	0.03368149	0.4965005	0.07767391	0.13509894	0.15076289

Fonte: Autor.

Figura 14 - Instruções para utilização da funcionalidade de predição de grande quantidade de dados.

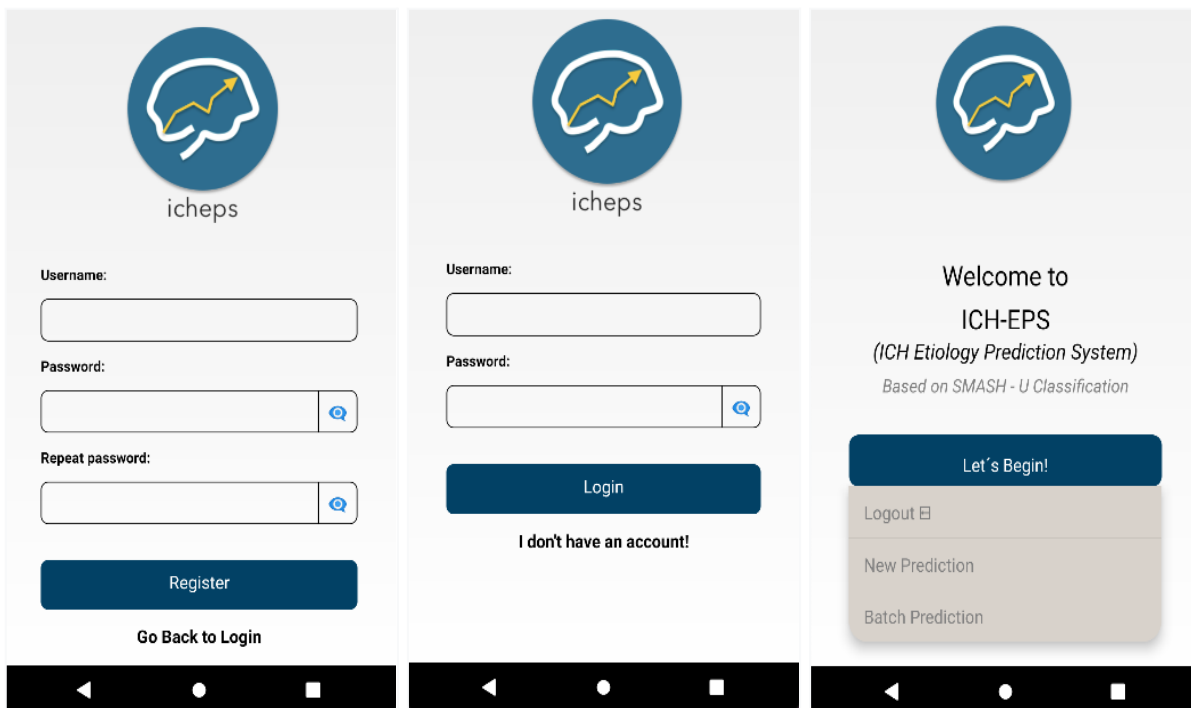


Fonte: Autor.

Além das funcionalidades principais, o aplicativo desenvolvido tem as funcionalidades de acesso, que são “Login” e “Cadastro” e “Login”. Para se cadastrar, o usuário deve inserir um nome de usuário, uma senha e a confirmação da senha. Caso haja um problema de rede, ele receberá uma mensagem de erro, em caso contrário ele receberá uma mensagem de sucesso. Caso o usuário já exista no sistema, ele receberá uma mensagem avisando que o usuário já foi cadastrado. Para fazer login, o usuário deve entrar com um nome de usuário e senha cadastrados. Caso esses dados não existam no sistema ou haja um problema de rede, o usuário receberá uma mensagem de erro.

Após o login com sucesso do usuário cadastrado, a tela “Home” aparece e permite que o usuário escolha a funcionalidade desejada. São três opções: logout, predição para 1 paciente (*new prediction*) e predição em lote (Figura 15).

Figura 15 - Tela de Cadastro de Usuário, de Login e tela Inicial (*Home*).



Fonte: Autor.

O aplicativo foi testado em emuladores e aparelhos, obtendo bom funcionamento em todos eles. Contudo, devem ser executados mais testes para validar o correto funcionamento do aplicativo na maioria dos dispositivos utilizados atualmente. Os emuladores e aparelhos que foram usados para teste podem ser observados nas Tabela 11 e 12.

Tabela 11 - Emuladores e aparelhos usados para teste, com as respectivas versões do sistema operacional neles instalados no sistema operacional Android.

Emulador ou Aparelho Físico	Modelo	Versão do Software
Emulador	Android Emulator (versão simplificada para M1)	30
Emulador	Galaxy Nexus	31
Emulador	Galaxy Nexus	30
Emulador	Galaxy Nexus	29
Emulador	Galaxy Nexus	28
Emulador	Pixel 4 XL	31

Emulador	Nexus S	31
Emulador	6.7 Horizontal Foldin	31

Fonte: Autor.

Tabela 12 - Emuladores e aparelhos usados para teste, com as respectivas versões do sistema operacional neles instalados no sistema operacional iOS.

Emulador ou Aparelho Físico	Modelo	Versão do Software
Emulador	Iphone 12 PRO MAX	14.4
Emulador	Iphone 12 MINI	14.4
Emulador	Iphone 8	14.4
Emulador	Iphone SE	14.4
Emulador	Iphone 11	14.4
Aparelho físico	iphone 6s	15.4

Fonte: Autor.

5.5 - Discussão

A inteligência artificial (IA) vem ocasionando mudanças importantes na área da saúde, tornando o diagnóstico médico mais rápido e preciso em diversas situações. Dentro da área da IA, métodos de AM permitem a classificação, o agrupamento e a redução de dimensionalidade, agilizando o processamento de grandes quantidades de dados biomédicos e auxiliando no diagnóstico clínico de diversas doenças. Adicionalmente, a integração de modelos de AM às aplicações móveis favoreceu a acessibilidade à informação clínica e a serviços médicos de qualidade.

Como objetivo principal, este trabalho almejou a construção de um aplicativo médico para predição da etiologia de episódios de HIC, com o intuito de facilitar classificação de grande quantidade de dados e dar apoio à investigação clínica. Para isso, foram comparados pipelines compostos pelas etapas de rotulação de dados faltantes, seleção de atributos e classificação. O melhor pipeline foi utilizado para servir ao aplicativo desenvolvido.

Como objetivo secundário, este trabalho procurou avaliar o uso de dados clínicos de HIC em modelos semi-supervisionados de AM. A vantagem de usar modelos semi-supervisionados para a classificação etiológica da HIC é principalmente validar seu uso neste contexto. Dessa forma, pretende-se endossar o desenvolvimento de modelos semi-supervisionados, que necessitam de menor quantidade de dados rotulados por especialistas, e assim tornar os processos de pesquisa e produção de softwares de AM mais ágeis e eficientes.

Para treinar os pipelines, foram coletados dados de forma retrospectiva de 66 pacientes com HIC consecutivos do Registro de Acidente Vascular Cerebral de Ribeirão Preto (REAVR).

Observando-se as estatísticas descritivas dos dados, os valores das variáveis coletadas aparentam ter grande variabilidade dentre as classes etiológicas. Essa informação ressalta a importância de determinar a etiologia de um episódio de HIC, que consistem principalmente em :

- Definir o manejo do paciente. Essa conduta é uma possível forma de superar as limitações de prognóstico e terapias para o AVCH.
- Auxiliar na estratificação de grupos de dados nas pesquisas clínicas, já que há relatos de pesquisas que falharam devido à ausência de análise de subgrupos. Por exemplo, a maioria dos estudos epidemiológicos falhou em associar a

concentração sérica dos níveis de colesterol com AVC porque não procuraram essa associação no subtipo aterotrombótico do AVC isquêmico. Já os estudos que investigaram essa associação separadamente no subtipo aterotrombótico, encontraram evidência de associação. [1]

Quanto aos dados coletados, as variáveis preditoras foram escolhidas devido a sua relação com certas classes de etiologia da HIC ou por serem fatores de risco conhecidos para HIC, conforme encontrado na literatura. As variáveis raça, gênero, tabagismo, etilismo e uso de drogas foram escolhidas por serem fatores de risco comumente associados à HIC [23,24,25,26]. A hipertensão arterial sistêmica (HAS) e a angiopatia amilóide estão associadas à etiologias de causa primária da HIC. Devido a isso, a medida da pressão arterial juntamente com os diagnósticos para hipertensão e angiopatia amilóide (ausente ou presente) foram utilizados como variáveis preditoras [23]. Já as coagulopatias, uso de anticoagulantes orais, diabetes, vasculopatias e aneurismas estão associadas à etiologias secundárias da HIC, e por isso foram incluídas no grupo de variáveis preditoras, assim como a medição da glicemia e testes de triagem da coagulação (TP e TTPA) de admissão [27,28]. A localização, lado e volume da lesão no exame de imagem foram escolhidos porque também podem indicar a etiologia do sangramento. Por exemplo, o diagnóstico mais provável quando a hemorragia está localizada nos núcleos da base é o hematoma hipertensivo [8]. Adicionalmente, as escalas SICH e ICH têm mostrado sucesso ao prever o risco de um determinado paciente com HIC de possuir uma determinada etiologia, e por isso também foram escolhidas como variáveis preditoras.

Ao comparar os dados coletados com os encontrados em Meretoja et al (2012), as proporções de pacientes nos subgrupos de etiologia foram similares para alguns dos subgrupos (estrutural e sistêmico), apresentando diferenças nos demais. Por exemplo, enquanto no estudo realizado por Meretoja et al (2012) 20% dos pacientes estavam no subgrupo de angiopatia amilóide e 14% no de anticoagulação, no presente estudo os pacientes de cada um destes grupos representam apenas 4.6%. Em Meretoja et al (2012), o subgrupo hipertensivo representou apenas 35%, enquanto neste estudo apresentou quase 70%. O grupo de etiologia indefinida também apresentou diferença em ambos os estudos - 13% neste estudo versus 21% em Meretoja et al (2012).

Os pacientes com etiologia primária (hipertensiva ou por angiopatia amilóide) constituíram 83,3 % do total (55 pacientes) e os pacientes com etiologias secundárias constituíram 16,7 %. Esses dados são compatíveis com os mencionados por Pontes et al

(2009), no qual se afirma que as etiologias primárias estão presentes em 80-85% dos casos e as secundárias, em 15-25% dos casos.

Fazendo novamente uma comparação com Meretoja et al (2012), foram encontradas diferenças na distribuição dos dados agrupados por etiologia em ambos os estudos. Uma tabela comparativa entre as estatísticas descritivas dos dados agrupados deste estudo e de Meretoja et al (2012), pode ser encontrada na Tabela 13 (contendo atributos em comum para ambos os estudos).

Tabela 13 - Tabela comparativa com Meretoja et al (2012) entre as estatísticas descritivas agrupadas por etiologia para os atributos em comum.

Meretoja et al (2012)

	ESTRU- TURAL	MEDI- CAÇÃO	AAC	SISTÊMICO	HAS	U
Idade	55 (38–64)	76 (68–82)	73 (66–80)	60 (51–68)	66 (57–78)	62 (54–73)
Sexo (Masculino)	22 (44%)	88 (62%)	100 (49%)	32 (67%)	204 (58%)	136 (64%)
Hipertensão	14 (28%)	111 (78%)	117 (57%)	24 (50%)	354 (100%)	17 (8%)
Diabetes	4 (8%)	31 (22%)	26 (13%)	7 (15%)	64 (18%)	10 (5%)
Anticoagulação (AAS)	2 (4%)	122 (85%)	2 (1%)	1 (2%)	5 (1%)	0 (0%)
Volume da hemorragia, mL	2,8 (0,8–5,7)	14 (4,8–46)	14 (4,5–34,4)	9,8 (3,0–32)	9,3 (3,6–22)	9,1 (4,4–24,6)
Presente estudo						
	ESTRU- TURAL	MEDI- CAÇÃO	AAC	SISTÊMICO	HAS	U
Idade	57 (46,5–60,0)	61 (59,5–62,5)	71 (66,5–71,0)	62 (47,5–63,5)	57 (50,0–66,0)	78 (69,0–85,0)
Sexo (Masculino)	3 (100%)	0 (0%)	2 (67%)	1 (33%)	31(67%)	3 (33%)
Hipertensão	2 (67%)	1 (50%)	1 (33%)	2 (67%)	33 (72%)	8 (89%)
Diabetes	0 (0%)	0 (0%)	0 (0%)	2 (67%)	9 (20)%	1 (11%)
Anticoagulação (varfarina)	0 (0%)	1 (50%)	0 (0%)	1 (33%)	0 (0%)	3 (33%)
Volume de hemorragia, cm ³	10,5 (6,05-60,0)	20,225 (10,2875–62,5)	63 (31,8-71,0)	1,79 (1,395-63,5)	7,95 (2,6–66,0)	42 (9,0–85,0)

Fonte: Autor. Variáveis categóricas : n (%). Variáveis numéricas : mediana (variação interquartil). AAS: ácido acetilsalicílico, AAC: angiopatia amilóide, HAS: hipertensão, U: indefinido.

Ao compararmos os resultados dos pipelines construídos com os de trabalhos científicos que aplicaram métodos de AM em dados clínicos na predição de diagnósticos/condições médicas, obtemos resultados consistentes com a literatura. Nos trabalhos avaliados na revisão bibliográfica encontrada na seção 2, os algoritmos obtiveram AUC variando entre 0,58 e 0,83 e acurácia variando entre 36% e 99.5%. Os resultados do presente estudo, que obteve AUC variando entre 0,495 e 0,986, aparentam ser similares. Quanto à validação externa, o resultado foi inferior ao esperado, podendo ter ocorrido um *underfitting* do modelo, provavelmente devido ao número insuficiente de dados em cada classe do conjunto de treino.

A relevância deste trabalho está na possível incorporação do aplicativo na prática clínica. Além disso, este é um dos primeiros estudos que utilizou métodos de AM para predição da etiologia de episódios de HIC, podendo ter seus resultados comparados com futuros estudos para desenvolver um conhecimento específico à essa aplicação. Esta relevância reside no fato de estudos experimentais serem importantes na área de AM devido à inexistência de uma análise única e independente de dados que permita analisar o desempenho de um modelo.

Adicionalmente, os estudos que validam as escalas SMASH-U e H-ATOMIC utilizaram dados de populações diferentes da população brasileira, logo a confiabilidade entre avaliadores nestes estudos pode variar da encontrada na prática clínica no Brasil. Por isso estudos específicos para a população brasileira, como o presente estudo, são necessários para caracterizar populações (no caso, a da cidade de Ribeirão Preto nos anos de 2018 e 2019) quanto à etiologia dos episódios de HIC. Além disso, este trabalho propõe uma ferramenta de classificação da etiologia HIC com o intuito de auxiliar na prática clínica, para que os resultados da investigação etiológica sejam mais precisos e específicos à população em que é utilizada.

Como limitação do trabalho, pode ser citada a pequena quantidade de dados, que resultou em um pequeno número de dados de cada classe etiológica para treinar o modelo, tendo ocorrido um possível *underfitting*. Para superarmos essa limitação, planejamos aumentar a coleta de dados e retreinar os modelos. Além disso, a remoção de classes etiológicas para teste dos modelos semi supervisionados será realizada de forma estratificada (isto não pode ser feito no presente trabalho devido ao pequeno número de dados em cada classe, o que resultou em todos os rótulos de etiologias removidos serem da classe etiológica).

5.6 - Sugestões para trabalhos futuros

Para a evolução do aplicativo desenvolvido e melhoria no modelo de AM utilizado, os seguintes trabalhos futuros são sugeridos:

- coleta de maior número de dados em um estudo multicêntrico;
- tornar o modelo de AM mais dinâmico, através da incorporação automática de dados inseridos no aplicativo
- ampliar as funcionalidades do app com processamento e extração de neuroimagens para predição da etiologia e com cálculo automático de escalas neurológicas SICH e ICH.
- validação do aplicativo para que este seja incorporado na prática clínica, possivelmente com uma escala de validação de aplicativos como a *Mobile Application Rating Scale* (MARS)

6 - Conclusões

- Foram desenvolvidos e comparados 18 pipelines compostos pelas tarefas de rotulação por métodos semi supervisionados, seleção de atributos e classificação.
- Os pipelines de modelos semi-supervisionados obtiveram desempenho parecido com estudos similares encontrados na literatura para dados clínicos aplicados à classificação de uma condição médica: os algoritmos encontrados na literatura obtiveram AUC variando entre 0,58 e 0,83 e os do presente estudo obtiveram AUC variando entre 0,495 e 0,986.
- O melhor pipeline (AUC média = 0,986) foi selecionado para servir ao aplicativo.
- O aplicativo de celular foi desenvolvido, apresentando interface de fácil usabilidade e desempenho satisfatório para ambas as plataformas iOS e Android. Duas funcionalidades principais foram incorporadas para facilitar classificação da etiologia de pacientes na prática clínica e na pesquisa.

7 - Referências

- 1 AMARENCO, P. et al. Classification of stroke subtypes. *Cerebrovasc Dis*, v. 27, n. 5, p. 493-501, 2009. ISSN 1421-9786. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/19342825> >.
- 2 ORGANIZAÇÃO PAN AMERICANA DA SAÚDE. Rede interagencial de informações para a Saúde. Disponível em: <https://opas.org.br/>.
- 3 STRONG, K.; MATHERS, C.; BONITA, R. Preventing stroke: saving lives around the world. *The Lancet Neurology*, v. 6, n. 2, p. 182-187, 2007. ISSN 1474-4422.
- 4 DENNIS, M. Charles Warlow, Cathie Sudlow, Martin Dennis, Joanna Wardlaw, Peter Sandercock. *THE LANCET*, v. 362, p. 1211-24, 2003.
- 5 VAN ASCH, C. J. et al. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *The Lancet Neurology*, v. 9, n. 2, p. 167-176, 2010. ISSN 1474-4422.
- 6 POON, M. T. C.; FONVILLE, A. F.; SALMAN, R. A.-S. Long-term prognosis after intracerebral haemorrhage: systematic review and meta-analysis. *Journal of Neurology, Neurosurgery & Psychiatry*, v. 85, n. 6, p. 660-667, 2014. ISSN 0022-3050.
- 7 POUNGVARIN, N.; VIRIYAVEJAKUL, A.; KOMONTRI, C. Siriraj stroke score and validation study to distinguish supratentorial intracerebral haemorrhage from infarction. *British Medical Journal*, v. 302, n. 6792, p. 1565-1567, 1991. ISSN 0959-8138.
- 8 MASSARO, A. et al. Clinical discriminators of lobar and deep hemorrhages: the Stroke Data Bank. *Neurology*, v. 41, n. 12, p. 1881-1881, 1991. ISSN 0028-3878.

- 9 LAISSY, J. P. et al. Spontaneous intracerebral hematomas from vascular causes. Predictive value of CT compared with angiography. *Neuroradiology*, v. 33, n. 4, p. 291-5, 1991. ISSN 0028-3940. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/1922741> >.
- 10 KOTHARI, R. U. et al. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke*, v. 27, n. 8, p. 1304-1305, 1996. ISSN 0039-2499.
- 11 GEBEL, J. M. et al. Comparison of the ABC/2 estimation technique to computer-assisted volumetric analysis of intraparenchymal and subdural hematomas complicating the GUSTO-1 trial. *Stroke*, v. 29, n. 9, p. 1799-1801, 1998. ISSN 0039-2499.
- 12 BROTT, T. et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*, v. 20, n. 7, p. 864-870, 1989. ISSN 0039-2499.
- 13 GOLDSTEIN, L. B.; CHILUKURI, V. Retrospective assessment of initial stroke severity with the Canadian Neurological Scale. *Stroke*, v. 28, n. 6, p. 1181-1184, 1997. ISSN 0039-2499.
- 14 BRODERICK, J. et al. Lobar hemorrhage in the elderly. The undiminishing importance of hypertension. *Stroke*, v. 24, n. 1, p. 49-51, 1993. ISSN 0039-2499.
- 15 HEMPHILL III, J. C. et al. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. *Stroke*, v. 32, n. 4, p. 891-897, 2001. ISSN 0039-2499.
- 16 DAVIS, S. et al. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. *Neurology*, v. 66, n. 8, p. 1175-1181, 2006. ISSN 0028-3878.
- 17 SALOHEIMO, P. et al. Regular aspirin-use preceding the onset of primary intracerebral hemorrhage is an independent predictor for death. *Stroke*, v. 37, n. 1, p. 129-133, 2006. ISSN 0039-2499.
- 18 MAYER, S. A.; RINCON, F. Treatment of intracerebral haemorrhage. *The Lancet Neurology*, v. 4, n. 10, p. 662-672, 2005. ISSN 1474-4422.

- 19 FRANKE, C. et al. Prognostic factors in patients with intracerebral haematoma. *Journal of Neurology, Neurosurgery & Psychiatry*, v. 55, n. 8, p. 653-657, 1992. ISSN 0022-3050.
- 20 GREENBERG, S. M.; BRIGGS, M. E.; HYMAN, B. T.; KOKORIS, G. J. et al. Apolipoprotein E ϵ 4 is associated with the presence and earlier onset of hemorrhage in cerebral amyloid angiopathy. *Stroke*, 27, n. 8, p. 1333-1337, 1996.
- 21 MERETOJA, A. et al. SMASH-U: a proposal for etiologic classification of intracerebral hemorrhage. *Stroke*, v. 43, n. 10, p. 2592-7, Oct 2012. ISSN 1524-4628. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22858729> >.
- 22 MARTÍ-FÀBREGAS, J. et al. The H-ATOMIC Criteria for the Etiologic Classification of Patients with Intracerebral Hemorrhage. *PLOS ONE*, v. 11, n. 6, p. e0156992, 2016. Disponível em: < <https://doi.org/10.1371/journal.pone.0156992> >.
- 23 QURESHI, A. I.; TUHRIM, S.; BRODERICK, J. P.; BATJER, H. H. et al. Spontaneous intracerebral hemorrhage. *New England Journal of Medicine*, 344, n. 19, p. 1450-1460, 2001.
- 24 ABBOTT, R. D. et al. Risk of stroke in male cigarette smokers. *N Engl J Med*, v. 315, n. 12, p. 717-20, Sep 18 1986. ISSN 0028-4793. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/3748080> >.
- 25 JUVELA, S.; HILLBOM, M.; PALOMÄKI, H. Risk factors for spontaneous intracerebral hemorrhage. *Stroke*, 26, n. 9, p. 1558-1564, 1995.
- 26 LAKE, C. R. et al. Adverse drug effects attributed to phenylpropanolamine: a review of 142 case reports. *Am J Med*, v. 89, n. 2, p. 195-208, Aug 1990. ISSN 0002-9343. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/2200264> >.

- 27 ROSAND, J. et al. Warfarin-associated hemorrhage and cerebral amyloid angiopathy: a genetic and pathologic study. *Neurology*, v. 55, n. 7, p. 947-51, Oct 10 2000. ISSN 0028-3878. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/11061249> >.
- 28 TOYODA, K. et al. Antiplatelet therapy contributes to acute deterioration of intracerebral hemorrhage. *Neurology*, v. 65, n. 7, p. 1000-4, Oct 11 2005. ISSN 1526-632X. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/16217049> >.
- 29 GOMES, L. A história dos apps, uma breve linha do tempo. Disponível em: < <https://www.showmetech.com.br/a-historia-dos-apps/> >.
- 30 MOSA, A. S.; YOO, I.; SHEETS, L. A systematic review of healthcare applications for smartphones. *BMC Med Inform Decis Mak*, v. 12, p. 67, Jul 10 2012. ISSN 1472-6947. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/22781312> >.
- 31 WALLACE, S.; CLARK, M.; WHITE, J. 'It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. *BMJ Open*, v. 2, n. 4, 2012. ISSN 2044-6055. Disponível em: <<https://www.ncbi.nlm.nih.gov/pubmed/22923627>>.
- 32 DIVALL, P.; CAMOSSO-STEFINOVIC, J.; BAKER, R. The use of personal digital assistants in clinical decision making by health care professionals: a systematic review. *Health Informatics J*, v. 19, n. 1, p. 16-28, Mar 2013. ISSN 1741-2811. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/23486823> >.
- 33 GARGENTA, M. *Learning Android*. 2011.
- 34 SINICKI, A. *Developing for Android vs developing for iOS—in 5 rounds*. Android Authority, 2017.
- 35 RAJPUT, M. *Top Mobile App Development Framework in 2019*. 2018. Disponível em: . Acesso em: 18 abr. 2019. ROCKCONTENT.

- 36 HEITKÖTTER, H.; HANSCHKE, S.; MAJCHRZAK, T. A. Evaluating cross-platform development approaches for mobile applications. International Conference on Web Information Systems and Technologies: Springer, 2012. 120-138 p.
- 37 HANSSON, N.; VIDHALL, T. Effects on performance and usability for cross-platform application development using React Native 2016.
- 38 ORACLE. O que é inteligência artificial (IA)? Saiba mais sobre Inteligência Artificial. Disponível em: < <https://www.oracle.com/br/artificial-intelligence/what-is-ai/> >.
- 39 GOMES, D. D. S. Inteligência Artificial: conceitos e aplicações. Olhar Científico. v1, n. 2, p. 234-246, 2010.
- 40 WANG, F.; PREININGER, A. AI in Health: State of the Art, Challenges, and Future Directions. Yearb Med Inform, v. 28, n. 1, p. 16-26, Aug 2019. ISSN 2364-0502. Disponível em: < <https://www.ncbi.nlm.nih.gov/pubmed/31419814> >.
- 41 JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. Science, 349, n. 6245, p. 255-260, 2015.
- 42 MARSLAND, S. Machine Learning, CRC Press. 2014.
- 43 SYDORENKO, I. How to Use Unlabeled Data in Machine Learning. 2020. Disponível em: <https://labeyourdata.com/articles/unlabeled-data-in-machine-learning>.
- 44 COZMAN, F. G.; COHEN, I.; CIRELO, M. C. Semi-supervised learning of mixture models. ICML, 2003. 24 p.
- 45 SUTTON, R. S.; BARTO, A. G. Reinforcement learning: An introduction. MIT press, 2018. ISBN 0262352702.
- 46 IZBICKI, R.; DOS SANTOS, T. M. Aprendizado de máquina: uma abordagem estatística. Rafael Izbicki, 2020. 6500024109.

- 47 BHANDARI , A. AUC-ROC Curve in Machine Learning Clearly Explained. 2020. Disponível em: <https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/#:~:text=The%20Area%20Under%20the%20Curve,the%20positive%20and%20negative%20classes.>
- 48 SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. *Information processing & management*, v. 45, n. 4, p. 427-437, 2009. ISSN 0306-4573.
- 49 MARIANO, DIEGO; PAZ, F. J. . *Data Mining*. 1. ed. Porto Alegre: Sagah, 2020
- 50 AZANK, F.; KORZUNE GURGEL, G. Dados Desbalanceados — O que são e como lidar com eles. 2020. Disponível em: [https://medium.com/turing-talks/dados-desbalanceados-o-que-s%C3%A3o-e-como-evit%C3%A1-los-43df4f49732b.](https://medium.com/turing-talks/dados-desbalanceados-o-que-s%C3%A3o-e-como-evit%C3%A1-los-43df4f49732b)
- 51 CHANDRASHEKAR, G.; SAHIN, F. A survey on feature selection methods. *Computers & Electrical Engineering*, v. 40, n. 1, p. 16-28, 2014. ISSN 0045-7906.
- 52 GUYON, I.; ELISSEEFF, A. An introduction to variable and feature selection. *Journal of machine learning research*, v. 3, n. Mar, p. 1157-1182, 2003.
- 53 LIU, H.; SETIONO, R. A probabilistic approach to feature selection-a filter solution. *ICML: Citeseer*, 1996. 319-327 p.
- 54 LAZAR, C. et al. A survey on filter techniques for feature selection in gene expression microarray analysis. *IEEE/ACM transactions on computational biology and bioinformatics*, v. 9, n. 4, p. 1106-1119, 2012. ISSN 1545-5963.
- 55 STONE, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, v. 36, n. 2, p. 111-133, 1974. ISSN 0035-9246.

- 56 CARUANA, R.; NICULESCU-MIZIL, A., 2006a, An empirical comparison of supervised learning algorithms. 161-168.
- 57 FERNÁNDEZ-DELGADO, M.; CERNADAS, E.; BARRO, S.; AMORIM, D. Do we need hundreds of classifiers to solve real world classification problems? The journal of machine learning research, 15, n. 1, p. 3133-3181, 2014a.
- 58 Redes neurais. Disponível em: <https://www.ibm.com/br-pt/cloud/learn/neural-networks>.
- 59 O que são os pipelines do Azure Machine Learning? , 2022. Disponível em: <https://docs.microsoft.com/pt-br/azure/machine-learning/concept-ml-pipelines>.
- 60 ZHANG, H. et al. Artificial intelligence–based traditional chinese medicine assistive diagnostic system: Validation study. JMIR medical informatics, v. 8, n. 6, p. e17608, 2020.
- 61 NI, Y. et al. Towards phenotyping stroke: Leveraging data from a large-scale epidemiological study to detect stroke diagnosis. PloS one, v. 13, n. 2, p. e0192586, 2018. ISSN 1932-6203.
- 62 ZEIBERG, D. et al. Machine learning for patient risk stratification for acute respiratory distress syndrome. PloS one, v. 14, n. 3, p. e0214465, 2019. ISSN 1932-6203.
- 63 COOMBES, C. E.; COOMBES, K. R.; 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. ISSN 1472-6947.
- 64 ANOOJ, P. Implementing decision tree fuzzy rules in clinical decision support system after comparing with fuzzy based and neural network based systems. 2013 International Conference on IT Convergence and Security (ICITCS): IEEE, 2013. 1-6 p.

65 HASHI, E. K.; ZAMAN, M. S. U.; HASAN, M. R. An expert clinical decision support system to predict disease using classification techniques. 2017 International conference on electrical, computer and communication engineering (ECCE): IEEE, 2017. 396-400 p.

66 ZHANG, X.; SUN, Y., 2018, Breast cancer risk prediction model based on C5. 0 algorithm for postmenopausal women. IEEE. 321-325.

67 ZHANG, Y. et al. Ensemble learning based postpartum hemorrhage diagnosis for 5g remote healthcare. IEEE Access, v. 9, p. 18538-18548, 2021. ISSN 2169-3536.

68 THALLAM, C. et al. Early Stage Lung Cancer Prediction Using Various Machine Learning Techniques. 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA): IEEE, 2020. 1285-1292 p.

69 LEE, H. D. Seleção e construção de features relevantes para o aprendizado de máquina. 2000. -, Universidade de São Paulo.

70 KNUDSEN, K. A.; ROSAND, J.; KARLUK, D.; GREENBERG, S. M. Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. **Neurology**, 56, n. 4, p. 537-539, 2001.

8 - Anexos

etiologia	AAC	ESTRUTURAL	HAS	MEDICAÇÃO	SISTÊMICO	U
Vasculopatias	0% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE
Angiopatia Amilóide	33% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE	0% PRESENTE
Aneurisma	0% PRESENTE	0% PRESENTE	< 1% PRESENTE	0% PRESENTE	0% PRESENTE	22% PRESENTE
Volume	63 (31,8-96,50) cm ³	10,5 (6,05-18,60) cm ³	7,9 (2,6-13,45) cm ³	20,225 (10,29-30,16) cm ³	1,79 (1,39-11,54) cm ³	41 (9-81) cm ³
SICH	SICH 2 : 1 paciente	-----	SICH 1 : 11 pacientes SICH 2 : 9 pacientes SICH 0 : 1 paciente SICH 4 : 1 paciente	-----	SICH 4 : 1 paciente SICH 5 : 1 paciente	SICH 1 : 3 pacientes SICH 2 : 1 paciente SICH 3 : 1 paciente
ICH	ICH 0 : 1 paciente ICH 4 : 1 paciente	ICH 0 - 1 paciente	ICH 1 : 15 pacientes ICH 0 : 9 pacientes ICH 2 : 5 pacientes ICH 3 : 4 pacientes	-----	ICH 0 : 1 paciente ICH 1 : 1 paciente	ICH 0 : 1 paciente ICH 1 : 1 paciente ICH 3 : 1 paciente ICH 4 : 1 paciente ICH 6 : 1 paciente
Lado	33% DIREITO 67% ESQUERDO 0% BILATERAL	0% DIREITO 100% ESQUERDO 0% BILATERAL	42% DIREITO 42% ESQUERDO 16% BILATERAL	50% DIREITO 50% ESQUERDO 0% BILATERAL	33% DIREITO 67% ESQUERDO 0% BILATERAL	55% DIREITO 33% ESQUERDO 12% BILATERAL

etiologia	AAC	ESTRUTURAL	HAS	MEDICAÇÃO	SISTÊMICO	U
Localização	FRONTAL : 1 paciente FRONTOTEMPORAL : 1 paciente OCCIPITAL : 1 paciente	FRONTOPARIETAL : 1 paciente TALÂMICO : 1 paciente TEMPORAL : 1 paciente	NUCLEOCAPSULA R : 24 pacientes TALAMOCAPSULR : 9 pacientes TALÂMICO : 4 pacientes TRONCO : 4 pacientes PONTE : 3 pacientes NUCLEOCAPSULA R/TRONCO : 1 paciente OCCIPITAL : 1 paciente	NUCLEOCAPSULAR : 1 paciente PARIETAL : 1 paciente	FRONTAL : 3 pacientes	NUCLEOCAPSULA R : 3 pacientes OCCIPITAL : 2 pacientes F-P-T : 1 paciente PARIETOOCCIPITAL : 1 paciente TEMPOROPARIETAL : 1 paciente TRONCO : 1 paciente
TTPA	22,90 (21,05-26,20)	26,40 (25,60-29,35)	29,50 (25,80-34,55)	52,1 (51,1-52,1)	25,90 (25,65-25,90)	30,95 (25,70-43,85)
TP	10,3 (9,5-10,75)	10,7 (10,65-12,30)	13,1 (10,3-14,50)	77,5 (77,5-77,5)	11,20 (10,9-11,25)	12,45 (11,125-28,25)
Varfarina	0% ANTICOAG.	0% ANTICOAG.	0% ANTICOAG.	50% ANTICOAG.	33% ANTICOAG.	33% ANTICOAG.
AAS	0% ANTICOAG.	33% ANTICOAG.	20% ANTICOAG.	0% ANTICOAG.	33% ANTICOAG.	33% ANTICOAG.

Fonte: Autor. AAC: angiopatia amilóide, HAS: hipertensão, U: indefinido, AAS: ácido acetilsalicílico, dp: desvio padrão. Variáveis categóricas : %. Variáveis numéricas normais : média e desvio padrão. Variáveis numéricas não-normais: mediana (variação interquartil). Células representadas com “-” ocorrem quando todos os dados de uma determinada variável em uma determinada classe são faltantes.