Journal of Health Informatics

J Health Inform. 2026;18:e1797
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1797

ARTIGO COMPLETO
CBIS 2026

GeriRisco: sistema web com aprendizado de máquina para triagem de quedas em idosos

GeriRisco: web system with machine learning for fall risk screening in older adults

Mariana Lúcia Branco Zeitune[1] ✉, Karla Tereza Figueiredo Leite[1], Alexandre da Costa Sena[1], Roberto Alves Lourenço[1]

ORCID: 0009-0004-8461-8445 · 0000-0001-8420-3937 · 0000-0002-8931-9397 · 0000-0003-0838-1285

[1] Universidade do Estado do Rio de Janeiro, Rio de Janeiro (RJ), Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1797

DOI
10.59681/2175-4411.v18.iEspecial.2026.1797

RECEBIDO
17/05/2026

APROVADO
20/07/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
Mariana Lúcia Branco Zeitune
marianazeitune25@gmail.com

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

Objetivo: apresentar o GeriRisco, um sistema web de apoio à decisão para triagem de risco de quedas em idosos. **Método:** desenvolveu-se uma aplicação que integra um modelo de aprendizado de máquina (Support Vector Classifier com kernel RBF), treinado com dados da pesquisa FIBRA (N=5.931). O pipeline inclui pré-processamento, balanceamento com SMOTE, seleção de 26 variáveis e calibração isotônica. A pontuação de risco combina a probabilidade do modelo com um índice clínico baseado em evidências, classificando pacientes em três categorias. **Resultados:** o modelo obteve F1-macro de 0,607 (±0,018) e AUC de 0,658 em validação cruzada estratificada (5 folds). O sistema apresenta fatores de risco e apoia a tomada de decisão clínica. **Conclusão:** o GeriRisco demonstra a viabilidade de integrar aprendizado de máquina em sistemas de apoio à decisão para triagem geriátrica, com potencial de aplicação na atenção primária.

Descritores: Acidentes por Quedas; Sistemas de Apoio a Decisões Clínicas; Aprendizado de Máquina.

ABSTRACT

Objective: to present GeriRisco, a web-based clinical decision support system for fall risk screening in older adults. **Method:** a web application was developed integrating a machine learning model (Support Vector Classifier with RBF kernel) trained on data from the multicentric Brazilian FIBRA study (N=5,931). The pipeline includes preprocessing, SMOTE balancing, one-hot encoding of 26 clinical and sociodemographic variables selected via post-hoc interpretability analysis, and isotonic probability calibration. An integrated risk score combines model probability with an evidence-based clinical index, classifying patients into three risk categories. **Results:** the model achieved F1-macro of 0.607 (±0.018) and AUC of 0.658 in stratified 5-fold cross-validation. The system displays identified risk factors and guides clinical management. **Conclusion:** GeriRisco demonstrates the feasibility of integrating machine learning into clinical decision support for geriatric screening, with potential use in primary care.

Keywords: Accidental Falls; Decision Support Systems, Clinical; Machine Learning.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Artigo apresentado no XXI Congresso Brasileiro de Informática em Saúde (CBIS 2026), realizado de 23 a 25 de setembro de 2026, em Brasília (DF), e publicado na seção CBIS 2026 do Journal of Health Informatics.

COMO CITAR ESTE ARTIGO: Zeitune MLB, Leite KTF, Sena AC, Lourenço RA. GeriRisco: sistema web com aprendizado de máquina para triagem de quedas em idosos. J Health Inform 2026;18:e1797. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1797

1 INTRODUÇÃO

O envelhecimento populacional é um fenômeno global com implicações diretas para os sistemas de saúde. No Brasil, o Censo de 2022 registrou que 10,9% da população tem 65 anos ou mais, crescimento de 57,4% em 12 anos [1], o que impõe desafios relacionados à multimorbidade, à polifarmácia, ao declínio funcional e cognitivo e a eventos adversos, como quedas [2]. As quedas são o principal evento adverso na saúde do idoso, causando fraturas, hospitalizações, perda de autonomia e morte. Cerca de 30% dos idosos caem ao menos uma vez ao ano, prevalência similar à da pesquisa FIBRA (Fragilidade em Idosos Brasileiros), que registrou 31,1% nos últimos 12 meses em amostra de 5.931 indivíduos [3]. No Brasil, infraestrutura urbana precária, iniquidades em saúde e barreiras de acesso amplificam esse risco.

A identificação precoce de idosos em risco é essencial para intervenções preventivas eficazes; contudo, ferramentas de triagem acessíveis e sistematizadas ainda são escassas na atenção primária do Sistema Único de Saúde (SUS). Sistemas de Apoio à Decisão Clínica (SADC) baseados em aprendizado de máquina (AM) podem preencher essa lacuna, ao transformar dados rotineiramente coletados em estimativas de risco individualizadas. Este trabalho apresenta o GeriRisco, sistema web que integra modelo de AM treinado com dados da FIBRA, uma pontuação de risco integrada e uma interface acessível, com o objetivo de auxiliar profissionais de saúde na identificação de fatores de risco prioritários durante avaliações geriátricas.

A aplicação de AM à predição de quedas tem sido amplamente explorada na literatura internacional. Speiser *et al.* [4] utilizaram Árvore de Decisão e Random Forest com dados do estudo LIFE (N=1.635), alcançando precisão de 0,58 e 0,66, respectivamente. Thapa *et al.* [5] empregaram XGBoost, Regressão Logística e Perceptron Multicamadas (MLP) com registros eletrônicos de 2.785 idosos em instituições de longa permanência nos EUA, destacando medicamentos e condições crônicas como preditores relevantes. Wang *et al.* [6] obtiveram AUC de 0,96 com modelo *ensemble* treinado em características cinemáticas da marcha, registradas por sistema óptico de captura de movimento em laboratório, para prever quedas após tropeço induzido em 298 idosos saudáveis. Essa abordagem, porém, depende de equipamentos especializados, inadequados para triagem em larga escala.

No que concerne aos sistemas de apoio à decisão, Rajeevan *et al.* [7] descreveram sistemas web integrados ao prontuário eletrônico, entre eles o TRIM (*Tool to Reduce Inappropriate Medications*), que identifica medicamentos potencialmente inapropriados em idosos e gera recomendações para o regime terapêutico, enquanto Sherrington *et al.* [8] destacaram a necessidade de ferramentas de triagem baseadas em evidências para a atenção primária; e Fonseca e Matumoto [9] identificaram, na atenção primária brasileira, barreiras à prevenção de quedas em idosos, como falta de materiais e equipamentos, infraestrutura insuficiente e sistemas informatizados inadequados. São poucos, porém, os estudos brasileiros com AM para predição de quedas a partir de dados clínicos e sociodemográficos. Este trabalho diferencia-se por utilizar a base multicêntrica FIBRA (N=5.931), combinar modelo calibrado com índice clínico baseado em evidências e disponibilizar o resultado em sistema web orientado a profissionais de saúde.

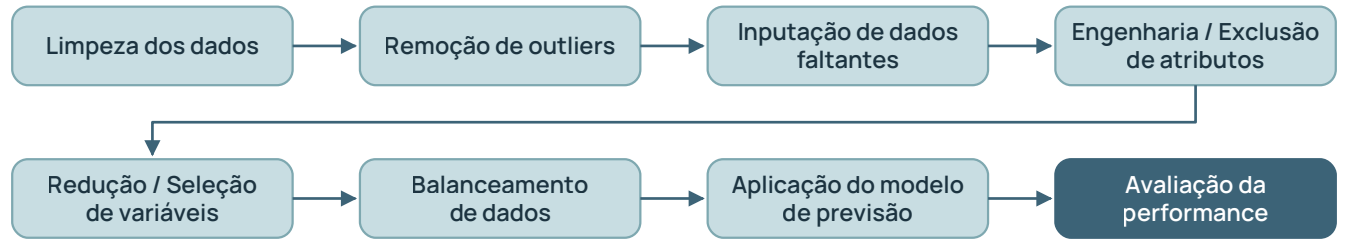
2 MÉTODOS

As subseções a seguir descrevem a metodologia adotada neste trabalho. Inicialmente, a base de dados utilizada é detalhada. Em seguida, as etapas para a implementação do modelo de AM são descritas. Ainda, na terceira subseção, é apresentada a arquitetura do sistema *Web* implementado. Por fim, é apresentado como é calculada a pontuação de risco integrada, que é a saída do sistema *Web*.

2.1 Base de Dados

Os dados provêm da pesquisa FIBRA (Fragilidade em Idosos Brasileiros), estudo multicêntrico realizado entre 2009 e 2010 em 17 cidades brasileiras [3]. A amostra (N=5.931) é composta por indivíduos com 65 anos ou mais, residentes dos setores censitários selecionados e com capacidade cognitiva preservada para as entrevistas. O questionário abordou aspectos sociodemográficos, condições de saúde, capacidade funcional (Atividades Instrumentais de Vida Diária, AIVD, e Atividades Avançadas de Vida Diária, AAVD), autoeficácia para prevenir quedas (Falls Efficacy Scale-International, FES-I), sintomas depressivos (Geriatric Depression Scale, GDS), percepção subjetiva de saúde e satisfação com a vida. A variável-alvo é a ocorrência de quedas nos últimos 12 meses (binária), com prevalência de 31,1% de casos positivos. Portanto, o modelo realiza uma classificação retrospectiva de perfis associados ao desfecho e não demonstra, nesta etapa, a predição prospectiva de quedas incidentes. A pesquisa FIBRA foi aprovada pelos Comitês de Ética em Pesquisa das instituições participantes; este estudo realizou análise secundária de dados anonimizados, não exigindo nova apreciação ética, e observou a Lei Geral de Proteção de Dados (LGPD) e a Resolução nº 466/2012 do Conselho Nacional de Saúde.

Figura 1. Etapas do pipeline proposto.



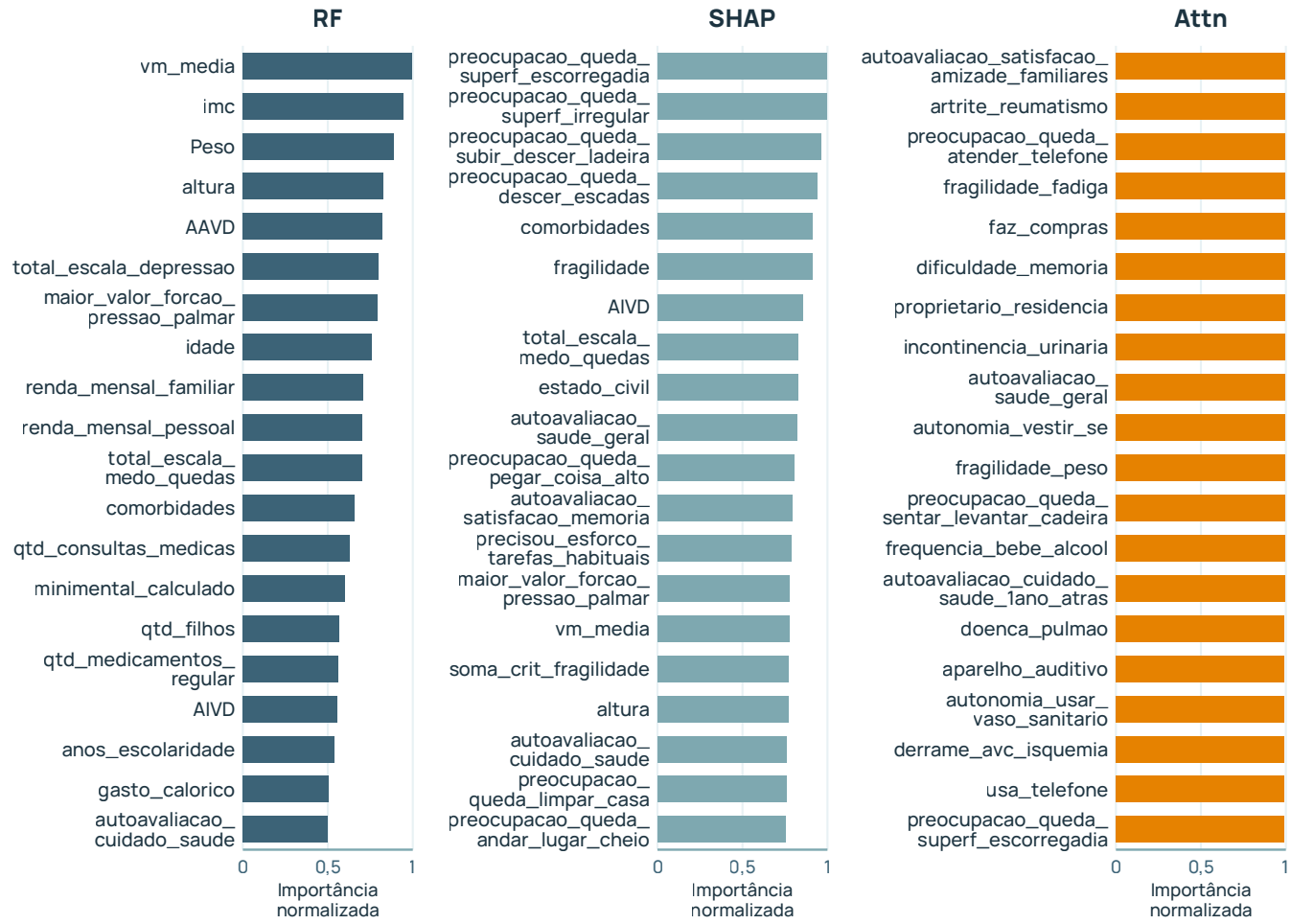
Fonte: Elaborado pelos autores (2026).

2.2 Pipeline de Aprendizado de Máquina

O *pipeline* de treinamento do modelo de implantação (Figura 1) compreende limpeza dos dados, remoção de *outliers*, codificação *one-hot* (OHE) com imputação de categorias ausentes, seleção de variáveis, padronização com StandardScaler, balanceamento com Synthetic Minority Oversampling Technique (SMOTE) aplicado dentro de cada *fold* e treinamento do classificador Support Vector Classifier (SVC) com *kernel* Radial Basis Function (RBF), class_weight "balanced" e C (*regularization parameter*) de 1,0. O desempenho foi estimado por validação cruzada estratificada de 5 *folds*: todos os transformadores foram ajustados nos dados de treino de cada *fold* e aplicados ao conjunto de teste.

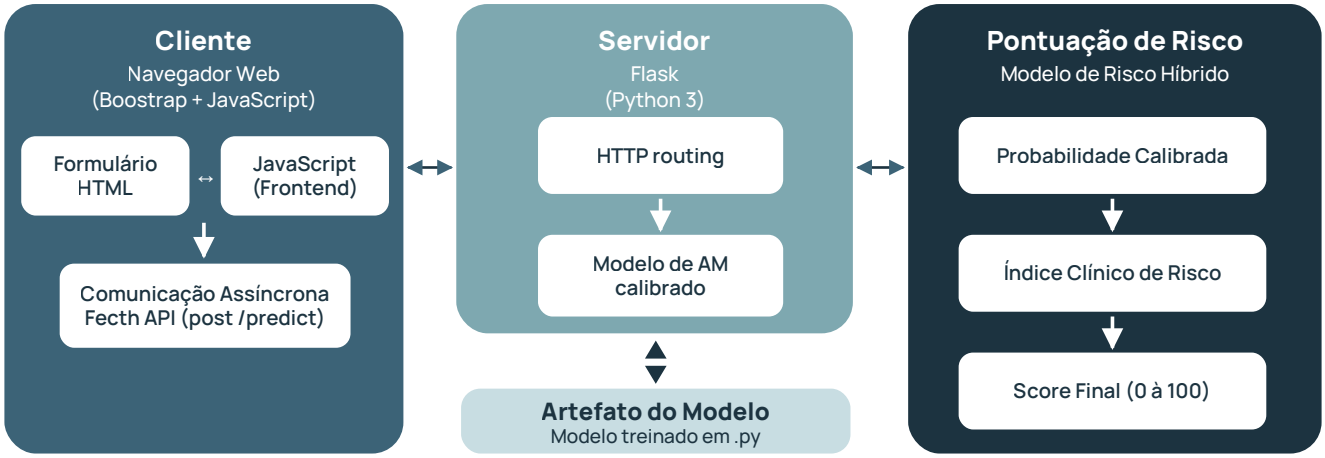
A seleção das variáveis de entrada do modelo de implantação foi realizada por análise de interpretabilidade *post hoc*, combinando três métodos aplicados aos modelos de pesquisa treinados na base completa: (i) importância de atributos por Random Forest, (ii) valores SHapley Additive exPlanations (SHAP) via KernelSHAP e (iii) pesos de atenção do TabTransformer. Os atributos foram ranqueados pela média dos três métodos, e as 26 variáveis mais relevantes foram pré-especificadas como entradas do sistema, abrangendo saúde geral, fragilidade, capacidade funcional (AIVD e AAVD), medo de cair, bem-estar e dados sociodemográficos (Figura 2).

Figura 2. Comparação dos métodos de interpretabilidade.



Fonte: Elaborado pelos autores (2026).

Figura 3. Arquitetura Cliente-Servidor do Sistema GeriRisco



Fonte: Elaborado pelos autores (2026).

O modelo final foi treinado em 80% dos dados e calibrado com regressão isotônica em conjunto independente de 20%, corrigindo estimativas de probabilidade que tendem a ser mal calibradas em dados desbalanceados após SMOTE. Os limiares de risco (baixo/médio/alto) derivam dos percentis 33 e 66 das probabilidades calibradas nesse conjunto.

2.3 Arquitetura do Sistema Web

O GeriRisco (https://geririsco.web.app/index.html) foi implementado como aplicação web cliente-servidor (Figura 3). O *backend* utiliza o *microframework* Flask (Python) em três camadas: (i) roteamento HTTP; (ii) interface de predição, responsável pela inferência e pelo cálculo da pontuação de risco; e (iii) utilitários de pré-processamento. O modelo treinado é serializado com joblib e instanciado uma única vez no início da aplicação.

O *frontend* foi construído com HTML5, CSS3 (Bootstrap) e JavaScript puro. O formulário apresenta as 26 variáveis em oito categorias clínicas: Dados Pessoais, Saúde Geral, Condições de Saúde, Fragilidade, AIVD, AAVD, Escala de Medo de Quedas e Bem-estar e Satisfação. As questões de AIVD e AAVD são apresentadas individualmente para facilitar o preenchimento, e o JavaScript computa os escores agregados antes do envio assíncrono ao *backend*, por meio da fetch API.

2.4 Pontuação de Risco Integrada

A pontuação de risco varia de 0 a 100 e combina uma probabilidade calibrada normalizada e um Índice Composto de Risco (ICR), no formato de escore clínico por pontos [10]. O ICR atua como piso de segurança frente a dois modos de falha do SVC-RBF: subestimação por extrapolação do *kernel* fora do suporte observado [11] e artefatos em degraus da calibração isotônica [12]. É também monotônico (piora clínica nunca reduz o escore), característica que o classificador probabilístico não garante. Seu cálculo está expresso nas equações (a), (b), (c) e (d).

$$\mathrm{ICR} = 100 \cdot \frac{\sum_i (v_i \cdot w_i)}{\sum_i w_i} \in [0, 100] \quad \text{(a)}$$

$$p_{norm} = \mathrm{linear}(p_{modelo}; \tau_{low}, \tau_{high}) \quad \text{(b)}$$

$$S_{pond} = 0,6 \cdot \mathrm{ICR} + 0,4 \cdot p_{norm} \quad \text{(c)}$$

$$S_{final} = \max(\mathrm{ICR}, S_{pond}) \quad \text{(d)}$$

onde v_i é a variável normalizada e w_i é o peso atribuído a cada uma delas, com $v_i \in [0, 1]$ e $w_i \in \{1, 2, 3\}$. A função linear (b) reescala a probabilidade calibrada entre os limites τ_low e τ_high. A razão 0,6/0,4 e os cortes 33/67 sobre S_final são hiperparâmetros fixados na prova de conceito e reportados para reprodutibilidade.

Os pesos refletem o grau de evidência disponível na literatura conforme ilustrado na Tabela 1. O peso 3 foi atribuído aos preditores centrais com associação independente a quedas em revisões sistemáticas ou meta-análises, incluindo fragilidade, medo de cair e dependência em AIVD [13-16]. O peso 2 foi aplicado aos indicadores com associação consistente em coortes prospectivas e magnitude intermediária. O peso 1 foi reservado aos indicadores suplementares ou contextuais, com evidência sugestiva, efeito heterogêneo ou informação parcialmente representada por componentes de maior peso. Foram escolhidos pesos inteiros para manter o índice auditável e monotônico. No entanto, os pesos foram fixados na implementação e não foram estimados ou validados como um escore clínico independente. Sua derivação estatística e sua avaliação prospectiva permanecem como trabalhos futuros.

Tabela 1. Distribuição de peso entre as variáveis para o cálculo do ICR

Peso	Variáveis	Critério de evidência
1	Saúde comparativa, consumo de álcool, satisfação com a vida, sexo e estado civil [2,14]	Evidência sugestiva, efeito pequeno/heterogêneo ou já capturado por outros componentes
2	Número total de comorbidades, saúde autoavaliada, autocuidado em saúde, declínio de atividade, artrite, queixa de memória, participação social com atividades avançadas [2,14]	Coorte prospectiva, efeito moderado replicado em populações independentes
3	Fragilidade, medo de cair, AIVD [13-16]	Associação independente em revisão sistemática ou meta-análise.

Fonte: Elaborado pelos autores (2026).

Tabela 2. Desempenho comparativo dos modelos, pipeline de implantação, validação cruzada estratificada 5-fold (média ± desvio padrão).

Modelo	Recall (classe 1)	Precisão (classe 1)	F1-macro	AUC-ROC
Voting Top-4[3]	0,431 ± 0,035	0,509 ± 0,033	0,626 ± 0,019	0,686 ± 0,016
Voting Top-3[2]	0,495 ± 0,037	0,478 ± 0,037	0,624 ± 0,025	0,683 ± 0,016
XGBoost	0,541 ± 0,032	0,450 ± 0,011	0,613 ± 0,010	0,683 ± 0,018
Elastic Net	0,576 ± 0,023	0,442 ± 0,026	0,611 ± 0,021	0,679 ± 0,017
SVC-RBF[1]	0,395 ± 0,048	0,488 ± 0,024	0,607 ± 0,018	0,658 ± 0,015
LightGBM	0,279 ± 0,025	0,612 ± 0,029	0,601 ± 0,010	0,683 ± 0,014
Gradient Boosting	0,290 ± 0,029	0,582 ± 0,049	0,600 ± 0,015	0,681 ± 0,014
CatBoost	0,241 ± 0,027	0,683 ± 0,053	0,592 ± 0,015	0,696 ± 0,023
MLP	0,411 ± 0,033	0,438 ± 0,019	0,587 ± 0,014	0,627 ± 0,019
Decision Tree	0,311 ± 0,047	0,491 ± 0,047	0,584 ± 0,022	0,636 ± 0,014
Random Forest	0,225 ± 0,025	0,658 ± 0,044	0,580 ± 0,014	0,686 ± 0,013
Naive Bayes	0,172 ± 0,044	0,736 ± 0,055	0,552 ± 0,029	0,665 ± 0,025
SVC-Linear	0,432 ± 0,026	0,347 ± 0,026	0,528 ± 0,022	0,540 ± 0,032
KNN	0,793 ± 0,026	0,364 ± 0,010	0,505 ± 0,017	0,646 ± 0,016

Fonte: Elaborado pelos autores (2026).

3 RESULTADOS

3.1 Desempenho Comparativo dos Modelos

Todos os algoritmos foram avaliados com o mesmo pré-processamento e o mesmo conjunto de variáveis (26 variáveis pré-especificadas, SMOTE dentro do *fold*, validação cruzada estratificada de 5 *folds*), conforme a Tabela 2.

[1]*Modelo selecionado para implantação.* [2]*Voting Top-3: XGBoost + Elastic Net + SVC-RBF.* [3]*Voting Top-4: + LightGBM.*

Os modelos apresentaram desempenho semelhante (F1-macro entre 0,505 e 0,626), com destaque para métodos *ensemble* e *boosting*. Embora os *ensembles* (Voting Top-3 e Top-4) e o XGBoost tenham obtido F1-macro superior (diferença ≤ 0,02), o SVC-RBF foi selecionado para implantação por dois critérios clínicos. Primeiro, sua combinação com calibração isotônica é eficaz na produção de probabilidades bem calibradas [12], propriedade importante quando a saída é apresentada como escore contínuo, ao passo que modelos baseados em árvore tendem a manter artefatos de distribuição mesmo após a calibração. Segundo, o *kernel* RBF produz fronteiras de decisão suaves, de modo que pequenas variações de entrada resultam em pequenas variações no escore (desejável para o acompanhamento clínico longitudinal). Como o ICR atua como piso de segurança, o componente de AM é otimizado para discriminação fina, e não para recall máximo.

3.2 Desempenho do Modelo

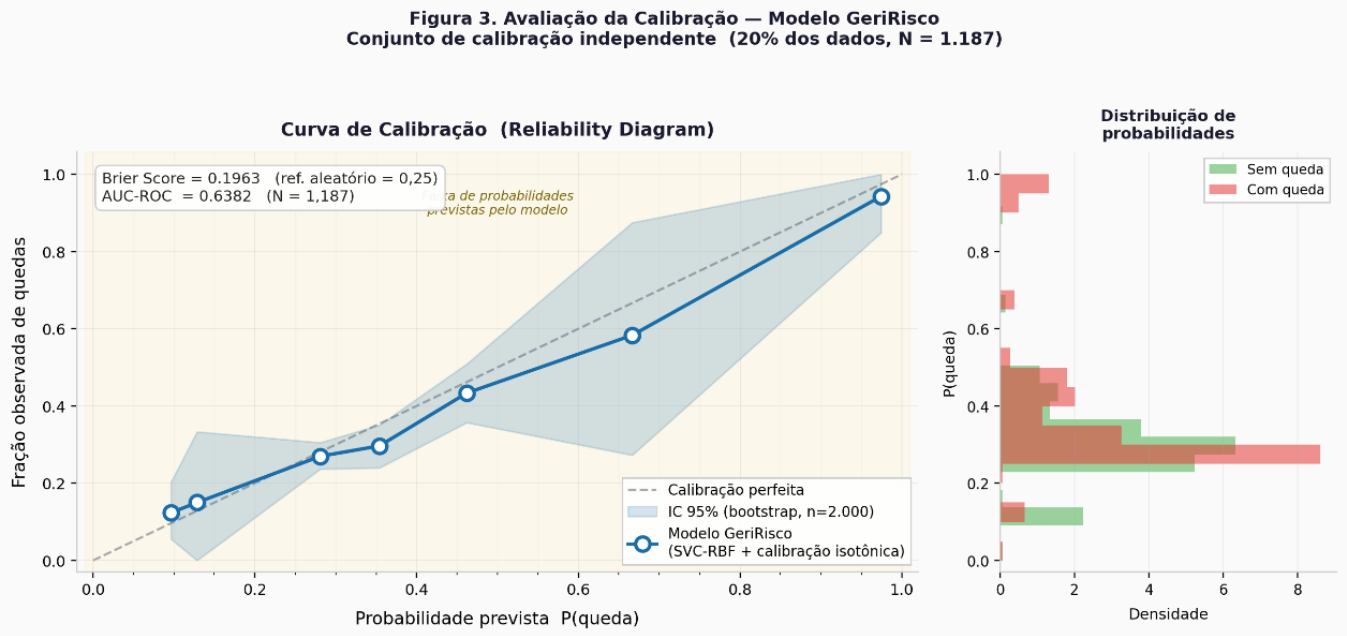
A Tabela 3 apresenta a validação cruzada estratificada de 5 *folds* para as duas variantes do modelo selecionado (SVC-RBF): variante A (sem Análise de Componentes Principais, PCA) e variante B (com PCA, 95% da variância preservada), que diferem apenas pela inclusão do PCA.

Tabela 3. Desempenho em validação cruzada estratificada de 5 folds (média ± desvio padrão).

Variante	F1-macro	Recall (classe 1)	Precisão (classe 1)	AUC-ROC
A — Sem PCA	0,607 ± 0,018	0,395 ± 0,048	0,488 ± 0,024	0,658 ± 0,015
B — Com PCA	0,606 ± 0,017	0,407 ± 0,046	0,479 ± 0,022	0,661 ± 0,013

Fonte: Elaborado pelos autores (2026).

Figura 4. Gráficos para avaliação da calibração do modelo GeriRisco.



Fonte: Elaborado pelos autores (2026).

As duas variantes apresentaram desempenho equivalente, com diferença inferior a 1 ponto percentual no F1-macro (0,607 versus 0,606). A preferência pela variante sem PCA (a Variante A, implantada) deve-se à maior interpretabilidade clínica: as 26 variáveis permanecem no espaço original, permitindo rastrear a contribuição de cada característica. No conjunto de calibração (20% dos dados, N=1.187), o modelo obteve F1-macro de 0,608, área sob a curva ROC (AUC) de 0,638 e Brier Score de 0,1963 (referência aleatória: 0,25), e a curva de calibração confirmou boa concordância entre probabilidades previstas e frequências observadas (Figura 4).

3.3 Variáveis de Entrada do Sistema

O *ranking* das 26 variáveis, definido pela importância média dos três métodos de interpretabilidade, aponta como mais relevantes: autoavaliação da saúde geral, preocupação com queda em superfície escorregadia, cidade de residência, AIVD, nível de atividade comparado a um ano atrás e preocupações com quedas, coerentes com a literatura, que identifica capacidade funcional, autoeficácia para prevenir quedas e fatores contextuais como determinantes do risco [13,14].

3.4 Comparação com escalas

Escalas tradicionais, como Morse, STRATIFY, Downton, Hendrich II e TUG, foram majoritariamente desenvolvidas em contexto hospitalar e com horizonte de curto prazo, condições em que apresentam sensibilidade entre 70% e 90% e AUC entre 0,60 e 0,76 (Tabela 4). O GeriRisco, avaliado em idosos comunitários com horizonte de 12 meses, obteve AUC de 0,658. As diferenças refletem variações de contexto e de desfecho, não sendo diretamente comparáveis.

3.5 Interface do Sistema

A interface é composta por três páginas: (i) formulário de avaliação, com as 26 questões organizadas em oito seções; (ii) painel de resultados, exibido na mesma página após o envio, com a pontuação de risco integrada (0–100), a classificação (Baixo/Médio/Alto Risco), a probabilidade bruta do modelo e os fatores de risco identificados; e (iii) página “Sobre o Modelo”, com informações técnicas sobre algoritmo, métricas, calibração e proveniência das variáveis (Figura 5).

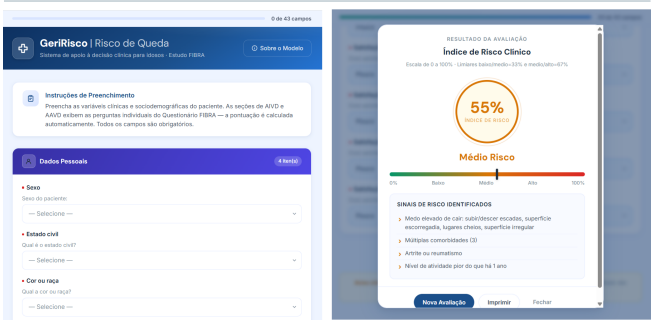
Tabela 4. Comparação com escalas tradicionais de risco de queda.

Ferramenta	Contexto	Sensibilidade	Especificidade	AUC
Morse (MFS)	Hospital	73–83%	36–49%	0,66–0,76
STRATIFY	Hospital	73–93%	70–87%	0,66–0,76
Downton	Hospital/ILPI	78–92%	35–63%	0,60–0,72
Hendrich II	Hospital	70–75%	65–74%	0,68–0,73
TUG	Comunidade	31–87%	74–87%	0,58–0,66
GeriRisco	Comunidade			0,658

Fonte: Elaborado pelos autores (2026).

O painel de resultados apresenta o escore em escala visual com três zonas de cor (verde, amarelo e vermelho) e exibe os fatores de risco como alertas específicos (por exemplo, "Fragilidade: 4 critérios de Fried atendidos" ou "Autoavaliação de saúde ruim ou muito ruim"), direcionando a atenção clínica para os domínios de maior vulnerabilidade.

Figura 5. Interface do GeriRisco.



Fonte: Elaborado pelos autores (2026).

4 DISCUSSÃO

Os resultados demonstram a viabilidade de integrar um modelo de AM calibrado em um sistema web de apoio à decisão para triagem de quedas. O F1-macro de 0,607 reflete o desafio do problema: a base FIBRA é naturalmente desbalanceada (68,9% versus 31,1%) e as variáveis são autorrelatadas, o que pode introduzir vieses de percepção e memória. Ainda assim, o desempenho é comparável ao de estudos com dados clínicos e questionários sem sensores especializados [4, 5]. A escolha do SVC apoia-se também na etapa exploratória, em que o SVC linear alcançou o maior recall para quedas (0,640); a versão implantada usa *kernel* RBF com calibração isotônica, priorizando a qualidade das probabilidades sobre o recall máximo.

A pontuação de risco integrada representa uma contribuição metodológica relevante: o ICR atua como mecanismo de segurança clínica, evitando que pacientes com múltiplos fatores de risco evidentes sejam subavaliados quando o modelo subestima a probabilidade, comportamento conhecido do SVC-RBF para instâncias fora da distribuição de treinamento [12].

A organização do formulário em categorias familiares ao profissional de saúde (fragilidade, AIVD, AAVD, medo de cair) e a geração automática de fatores de risco nomeados visam aumentar a aceitabilidade e a utilidade clínica do sistema. Estudos de implementação de SADC destacam que apresentar justificativas e fatores específicos, em vez de apenas um escore, aumenta a confiança e a adoção por profissionais [17]. Como detalhado na Seção 3.4, o GeriRisco distingue-se das escalas tradicionais por fornecer probabilidade calibrada, explicações individuais e ajuste do limiar de decisão conforme o contexto, posicionando-se como ferramenta complementar de triagem na atenção primária.

Apesar do desempenho global consistente, foram identificadas três situações de comportamento inferior ao esperado. Primeiro, há heterogeneidade em idosos frágeis, com aumento da sensibilidade e redução da especificidade, o que pode gerar excesso de falsos positivos. Segundo, há subestimação de risco em perfis clinicamente graves, com probabilidades baixas atribuídas a pacientes que efetivamente caíram. Terceiro, cerca de 9,6% dos indivíduos situam-se em zona limítrofe ao limiar de decisão, na qual pequenas variações de entrada alteram a classificação. Para mitigar esses efeitos, o sistema usa o ICR como piso de segurança e fornece explicações locais, permitindo ao profissional revisar a decisão automatizada.

Outras limitações merecem consideração. A ausência de validação prospectiva em ambiente clínico real impede confirmar impacto sobre desfechos. O sistema não possui módulo de autenticação nem registro de avaliações, necessários para uso assistencial formal. Por fim, as 26 variáveis incluem perguntas sobre cidades específicas do estudo FIBRA, o que pode limitar a generalização para outros municípios. Como os dados da FIBRA foram coletados entre 2009 e 2010, mudanças na população idosa brasileira podem afetar a validade do modelo: aumento da expectativa de vida e da escolaridade, expansão da Atenção Primária e do

acesso a medicamentos, alterações na multimorbidade e na polifarmácia e os efeitos da pandemia de COVID-19 sobre mobilidade e participação social. Esse deslocamento distribucional (*dataset shift*) afeta sobretudo a calibração, o nível absoluto das probabilidades e, por consequência, os limiares entre as faixas de risco, o que sugere recalibração periódica em dados contemporâneos e validação externa em coorte mais recente.

5 CONCLUSÃO

Este trabalho apresentou o GeriRisco, sistema web de apoio à decisão clínica para triagem de risco de quedas em idosos, que integra um modelo SVC calibrado com 26 variáveis clínicas e sociodemográficas derivadas da FIBRA, com F1-macro de 0,607 em validação cruzada. Os principais diferenciais são: (i) uso de base multicêntrica brasileira de grande porte; (ii) calibração explícita das probabilidades; (iii) mecanismo de segurança clínica que previne subestimação de risco em pacientes vulneráveis; e (iv) exibição de fatores de risco nomeados que direcionam a conduta clínica.

Como trabalhos futuros, duas frentes estão previstas. A primeira é a validação externa no Estudo Longitudinal da Saúde dos Idosos Brasileiros (ELSI-Brasil) [18], coorte nacional de base domiciliar, com uso das ondas 1, 2 e 3, o que permite avaliar a transportabilidade do modelo para amostra independente e mais recente, a estabilidade temporal do desempenho e a predição de quedas incidentes entre ondas. A etapa demanda a compatibilização prévia das variáveis entre as duas coortes e deve reportar discriminação e calibração (*calibration-in-the-large*, inclinação e Brier Score), com eventual recalibração dos limiares, seguindo o TRIPOD+AI [19].

A segunda frente reúne a avaliação clínica prospectiva e a avaliação de usabilidade em cenário real [20], por meio de estudo-piloto em ambulatório de geriatria do Hospital Universitário Pedro Ernesto (HUPE/UERJ), mediante aprovação do Comitê de Ética em Pesquisa. O piloto prevê a aplicação do GeriRisco em atendimentos de rotina, a comparação da estratificação obtida com a avaliação geriátrica ampla e o seguimento dos participantes para verificação de quedas incidentes.

Contribuição dos autores. Os autores contribuíram igualmente em todas as etapas de construção do artigo: concepção e delineamento do estudo; coleta, análise e interpretação dos dados; redação e revisão crítica do manuscrito; e aprovação da versão final.

Agradecimentos. Agradeço à equipe responsável pelo Estudo FIBRA Brasil e, em especial, ao Dr. Roberto Lourenço, pela disponibilização da base de dados e pelas valiosas orientações e contribuições ao desenvolvimento desta pesquisa. Expresso minha gratidão aos meus orientadores, Prof. Dr. Alexandre Sena e Profa. Dra. Karla Figueiredo, pelo acompanhamento, apoio e orientação ao longo de todas as etapas do projeto. Agradeço também à Universidade do Estado do Rio de Janeiro (UERJ) e aos seus profissionais pelo suporte institucional e acadêmico.

Financiamento. Os autores não declararam fonte de financiamento.

Disponibilidade de dados e materiais. O código-fonte do pipeline final de aprendizado de máquina, o questionário utilizado no Estudo FIBRA e o código-fonte da aplicação web desenvolvida neste trabalho estão disponíveis publicamente no seguinte repositório: https://github.com/MarianaZeitune/projectfibra.git

Conflitos de interesse. Os autores não declararam conflitos de interesse.

Declaração de uso de IA. Ferramentas de inteligência artificial generativa foram utilizadas como suporte à revisão ortográfica e gramatical do texto. Toda a concepção, análise, interpretação dos resultados e desenvolvimento do sistema são de autoria exclusiva dos pesquisadores.

REFERÊNCIAS

1. Instituto Brasileiro de Geografia e Estatística. Censo Demográfico 2022: primeiros resultados [Internet]. Rio de Janeiro: IBGE; 2023 [citado 2026 out 18]. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/livros/liv102011.pdf
2. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing. 2022;51(9):afac205. doi: https://doi.org/10.1093/ageing/afac205
3. Lourenço RA, Sanchez MA, Moreira VG, Ribeiro PCC, Perez M, Campos GC, et al. Fragilidade em Idosos Brasileiros - Fibra-RJ: metodologia de pesquisa dos estudos de fragilidade, distúrbios cognitivos e sarcopenia. Revista Hospital Universitário Pedro Ernesto. 2015;14(4). doi: https://doi.org/10.12957/rhupe.2015.20066
4. Speiser JL, Callahan KE, Houston DK, Fanning J, Gill TM, Guralnik JM, et al. Machine learning in aging: an example of developing prediction models for serious fall injury in older adults. J Gerontol A Biol Sci Med Sci. 2021;76(4):647-54. doi: https://doi.org/10.1093/gerona/glaa138
5. Thapa R, Garikipati A, Shokouhi S, Hurtado M, Barnes G, Hoffman J, et al. Predicting falls in long-term care facilities: machine learning study. JMIR Aging. 2022;5(2):e35373. doi: https://doi.org/10.2196/35373
6. Wang S, Nguyen TK, Bhatt T. Trip-related fall risk prediction based on gait pattern in healthy older adults: a machine-learning approach. Sensors (Basel). 2023;23(12):5536. doi: https://doi.org/10.3390/s23125536

7. Rajeevan N, Niehoff KM, Charpentier P, Levin FL, Justice A, Brandt CA, et al. Utilizing patient data from the Veterans Administration electronic health record to support web-based clinical decision support: informatics challenges and issues from three clinical domains. BMC Med Inform Decis Mak. 2017;17(1):111. doi: https://doi.org/10.1186/s12911-017-0501-x
8. Sherrington C, Fairhall NJ, Wallbank GK, Tiedemann A, Michaleff ZA, Howard K, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev. 2019;1(1):CD012424. doi: https://doi.org/10.1002/14651858.CD012424.pub2
9. Fonseca RFMR, Matumoto S. Falls prevention in community-dwelling elderly in Brazil: strategies and difficulties in primary health care. Acta Sci Health Sci. 2024;46:e64932. doi: https://doi.org/10.4025/actascihealthsci.v46i1.64932
10. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med. 2004;23(10):1631-60. doi: https://doi.org/10.1002/sim.1742
11. Hsu, C.W., Chang, C.C., & Lin, C.J. (2003). A practical guide to support vector classification. Technical Report, National Taiwan University.
12. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: Proceedings of the 22nd international conference on Machine learning - ICML '05. New York, New York, USA: ACM Press; 2005. p. 625-32. doi: https://doi.org/10.1145/1102351.1102430
13. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):M146-56. doi: https://doi.org/10.1093/gerona/56.3.m146
14. Deandrea S, Lucenteforte E, Bravi F, Foschi R, La Vecchia C, Negri E. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. Epidemiology. 2010;21(5):658-68. doi: https://doi.org/10.1097/EDE.0b013e3181e89905
15. Kojima G. Frailty as a Predictor of Future Falls Among Community-Dwelling Older People: A Systematic Review and Meta-Analysis. J Am Med Dir Assoc. 2015;16(12):1027-33. doi: https://doi.org/10.1016/j.jamda.2015.06.018
16. Delbaere K, Close JCT, Brodaty H, Sachdev P, Lord SR. Determinants of disparities between perceived and physiological risk of falling among elderly people: cohort study. BMJ. 2010;341(aug18 4):c4165. doi: https://doi.org/10.1136/bmj.c4165
17. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330(7494):765. doi: https://doi.org/10.1136/bmj.38398.500764.8F
18. Lima-Costa MF, de Andrade FB, de Souza PRB, Neri AL, Duarte YAO, Castro-Costa E, et al. The Brazilian Longitudinal Study of Aging (ELSI-Brazil): Objectives and Design. Am J Epidemiol. 2018;187(7):1345-53. doi: https://doi.org/10.1093/aje/kwx387
19. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378. doi: https://doi.org/10.1136/bmj-2023-078378
20. Yusof MM, Kuljis J, Papazafeiropoulou A, Stergioulas LK. An evaluation framework for Health Information Systems: human, organization and technology-fit factors (HOT-fit). Int J Med Inform. 2008;77(6):386-98. doi: https://doi.org/10.1016/j.ijmedinf.2007.08.011

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde