Journal of Health Informatics

J Health Inform. 2026;18:e1791
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1791

ARTIGO COMPLETO
CBIS 2026

Método de auxílio ao diagnóstico de câncer de próstata utilizando inteligência artificial e dados clínicos

Method to support prostate cancer diagnosis using artificial intelligence and clinical data

Wesley Batista Dominices de Araujo[1], Ewaldo Eder Carvalho Santana[1], Giullianno Lopes de Moura[2]

ORCID: 0009-0007-0953-8041 · 0000-0002-8894-5353 · 0009-0008-2885-7301

[1] Universidade Estadual do Maranhão, São Luís (MA), Brasil.

[2] Hospital Universitário da Universidade Federal do Maranhão, São Luís (MA), Brasil.

FICHA DO ARTIGO

TIPO
Artigo completo · CBIS 2026

IDENTIFICADOR
e1791

DOI
10.59681/2175-4411.v18.iEspecial.2026.1791

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
Wesley Batista Dominices de Araujo
wesleydominices@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: propor um método de auxílio à triagem para biópsia de câncer de próstata com base em inteligência artificial e dados clínicos. **Método:** estudo observacional retrospectivo com 274 pacientes, utilizando idade, raça, hipertensão arterial sistêmica, diabetes mellitus, tabagismo, etilismo, toque retal e antígeno prostático específico (PSA) total. Foram avaliados Naive Bayes, Support Vector Machine, K-Nearest Neighbor, árvore de decisão e rede neural artificial, com otimização de hiperparâmetros e validação cruzada 10-fold. **Resultados:** o melhor desempenho foi do Naive Bayes, com acurácia de 89,09%, sensibilidade de 92%, especificidade de 86,67% e área sob a curva (AUC) de 0,9187. **Conclusão:** o método mostrou potencial para reduzir biópsias desnecessárias e apoiar a decisão clínica.

Descritores: Neoplasias da próstata; Inteligência artificial; Programas de rastreamento.

ABSTRACT

Objective: to propose a method to support prostate biopsy screening based on artificial intelligence and clinical data. **Method:** retrospective observational study with 274 patients using age, race, systemic arterial hypertension, diabetes mellitus, smoking, alcoholism, digital rectal examination, and total PSA. Naive Bayes, Support Vector Machine, K-Nearest Neighbor, decision tree, and artificial neural network were evaluated with hyperparameter optimization and 10-fold cross-validation. **Results:** Naive Bayes achieved the best performance, with accuracy of 89.09%, sensitivity of 92%, specificity of 86.67%, and AUC of 0.9187. **Conclusion:** the method showed potential to reduce unnecessary biopsies and support clinical decision-making.

Keywords: Prostatic neoplasms; Artificial intelligence; Mass screening.

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: Araujo WBD, Santana EEC, Moura GL. Método de auxílio ao diagnóstico de câncer de próstata utilizando inteligência artificial e dados clínicos. J Health Inform 2026;18:e1791. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1791

1 INTRODUÇÃO

O câncer de próstata apresenta alta prevalência e elevada mortalidade entre homens em nível mundial e no Brasil, o que mantém a doença como um importante problema de saúde pública [1-5]. O rastreamento inicial é realizado, principalmente, por meio do antígeno prostático específico (PSA) total e do toque retal. Quando esses exames apresentam alterações, podem ser solicitados exames complementares, como ressonância magnética e biópsia. Entretanto, a biópsia prostática permanece como o único procedimento confirmatório, apesar de seu caráter invasivo, doloroso e do custo financeiro associado [6-9].

Nesse contexto, métodos computacionais de apoio à decisão têm sido propostos para auxiliar o médico na indicação de exames adicionais e na triagem para biópsia. O aprendizado de máquina permite identificar padrões em bases clínicas e construir modelos preditivos capazes de classificar novos casos a partir de exemplos previamente rotulados [10]. Estudos recentes têm aplicado aprendizado de máquina ao câncer de próstata com dados de imagem, dados multivariados e diferentes estratégias de modelagem, demonstrando o potencial da área para apoiar o diagnóstico [11-18].

Além disso, a interpretação clínica desses exames nem sempre é suficiente para definir, de forma objetiva, quais pacientes realmente devem ser submetidos à biópsia prostática. Muitos homens são encaminhados para investigação invasiva mesmo sem confirmação da doença, enquanto outros podem ter o diagnóstico retardado. Essa limitação decorre da sobreposição entre achados benignos e malignos, o que dificulta a estratificação de risco apenas com base na análise convencional. Nesse contexto, o uso de técnicas de inteligência artificial aplicadas a variáveis clínicas rotineiras surge como uma alternativa promissora para apoiar a decisão médica. Ao integrar múltiplos fatores clínicos em modelos preditivos, torna-se possível aumentar a precisão da triagem e reduzir procedimentos desnecessários.

A principal motivação deste estudo foi a escassez de métodos de baixo custo que utilizem variáveis clínicas não invasivas para apoiar a triagem de pacientes com risco de câncer de próstata. Assim, o objetivo foi desenvolver um método de auxílio ao diagnóstico utilizando inteligência artificial e dados clínicos, buscando identificar o modelo preditivo mais eficiente para classificar pacientes entre as classes câncer e normal.

Do ponto de vista clínico, o rastreamento do câncer de próstata permanece desafiador porque os exames tradicionalmente utilizados apresentam limitações importantes quando empregados de forma isolada. O PSA total pode se elevar por causas não malignas, como hiperplasia prostática benigna, prostatite, idade avançada e procedimentos urológicos, enquanto o toque retal depende da experiência do examinador e não detecta todos os tumores clinicamente relevantes. Esse cenário favorece resultados falso-positivos e falso-negativos, produzindo incerteza diagnóstica, ansiedade no paciente e aumento do número de exames subsequentes, inclusive procedimentos invasivos [6-9].

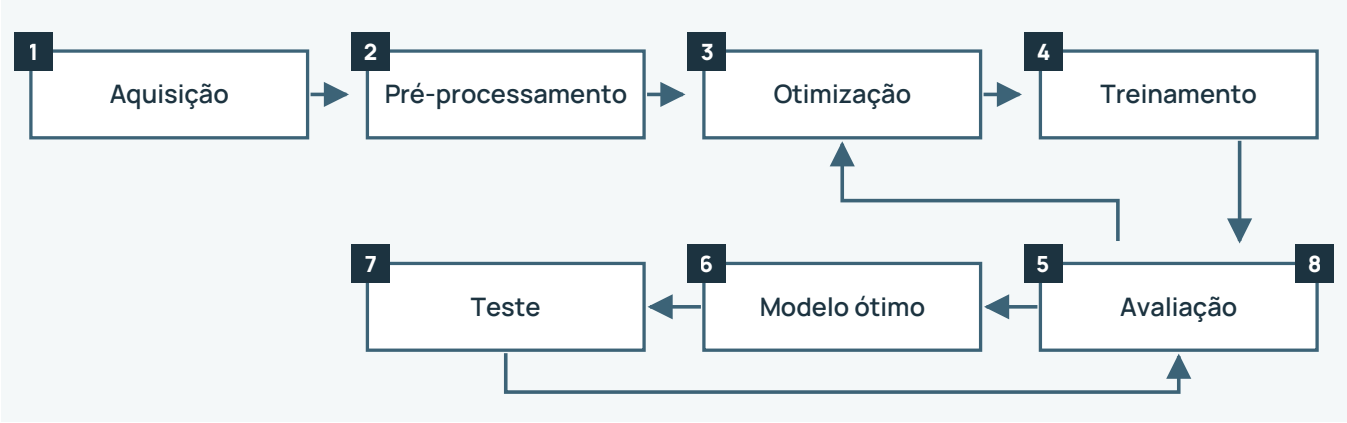
A literatura recente descreve diferentes estratégias computacionais para apoiar o diagnóstico do câncer de próstata, incluindo modelos baseados em ressonância magnética, ultrassonografia, fotoacústica e combinações multivariadas de parâmetros laboratoriais e clínicos [11-18]. Embora esses estudos mostrem resultados promissores, muitos dependem de exames de imagem sofisticados, biomarcadores adicionais ou contextos institucionais com maior disponibilidade tecnológica. Em sistemas de saúde com restrições de custo e acesso, especialmente em cenários assistenciais públicos, torna-se relevante investigar modelos que utilizem dados clínicos rotineiramente disponíveis no prontuário.

Nesse sentido, esta pesquisa se diferencia ao concentrar-se em oito variáveis clínicas de obtenção relativamente simples e baixo custo: idade, raça, hipertensão arterial sistêmica, diabetes mellitus, tabagismo, etilismo, toque retal e PSA total. A hipótese central é que a integração dessas variáveis por meio de algoritmos de aprendizado de máquina pode oferecer suporte à tomada de decisão médica quanto à indicação de biópsia, reduzindo procedimentos desnecessários e mantendo desempenho diagnóstico compatível com aplicações clínicas. A proposta não substitui a avaliação especializada, mas atua como ferramenta adicional de apoio, com foco em triagem e priorização de risco.

2 MÉTODOS

Este trabalho propõe o desenvolvimento de um método de auxílio ao diagnóstico de câncer de próstata através da aplicação das técnicas de aprendizado de máquina e dados clínicos, para que sirva de triagem para a biópsia de câncer de próstata. Para isto foi desenvolvida uma metodologia composta de oito etapas. A primeira etapa é a aquisição dos dados, a segunda etapa é o pré-processamento, a terceira etapa é a otimização dos parâmetros, a quarta etapa é o treinamento de cada modelo, a quinta etapa é a avaliação dos modelos treinados, a sexta etapa é a geração do modelo ótimo para cada algoritmo de aprendizado de máquina, a sétima etapa compõe a fase de testes com novos dados não utilizados no treinamento, e a última etapa é a avaliação dos modelos após os testes. O diagrama em blocos do método proposto é mostrado na Figura 1.

Figura 1. Diagrama esquemático do método proposto



Fonte: Elaborado pelos autores (2026).

Do ponto de vista ético, o estudo foi aprovado pela Comissão Científica e pelo Comitê de Ética em Pesquisa (CEP) do HU-UFMA, conforme CAAE 45444621.6.0000.5086 e parecer técnico consubstanciado ID 4.679.671. O hospital autorizou o uso retrospectivo dos dados e dispensou o termo de consentimento livre e esclarecido, mantendo-se a confidencialidade dos participantes.

O projeto foi conduzido com foco em pacientes submetidos à investigação diagnóstica para câncer de próstata, sendo adotado o resultado da biópsia como referência para definição das classes normal e câncer. Para reduzir viés de classe e permitir comparação equilibrada entre os algoritmos, foi utilizada amostra balanceada composta por 137 pacientes com câncer e 137 sem câncer. Os prontuários abrangem o período de janeiro de 2017 a outubro de 2023.

Na etapa de preparação do conjunto de dados, as variáveis foram parametrizadas de forma padronizada para permitir sua utilização pelos modelos computacionais. As categorias de raça, hipertensão arterial sistêmica, diabetes mellitus, tabagismo e etilismo foram codificadas numericamente; idade e toque retal foram mantidos como variáveis quantitativas; e o PSA total foi utilizado em sua escala contínua. O pré-processamento buscou assegurar consistência entre os registros, minimizar ambiguidades semânticas e organizar o *dataset* para treinamento supervisionado.

Para a seleção do classificador mais adequado, foram avaliados cinco algoritmos amplamente empregados em problemas de classificação biomédica: Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbor (KNN), Árvore de Decisão (*decision tree*, DT) e Rede Neural Artificial (RNA). Os experimentos foram conduzidos no ambiente MATLAB® R2023b, utilizando o aplicativo *Classification Learner*. A otimização dos hiperparâmetros foi realizada por meio da funcionalidade *Optimize Hyperparameters*, baseada em Otimização Bayesiana (*Bayesian optimization*), que explora iterativamente o espaço de busca para identificar as combinações de parâmetros que maximizam o desempenho preditivo dos classificadores. Para cada algoritmo, foram otimizados automaticamente os hiperparâmetros específicos disponibilizados pelo MATLAB, incluindo parâmetros de *kernel* e regularização para o SVM, número de vizinhos para o KNN, critérios de divisão para a árvore de decisão, arquitetura da rede neural artificial e parâmetros de distribuição para o Naive Bayes. Ao final do processo, foi selecionado, para cada algoritmo, o conjunto de hiperparâmetros que apresentou o melhor desempenho durante a etapa de treinamento.

Os dados foram divididos em 80% para treinamento e 20% para teste, mantendo-se o conjunto de teste completamente independente durante todo o processo de desenvolvimento dos modelos. Para assegurar a reprodutibilidade dos experimentos, foi fixada a semente do gerador de números aleatórios (rng = 100). Durante o treinamento foi empregada validação cruzada estratificada *10-fold*, preservando a proporção entre pacientes com câncer e pacientes sem câncer de próstata em cada partição e reduzindo a variabilidade decorrente de uma única divisão da amostra [19]. A otimização dos hiperparâmetros e o treinamento dos classificadores foram realizados exclusivamente sobre o conjunto de treinamento, enquanto o conjunto de teste foi utilizado apenas na avaliação final, minimizando o risco de sobreajuste (*overfitting*) e de vazamento de informações (*data leakage*). O desempenho dos modelos foi avaliado por meio das métricas de acurácia, sensibilidade, especificidade e área sob a curva ROC (AUC), por serem amplamente empregadas na avaliação de classificadores em aplicações biomédicas e permitirem analisar simultaneamente a capacidade de detecção da doença e o controle de falsos positivos.

Tabela 1. Correlação de Spearman do dataset

Variável	Tipo	Idade	Raça	HAS	DM	Tabagismo	Etilismo	Toque	PSA total
Idade	Coef.	1,000	-0,001	-0,157**	-0,098	0,155*	0,136*	0,293**	-0,014
	Sig.		0,993	0,009	0,105	0,010	0,024	0,000	0,821
Raça	Coef.	-0,001	1,000	-0,121*	0,057	-0,023	-0,013	0,087	0,140*
	Sig.	0,993		0,045	0,347	0,704	0,830	0,149	0,020
HAS	Coef.	-0,157**	-0,121*	1,000	0,226**	-0,062	0,016	-0,059	-0,065
	Sig.	0,009	0,045		0,000	0,310	0,797	0,329	0,283
DM	Coef.	-0,098	0,057	0,226**	1,000	-0,016	-0,042	-0,067	-0,022
	Sig.	0,105	0,347	0,000		0,786	0,491	0,266	0,716
Tabagismo	Coef.	0,155*	-0,023	-0,062	-0,016	1,000	0,187**	0,073	-0,036
	Sig.	0,010	0,704	0,310	0,786		0,002	0,227	0,553
Etilismo	Coef.	0,136*	-0,013	0,016	-0,042	0,187**	1,000	0,155*	-0,095
	Sig.	0,024	0,830	0,797	0,491	0,002		0,010	0,117
Toque	Coef.	0,293**	0,087	-0,059	-0,067	0,073	0,155*	1,000	0,058
	Sig.	0,000	0,149	0,329	0,266	0,227	0,010		0,338
PSA total	Coef.	-0,014	0,140*	-0,065	-0,022	-0,036	-0,095	0,058	1,000
	Sig.	0,821	0,020	0,283	0,716	0,553	0,117	0,338	

Abreviações: Coef. = Coeficiente; Sig. = Significância; HAS = Hipertensão Arterial Sistêmica; DM = *Diabetes Mellitus*; PSA = *Prostate-Specific Antigen*. ** A correlação é significativa no nível 0,01 (2 extremidades). * A correlação é significativa no nível 0,05 (2 extremidades).

Fonte: Elaborado pelos autores (2026).

3 RESULTADOS

Nesta seção, os resultados obtidos serão apresentados. O método proposto foi implementado usando os softwares MATLAB® R2023b e IBM SPSS®. Foi realizada uma pesquisa observacional retrospectiva através dos prontuários de 274 pacientes (137 sem câncer de próstata e 137 com câncer de próstata) do setor de Urologia do HU-UFMA, e foi criado um *dataset* a partir dos mesmos. Foram utilizadas oito variáveis clínicas (idade, raça, hipertensão arterial sistêmica (HAS), diabetes mellitus (DM), etilismo, tabagismo, toque e PSA total), como variáveis de entrada dos algoritmos utilizados neste trabalho. Foi criada uma variável chamada de "Rótulo", a qual identifica se o resultado da biópsia foi positivo (Gleason > 0), ou seja, o paciente tem câncer de próstata, ou se o resultado foi negativo (normal), ou seja, o paciente não tem câncer de próstata. A amostra apresentou média de idade de 67,23 anos, predominância de pacientes pardos e distribuição clínica compatível com o perfil assistencial do serviço. A média do toque retal foi de 57,81 g e a média do PSA total foi de 8,25 ng/mL. Na análise por classes, os pacientes com câncer apresentaram média de PSA total mais elevada (10,89 ng/mL) do que os pacientes com biópsia negativa (5,61 ng/mL), enquanto a diferença de idade entre os grupos foi menos expressiva. Em termos de frequência, observou-se maior proporção de hipertensão arterial sistêmica, diabetes mellitus e etilismo entre os pacientes classificados com câncer. A análise de correlação de Spearman pode ser vista na Tabela 1.

A análise de correlação de Spearman apresentou associações monotônicas relevantes entre algumas variáveis clínicas, reforçando a utilidade da exploração estatística prévia para compreensão do conjunto de dados (Tabela 1).

A análise individual da AUC (Tabela 2) mostra que o PSA total foi a variável isolada mais representativa, enquanto as demais variáveis apresentaram contribuição complementar quando consideradas em conjunto. Esse achado sustenta a estratégia multivariada adotada, pois o desempenho preditivo decorre menos de um único marcador e mais da integração entre sinais clínicos e laboratoriais de baixo custo.

Após otimização e treinamento dos modelos, o Naive Bayes apresentou o melhor desempenho global no conjunto de teste, com acurácia de 89,09%, sensibilidade de 92%, especificidade de 86,67% e AUC de 0,9187. Entre os demais classificadores, a RNA apresentou a segunda melhor AUC (0,9027), com acurácia, sensibilidade e especificidade de 80%. O KNN também apresentou bom desempenho global, com acurácia de 80%, sensibilidade de 88%, especificidade de 73,33% e AUC de 0,8947, enquanto a árvore de decisão apresentou a melhor especificidade entre os modelos não probabilísticos (83,33%), porém com AUC inferior (0,8180). O SVM apresentou o menor desempenho relativo, com acurácia de 76,36% e AUC de

Tabela 2. AUC de cada variável

Variável	AUC	IC 95% Assintótico	
		Limite inferior	Limite superior
Idade	0,369	0,303	0,435
Raça	0,540	0,472	0,608
HAS	0,478	0,410	0,547
DM	0,464	0,395	0,532
Tabagismo	0,489	0,421	0,558
Etilismo	0,404	0,337	0,471
Toque	0,322	0,259	0,385
PSA total	0,795	0,739	0,850

Abreviações: HAS = Hipertensão Arterial Sistêmica; DM = *Diabetes Mellitus*; PSA = *Prostate-Specific Antigen;* AUC = *Area Under the Curve.*

Fonte: Elaborado pelos autores (2026).

Tabela 3. Desempenho dos modelos de aprendizado de máquina no subconjunto de teste.

Modelo	Acurácia (%)	Sensibilidade (%)	Especificidade (%)	AUC (IC 95%)
Naive Bayes	89,09	92,00	86,67	0,9187 (0,838-0,999)
RNA	80,00	80,00	80,00	0,9027 (0,811-0,995)
SVM	76,36	80,00	73,33	0,8427 (0,722-0,963)
DT	81,82	80,00	83,33	0,8180 (0,695-0,941)

Abreviações: AUC = *Area Under the Curve*; IC = Intervalo de Confiança; RNA = Rede Neural Artificial; KNN = *K-Nearest Neighbor*; SVM = *Support Vector Machine*; DT = *Decision Tree.*

Fonte: Elaborado pelos autores (2026).

Tabela 4. Comparação indireta do método proposto com estudos relacionados publicados na literatura

Trabalho	Algoritmo	Método	Resultado	
			AUC	Acurácia (%)
Wang *et al.* [11]	SVM	IRM	0,700	-
Liu *et al.* [12]	Regressão Logística	Multivariáveis	0,775	-
Liu *et al.* [13]	Regressão Logística	Multivariáveis	0,790	-
Liu *et al.* [12]	Regressão Logística	Multivariáveis com *Gleason* ≥ 7	0,819	-
Liu *et al.* [13]	Regressão Logística	Multivariáveis com *Gleason* ≥ 7	0,840	-
Wang *et al.* [11]	DCNN	IRM	0,840	-
Park *et al.* [14]	Regressão Logística	Multivariáveis	0,840	-
Yoo *et al.* [15]	*Random Forest*	IRM	0,840	-
Chen *et al.* [16]	LDA	Fotoacústica	0,851	76,30
Lu *et al.* [17]	Bootstrap	-	-	80,46
Chen *et al.* [16]	QDA	Fotoacústica	0,862	81,70
Yoo *et al.* [15]	DCNN	IRM	0,870	-
Zheng *et al.* [18]	Regressão Logística	Multivariáveis	0,906	82,70
Araujo *et al.* [20]	RNA	Multivariáveis	0,9027	80,00
Método proposto	*Naive Bayes*	Multivariáveis	0,9187	89,09

Abreviações: AUC = Area Under the Curve; SVM = Support Vector Machine; IRM = Imagem por Ressonância Magnética; DCNN = Deep Learning Convolutional Neural Network; LDA = Linear Discriminant Analysis; QDA = Quadratic Discriminant Analysis; RNA = Rede Neural Artificial.

Fonte: Elaborado pelos autores (2026).

0,8427. Em conjunto, os resultados mostram que todos os algoritmos foram capazes de discriminar as classes, mas o Naive Bayes apresentou o melhor equilíbrio entre sensibilidade, especificidade e discriminação global. O desempenho dos modelos está apresentado na Tabela 3.

Numerosos estudos já foram realizados sobre o diagnóstico de câncer de próstata utilizando aprendizado de máquina. Esses estudos têm explorado diversas abordagens e algoritmos para melhorar a precisão e a eficiência do diagnóstico. No entanto, poucos estudos se concentraram especificamente na análise de variáveis clínicas no contexto do rastreamento da biópsia do câncer de próstata. Ao incluir variáveis clínicas detalhadas, como idade, raça, hipertensão arterial sistêmica, diabetes mellitus, tabagismo, etilismo, toque retal e PSA total, o método proposto apresentou uma perspectiva mais abrangente e personalizada para o diagnóstico. A Tabela 4 apresenta uma comparação indireta entre o método proposto e estudos relacionados encontrados na literatura. Os trabalhos comparados utilizaram bases de dados, populações, critérios de inclusão, variáveis de entrada, padrões de referência e estratégias de validação distintos. Portanto, os valores de AUC e acurácia não devem ser interpretados como resultado de uma comparação direta ou como evidência de superioridade absoluta entre os métodos. A finalidade da tabela é contextualizar o desempenho obtido neste estudo em relação a diferentes abordagens previamente descritas.

Embora o método proposto tenha apresentado os maiores valores de AUC e acurácia entre os estudos listados, essa observação deve ser interpretada com cautela, pois não houve aplicação dos métodos da literatura à mesma base de dados nem avaliação sob um protocolo experimental comum. Dessa forma, os resultados permitem posicionar o estudo no contexto da literatura, mas não permitem afirmar que o Naive Bayes seja universalmente superior aos métodos previamente publicados.

4 DISCUSSÃO

Os resultados indicam que a combinação de variáveis clínicas e PSA total com algoritmos de aprendizado de máquina pode auxiliar a estratificação de pacientes submetidos à investigação de câncer de próstata. No conjunto de dados analisado, o Naive Bayes apresentou o melhor desempenho no subconjunto de teste, com acurácia de 89,09%, sensibilidade de 92,00%, especificidade de 86,67% e AUC de 0,9187. Esse resultado sugere que a integração das variáveis clínicas avaliadas pode contribuir para diferenciar pacientes com biópsia positiva e negativa, especialmente em situações nas quais a decisão pela investigação invasiva permanece incerta.

O PSA total foi a variável isolada com maior capacidade discriminativa, apresentando AUC de 0,795. As demais variáveis apresentaram desempenho individual limitado, mas podem fornecer informação complementar quando analisadas conjuntamente. Esse achado reforça a importância de modelos multivariados, nos quais a classificação não depende exclusivamente de um único marcador laboratorial ou de um único achado clínico.

O desempenho superior do Naive Bayes, em comparação com os demais algoritmos avaliados neste estudo, pode estar relacionado à adequação de sua estrutura probabilística às características da base analisada e ao tamanho amostral disponível. Em comparação com publicação prévia do próprio grupo, baseada exclusivamente em Rede Neural Artificial e no mesmo conjunto de 274 prontuários, que alcançou acurácia de 80%, sensibilidade de 80%, especificidade de 80% e AUC de 0,9027, o Naive Bayes apresentou, nesta análise, valores superiores de acurácia, sensibilidade, especificidade e AUC [20]. Essa comparação sugere que a avaliação de diferentes algoritmos sobre a mesma base de dados pode contribuir para identificar o classificador mais adequado ao conjunto analisado. Entretanto, esse resultado não deve ser interpretado como evidência de superioridade geral do Naive Bayes em diferentes populações ou contextos clínicos. A escolha do classificador deve permanecer condicionada às características dos dados, à qualidade das variáveis, à calibração do modelo e ao desempenho observado em validações independentes.

A simplicidade operacional do Naive Bayes constitui uma vantagem potencial para futuras implementações, pois o algoritmo possui baixa complexidade computacional e pode ser executado com recursos relativamente modestos. Contudo, este estudo não realizou uma comparação formal de tempo de processamento, consumo de memória, custo de infraestrutura ou interpretabilidade entre os algoritmos. Assim, a afirmação sobre menor custo computacional deve ser entendida como uma consideração teórica e operacional associada à natureza do algoritmo, e não como uma característica experimentalmente demonstrada nesta pesquisa. Da mesma forma, a interpretabilidade do modelo não foi avaliada por técnicas específicas, como análise da contribuição das variáveis, explicações locais ou avaliação com profissionais de saúde.

A comparação apresentada na Tabela 4 é indireta. Os estudos incluídos utilizaram bases de dados, populações, variáveis, equipamentos, padrões de referência e protocolos de validação diferentes. Por essa razão, os valores de AUC e acurácia não permitem estabelecer uma hierarquia definitiva entre os métodos. O desempenho obtido pelo Naive Bayes foi superior aos valores apresentados em parte dos trabalhos relacionados, mas essa observação deve ser

interpretada apenas como contextualização da literatura, e não como comparação direta ou demonstração de superioridade universal.

A amostra utilizada foi balanceada, com 137 pacientes com câncer e 137 pacientes sem câncer. Embora o balanceamento facilite a comparação entre os algoritmos e reduza a influência da classe majoritária durante o treinamento, a proporção de casos da amostra não necessariamente representa a prevalência da doença na população geral ou no fluxo real de atendimento. Consequentemente, as métricas obtidas, especialmente acurácia, valor preditivo positivo e valor preditivo negativo, podem variar quando o modelo for aplicado a uma população com prevalência diferente. Em uma eventual implementação clínica, será necessário avaliar a calibração do modelo e seu desempenho em cenários com distribuição de classes mais próxima da prática assistencial.

Outra limitação importante é a ausência de validação externa. O estudo foi realizado retrospectivamente, em uma única instituição e com número moderado de pacientes. Dessa forma, ainda não é possível assegurar que o desempenho observado se mantenha em outros hospitais, regiões, populações ou períodos. Diferenças no perfil epidemiológico, na seleção dos pacientes, na qualidade dos registros, na indicação de biópsia e na prevalência de comorbidades podem afetar a capacidade de generalização do modelo. Estudos futuros devem incluir validação temporal, validação externa multicêntrica e, posteriormente, avaliação prospectiva antes de qualquer utilização assistencial.

Também devem ser consideradas as limitações relacionadas à disponibilidade e à qualidade dos dados clínicos. Variáveis potencialmente relevantes, como índice de massa corporal, hiperplasia prostática benigna, histórico familiar, vasectomia e hipercolesterolemia, não estavam disponíveis de forma suficientemente completa para inclusão no modelo final. Além disso, registros retrospectivos podem conter dados ausentes, inconsistências de preenchimento e diferenças na forma de documentação do toque retal e de outros fatores clínicos.

O método proposto não substitui a avaliação do urologista nem estabelece, isoladamente, a indicação de biópsia. Seu possível uso é o de ferramenta complementar de triagem e estratificação de risco. A decisão clínica deve considerar o conjunto de informações disponíveis, incluindo história clínica, exame físico, PSA, exames de imagem, preferência do paciente e avaliação especializada. Portanto, os resultados deste estudo demonstram potencial de auxílio à decisão na base analisada, mas não comprovam, por si só, redução efetiva do número de biópsias, redução de custos ou melhoria de desfechos clínicos.

Apesar dessas limitações, o uso de variáveis de obtenção relativamente simples pode favorecer a adoção de soluções de apoio à decisão em serviços com infraestrutura tecnológica limitada. O modelo poderá futuramente ser avaliado como calculadora clínica ou módulo integrado ao prontuário eletrônico. Antes disso, será necessário realizar validação externa, analisar a calibração, investigar possíveis vieses entre subgrupos, definir limiares de decisão e avaliar o impacto da ferramenta no fluxo de atendimento.

Estudos futuros também poderão incorporar novas variáveis clínicas, laboratoriais e de imagem, além de comparar o modelo com estratégias clínicas já utilizadas na prática. Uma avaliação prospectiva deverá verificar não apenas as métricas discriminativas, mas também os efeitos sobre encaminhamentos, realização de biópsias, detecção de casos clinicamente relevantes, tempo de atendimento, custos e segurança do paciente. Somente após essas etapas será possível determinar se o método produz benefício clínico efetivo e se pode ser incorporado de forma responsável à assistência, inclusive no contexto do Sistema Único de Saúde.

5 CONCLUSÕES

Foi desenvolvido e avaliado um método de auxílio à triagem para investigação de câncer de próstata utilizando oito variáveis clínicas e cinco algoritmos de aprendizado de máquina. Na base retrospectiva e balanceada analisada, o Naive Bayes apresentou o melhor desempenho no conjunto de teste, com acurácia de 89,09%, sensibilidade de 92,00%, especificidade de 86,67% e AUC de 0,9187.

Os resultados indicam que a integração de dados clínicos rotineiros, especialmente em conjunto com o PSA total, apresenta potencial para apoiar a estratificação de risco e a decisão médica quanto à necessidade de investigação adicional. Entretanto, o estudo não demonstrou diretamente redução de biópsias, redução de custos ou melhoria de desfechos clínicos. Essas possibilidades devem ser avaliadas em estudos posteriores.

A interpretação dos resultados deve considerar que a pesquisa utilizou uma amostra balanceada, retrospectiva e proveniente de um único centro, sem validação externa. Portanto, o desempenho observado não pode ser generalizado automaticamente para outras populações ou instituições. Além disso, a comparação com estudos da literatura foi indireta, pois os trabalhos utilizaram bases de dados e protocolos metodológicos diferentes.

Como etapas futuras, recomenda-se ampliar a amostra, incorporar variáveis clínicas adicionais, realizar validação temporal e multicêntrica, avaliar a calibração e investigar o desempenho do modelo em cenário prospectivo. Também poderá ser desenvolvido um protótipo de apoio à decisão para uso por

urologistas, desde que sua implementação seja precedida por avaliação clínica, ética, regulatória e de segurança da informação.

Assim, o método proposto deve ser considerado uma ferramenta experimental de apoio à triagem, e não um substituto da avaliação médica ou um instrumento diagnóstico autônomo.

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. Ao Hospital Universitário da UFMA e aos colaboradores que contribuíram para a execução do estudo, a coleta dos dados clínicos e o desenvolvimento da pesquisa.

Financiamento. Os autores não declararam fonte de financiamento.

Disponibilidade de dados e materiais. Os autores não declararam a disponibilidade dos dados e materiais da pesquisa.

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

Declaração de uso de IA. Os autores declaram o uso de ferramenta de inteligência artificial generativa (ChatGPT) exclusivamente para apoio à adequação formal do manuscrito ao *template* do evento, organização textual e síntese do conteúdo previamente produzido pelos autores, sem geração autônoma de dados, resultados ou conclusões. Todo o conteúdo científico, a validação metodológica e a responsabilidade intelectual permanecem integralmente sob responsabilidade dos autores.

REFERÊNCIAS

1. Kim MH, Yoo S, Choo MS, Cho MC, Son H, Jeong H. The role of the serum 25-OH vitamin D level on detecting prostate cancer in men with elevated prostate-specific antigen levels. Sci Rep. 2022;12(1):14089. doi: https://doi.org/10.1038/s41598-022-17563-8
2. Lee C, Light A, Alaa A, Thurtle D, van der Schaar M, Gnanapragasam VJ. Application of a novel machine learning framework for predicting non-metastatic prostate cancer-specific mortality in men using the Surveillance, Epidemiology, and End Results (SEER) database. Lancet Digit Health. 2021;3(3):e158-e165. doi: https://doi.org/10.1016/S2589-7500(20)30314-9
3. American Cancer Society. Key statistics for prostate cancer [Internet]. 2026 [citado 2026 out 18]. Disponível em: https://www.cancer.org/cancer/types/prostate-cancer/about/key-statistics.html
4. Instituto Nacional de Câncer. Câncer de próstata [Internet]. Rio de Janeiro: INCA; 2023 [citado 2026 out 18]. Disponível em: https://www.gov.br/inca/pt-br/assuntos/cancer/tipos/prostata
5. Instituto Nacional de Câncer. Atlas de mortalidade por câncer [Internet]. Rio de Janeiro: INCA; 2021 [citado 2023 dez 14]. Disponível em: https://www.inca.gov.br/app/mortalidade
6. Cosma G, McArdle SE, Foulds GA, Hood SP, Reeder S, Johnson C, et al. Prostate Cancer: Early Detection and Assessing Clinical Risk Using Deep Machine Learning of High Dimensional Peripheral Blood Flow Cytometric Phenotyping Data. Front Immunol. 2021;12:786828. doi: https://doi.org/10.3389/fimmu.2021.786828
7. Correas JM, Halpern EJ, Barr RG, Ghai S, Walz J, Bodard S, et al. Advanced ultrasound in the diagnosis of prostate cancer. World J Urol. 2021;39(3):661-76. doi: https://doi.org/10.1007/s00345-020-03193-0
8. Hicks RM, Simko JP, Westphalen AC, Nguyen HG, Greene KL, Zhang L, et al. Diagnostic Accuracy of 68Ga-PSMA-11 PET/MRI Compared with Multiparametric MRI in the Detection of Prostate Cancer. Radiology. 2018;289(3):730-7. doi: https://doi.org/10.1148/radiol.2018180788
9. Stabile A, Giganti F, Rosenkrantz AB, Taneja SS, Villeirs G, Gill IS, et al. Multiparametric MRI for prostate cancer diagnosis: current status and future directions. Nat Rev Urol. 2020;17(1):41-61. doi: https://doi.org/10.1038/s41585-019-0212-4
10. Faceli K, Lorena AC, Gama J, Almeida TA, Carvalho ACPLF. Inteligência artificial: uma abordagem de aprendizado de máquina. 2ª ed. Rio de Janeiro: LTC; 2021.
11. Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, et al. Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. Sci Rep. 2017;7(1):15415. doi: https://doi.org/10.1038/s41598-017-15720-y
12. Liu J, Wang ZQ, Li M, Zhou MY, Yu YF, Zhan WW. Establishment of two new predictive models for prostate cancer to determine whether to require prostate biopsy when the PSA level is in the diagnostic gray zone (4-10 ng ml-1). Asian J Androl. 2020;22(2):213-6. doi: https://doi.org/10.4103/aja.aja_46_19
13. Liu J, Dong B, Qu W, Wang J, Xu Y, Yu S, et al. Using clinical parameters to predict prostate cancer and reduce the unnecessary biopsy among patients with PSA in the gray zone. Sci Rep. 2020;10(1):5157. doi: https://doi.org/10.1038/s41598-020-62015-w
14. Park JY, Yoon S, Park MS, Choi H, Bae JH, Moon DG, et al. Development and External Validation of the Korean Prostate Cancer Risk Calculator for High-Grade Prostate Cancer: Comparison with Two Western Risk Calculators in an Asian Cohort. PLoS One. 2017;12(1):e0168917. doi: https://doi.org/10.1371/journal.pone.0168917
15. Yoo S, Gujrathi I, Haider MA, Khalvati F. Prostate Cancer Detection using Deep Convolutional Neural Networks. Sci Rep. 2019;9(1):19518. doi: https://doi.org/10.1038/s41598-019-55972-4
16. Chen Y, Xu C, Zhang Z, Zhu A, Xu X, Pan J, et al. Prostate cancer identification via photoacoustic spectroscopy and machine learning. Photoacoustics. 2021;23:100280. doi: https://doi.org/10.1016/j.pacs.2021.100280
17. Lu X, Liu X, Xiao Z, Zhang S, Huang J, Yang C, et al. Self-supervised dual-head attentional bootstrap learning network for prostate cancer screening in transrectal

ultrasound images. Comput Biol Med. 2023;165:107337. doi: https://doi.org/10.1016/j.compbiomed.2023.107337

18. Zheng S, Jiang S, Chen Z, Huang Z, Shi W, Liu B, et al. The roles of MRI-based prostate volume and associated zone-adjusted prostate-specific antigen concentrations in predicting prostate cancer and high-risk prostate cancer. PLoS One. 2019;14(11):e0218645. doi: https://doi.org/10.1371/journal.pone.0218645

19. Kohavi R. A study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. International joint Conference on artificial intelligence, v. 2, p. 1137-1145, 1995. doi: https://dl.acm.org/doi/10.5555/1643031.1643047.

20. De Araujo WBD, Santana EEC, Silva NP, Sousa Junior CM, Moura GL, Dos Santos JALM, et al. Rede neural artificial aplicada ao diagnóstico de câncer de próstata. J Health Inform. 2024;16(Especial). doi: https://doi.org/10.59681/2175-4411.v16.iespecial.2024.1371

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