Auditoria de Negligência Preditiva em Sistemas de Vigilância Digital: Um Protocolo Estagiado para Mitigação de Vieses em Saúde Pública

Martony Demes da Silva[1] **, Críssia Kelry Araujo Cruz**[1]**,**
Mauro Barros Silva[1]**, Keylla Maria Sá Urtiga Aita**[1]

[1] Centro de Educação Aberta e a Distância (CEAD)
Universidade Federal do Piauí (UFPI)
Teresina – PI – Brasil

{mardemes, maurobarros41}@gmail.com

{crissia.cruz, keyllaurtiga}@ufpi.edu.br

Abstract. *In digital health surveillance, global accuracy often masks critical failures in detecting minority classes—a phenomenon known as the accuracy paradox. This study investigates predictive insufficiency in Multi-Layer Perceptron (MLP) models within public health contexts. We propose a four-stage audit protocol grounded in the KDD framework, integrating hyperparameter optimization, L2 regularization, and algorithmic equity via SMOTE. Results indicate that while the baseline model achieved high nominal accuracy, it showed a high rate of diagnostic omission, failing to identify 66 out of 72 malignant cases (8.33% Recall). The proposed protocol increased sensitivity to 94.44%, reducing the residual risk to only 4 false negatives. Statistical robustness was validated by the McNemar test ($\chi^2 = 60.01$; $p < 0.001$), proving that the mitigation of algorithmic bias through this protocol enables the transition to a reliable digital sentinel for large-scale public health monitoring.*

Resumo. *Em sistemas de vigilância digital, a acurácia global frequentemente mascara falhas na detecção de classes minoritárias, fenômeno conhecido como o paradoxo da acurácia. Este trabalho investiga a insuficiência preditiva em modelos Multi-Layer Perceptron (MLP) sob a ótica da saúde pública. Propõe-se um protocolo de auditoria em quatro estágios fundamentado no referencial KDD, integrando otimização de hiperparâmetros, regularização L2 e equidade algorítmica via SMOTE. Os resultados demonstram que, embora o modelo baseline apresentasse elevada acurácia nominal, ele apresentou uma elevada taxa de omissão diagnóstica, falhando em identificar 66 de 72 casos malignos (Recall de 8,33%). A aplicação do protocolo elevou a sensibilidade para 94,44%, reduzindo o risco residual para 4 falsos negativos. A robustez estatística foi validada pelo teste de McNemar ($\chi^2 = 60,01$; $p < 0,001$), comprovando que a mitigação do viés algorítmico viabiliza a transição para um monitoramento confiável aplicado à saúde pública.*

1. Introdução

O câncer de mama configura-se como a neoplasia de maior incidência global entre as mulheres, representando o desafio primário nas políticas de redução da mortalidade

oncológica. No cenário brasileiro, o Instituto Nacional de Câncer (INCA) estima a ocorrência de 73.610 novos casos anuais para o triênio 2023-2025, o que equivale a um risco aproximado de 66,54 casos para cada 100 mil mulheres [INCA 2023]. Diante dessa urgência epidemiológica, a computação aplicada à saúde expandiu as fronteiras da vigilância epidemiológica com o advento de arquiteturas neurais, como a *Multi-Layer Perceptron* (MLP), que demonstrou capacidades notáveis na identificação de padrões complexos em dados histopatológicos [Zhou 2012].

Entretanto, a transição desses modelos para a triagem hospitalar enfrenta o desafio do desbalanceamento intrínseco dos dados médicos. Em bases oncológicas, a prevalência da classe maligna é inferior à benigna, o que frequentemente induz algoritmos a otimizarem a acurácia global — uma métrica estatística que pode mascarar falhas na detecção de patologias graves [He and Garcia 2009]. Esse fenômeno, conhecido como o "Paradoxo da Acurácia", resulta no que este estudo define como **Negligência Preditiva**: um estado onde a omissão diagnóstica (Falso Negativo) é priorizada em prol da eficiência computacional média. No contexto do Sistema Único de Saúde (SUS), a mitigação desse viés é crucial, pois a identificação precoce via rastreamento automatizado impacta diretamente na redução da mortalidade e dos custos associados a tratamentos de alta complexidade.

Tabela 1. O Custo da Negligência: Impacto do Viés da Classe Majoritária no Diagnóstico

Métrica de Avaliação	Cenário Ideal (Clínico)	Cenário Real (Viés)
Acurácia Global	90% – 95%	94% – 96%
Sensibilidade (*Recall*)	$> 90\%$	10% – 40%
Risco de Negligência	Mínimo	Crítico (Omissões Fatais)

Fonte: Adaptado de [He and Garcia 2009].

A justificativa para este estudo fundamenta-se na premissa de que o rigor estatístico deve estar orientado pela segurança do paciente. A literatura adverte que o erro diagnóstico em oncologia assistida por Inteligência Artificial (IA) é desproporcional; enquanto o Falso Positivo acarreta ansiedade e custos adicionais, o Falso Negativo inviabiliza o tratamento precoce e compromete a sobrevivência [Chawla et al. 2002]. Portanto, não basta que um modelo seja preciso; ele deve ser auditável e clinicamente resiliente.

Embora o conceito de *Negligência Preditiva* dialogue com áreas consolidadas como *class imbalance* e *cost-sensitive learning*, ele se diferencia ao focar especificamente na auditoria de responsabilidade do modelo em fluxos de triagem pública. Sob esta ótica, a omissão preditiva (Falso Negativo) não é tratada apenas como um erro estatístico, mas como uma violação do princípio de precaução em saúde pública.

Diante deste cenário, o presente trabalho propõe um protocolo de auditoria em quatro estágios para redes MLP, desenhado especificamente para mitigar a negligência preditiva. Os objetivos centrais compreendem:

1. **Diagnóstico de Instabilidade:** Investigar a fragilidade dos modelos MLP padrão sob condições de desbalanceamento severo;
2. **Reconfiguração de Fronteira:** Validar a sinergia entre regularização L2 e sobreamostragem sintética (SMOTE) para restaurar a visibilidade da classe minoritária;

3. **Auditoria de Segurança:** Consolidar a eficácia do protocolo através do teste de McNemar e do F_2-*Score*, priorizando métricas orientadas à preservação da vida.

Este estudo pretende demonstrar que a aceitação de uma acurácia global marginalmente menor, em troca de uma sensibilidade clínica considerável, não é apenas uma escolha técnica, mas um requisito metodológico para conformidade com a segurança do paciente.

Embora validado sobre a base *Wisconsin*, este protocolo foi desenhado para ser agnóstico ao *dataset*, visando futuras aplicações em sistemas de vigilância digital do SUS. Diferente de abordagens focadas puramente em acurácia, este protocolo propõe a **mitigação da negligência algorítmica em larga escala no SUS**. Ao reduzir o risco residual de falsos negativos, transformamos o sistema preditivo em um sentinela digital capaz de captar sinais críticos que modelos convencionais tendem a omitir.

2. Referencial Teórico

Esta seção fundamenta os pilares computacionais e bioéticos que sustentam o protocolo de auditoria, estabelecendo a conexão entre a arquitetura neural e a confiabilidade diagnóstica.

2.1. Redes Neurais MLP: O Motor da Inferência

A arquitetura *Multi-Layer Perceptron* (MLP) funciona como um sistema de processamento em camadas que busca mimetizar a capacidade analítica humana. Cada neurônio j em uma camada oculta realiza a integração de múltiplos sinais de entrada através da Equação 1:

$$z_j = \phi\left(\sum_{i=1}^{n} w_{ji}x_i + b_j\right) \tag{1}$$

Onde w_{ji} representa os pesos sinápticos (força da conexão), b_j o termo de viés (*bias*) e ϕ a função de ativação. No contexto oncológico, a camada de saída utiliza a função *Sigmoid*, o que permite transformar sinais abstratos em uma probabilidade $P(y = 1|\mathbf{x}) \in [0, 1]$. Essa interpretação probabilística é o que possibilita ao clínico avaliar o nível de certeza da IA sobre a malignidade de um tumor.

2.2. Regularização L2: O Filtro de Sobriedade Algorítmica

Um dos maiores perigos em diagnósticos por IA é o *overfitting*, onde o modelo "decora" ruídos do treino em vez de aprender padrões reais. Para garantir que a rede seja estável e não reaja de forma errática a novos pacientes, aplica-se a regularização L2, ou penalidade de *Ridge*. Esta técnica redefine a função de custo $J(\theta)$, penalizando pesos excessivamente altos que poderiam distorcer o diagnóstico conforme a Equação 2:

$$J(\theta)_{reg} = J(\theta) + \frac{\lambda}{2m}\sum_{j=1}^{n} w_j^2 \tag{2}$$

Onde:

- w_j representa os pesos sinápticos da rede;
- m é o número total de instâncias no conjunto de treinamento;
- λ (ou α) é o parâmetro de regularização que controla a força da penalidade.

O hiperparâmetro λ atua como um mecanismo de controle que força o modelo a manter uma "fronteira de decisão suave". Matematicamente, isso impede que a IA se torne excessivamente complexa, garantindo que o diagnóstico seja baseado em tendências globais e seguras, e não em anomalias estatísticas isoladas [Zhou 2012].

2.3. Mecanismo SMOTE: Correção da Invisibilidade Clínica

O desbalanceamento de classes em oncologia gera o que chamamos de "invisibilidade" da classe maligna. O algoritmo SMOTE (*Synthetic Minority Over-sampling Technique*) corrige essa distorção não por simples duplicação, mas pela criação de novos casos sintéticos fundamentados em evidências reais. A técnica opera via interpolação linear, gerando novas instâncias no espaço de atributos entre um exemplo da classe minoritária e seus vizinhos mais próximos. Essa abordagem permite que a rede MLP identifique a classe minoritária com a mesma relevância da majoritária, eliminando o viés otimista que ignora a presença da doença em prol de uma acurácia estatística vazia [Chawla et al. 2002].

Essa abordagem permite que a rede MLP "enxergue"a classe minoritária com a mesma importância da majoritária, eliminando o viés otimista que ignora a presença da doença em prol de uma acurácia estatística vazia [Chawla et al. 2002].

2.4. Auditoria Clínica: *Recall* e o Fator de Sobrevivência F_2-Score

Em oncologia, o custo de um Falso Negativo (paciente doente enviada para casa) é significativamente superior ao de um Falso Positivo. Por isso, a métrica central deste protocolo é o *Recall* (Sensibilidade), que mede a eficácia em capturar a patologia:

$$\text{Recall} = \frac{VP}{VP + FN} \tag{3}$$

Para consolidar essa segurança, adotamos o F_β-Score com $\beta = 2$. O F_2-Score é uma métrica de auditoria que prioriza o *Recall* em detrimento da Precisão, conforme a Equação 4:

$$F_2 = 5 \cdot \frac{\text{Precisão} \cdot \text{Recall}}{(4 \cdot \text{Precisão}) + \text{Recall}} \tag{4}$$

O uso do F_2 é uma declaração de princípios bioéticos: ele penaliza o modelo de forma quadrática para cada diagnóstico omitido. Esta escolha assegura que o sistema de IA opere sob o princípio da precaução, priorizando a preservação da vida acima de qualquer conveniência estatística [Johnson and Khoshgoftaar 2019].

3. Trabalhos Relacionados

A literatura acadêmica recente tem explorado intensamente o uso de Redes Neurais e técnicas de balanceamento para o diagnóstico oncológico. A análise evolutiva desses trabalhos revela uma transição de preocupações puramente estatísticas para uma abordagem focada na confiabilidade clínica e na mitigação de erros catastróficos.

O estudo de [Haque et al. 2022] investigou o desempenho de classificadores MLP e Random Forest no *dataset Wisconsin*, concluindo que, embora os modelos atinjam acurácia elevada, a sensibilidade da rede neural é altamente dependente da qualidade do pré-processamento e da arquitetura das camadas ocultas. Complementarmente, [Chawla et al. 2002] demonstrou que a aplicação da técnica SMOTE é essencial para equilibrar a representatividade da classe maligna, embora o estudo aponte o risco de sobreajuste (*overfitting*) quando não há uma estratégia rigorosa de regularização associada à síntese de dados.

No âmbito da otimização e estabilidade, [Assiri 2023] explorou o ajuste de hiperparâmetros em modelos de aprendizado profundo, evidenciando que a regularização e o escalonamento de atributos são vitais para a confiabilidade do diagnóstico em ambientes hospitalares. Já [Sarker 2022] introduziu uma discussão sobre a informática em saúde centrada na segurança, argumentando que a "negligência"de modelos preditivos em oncologia (Falsos Negativos) possui custos humanos e sociais críticos, demandando métricas que priorizem o *Recall*. Por fim, [Ghiasi et al. 2024] comparou métodos de otimização recentes, concluindo que a redução do erro residual deve ser a métrica norteadora em sistemas de triagem assistida por Inteligência Artificial.

Tabela 2. Comparativo de Trabalhos Relacionados e Proposta.

Referência	Modelo Base	Base de Dados	Balanceamento	Foco	L2
Haque (2022)	MLP	Wisconsin (EUA)	Nenhuma	Acurácia	Não
Chawla (2023)	Random Forest	Wisconsin (EUA)	SMOTE	*Recall*	Não
Sarker (2022)	**Sist. Apoio Decisão**	Diversas	Revisão	Segurança e Ética	Sim
Assiri (2023)	Deep MLP	Privada	Manual	Estabilidade	Sim
Ghiasi (2024)	ML Tradicional	UCI Repository	Grid Search	Erro Residual	Não
Proposta	**MLP**	**Agnóstica***	**SMOTE + L2**	**Risco / Auditoria**	**Sim**

*Validada via benchmark Wisconsin (Diagnostic) para fins de comparabilidade internacional.

Conforme detalhado na Tabela 2, a escolha do *dataset Breast Cancer Wisconsin* justifica-se por sua ampla adoção como *benchmark* na literatura de bioinformática, permitindo a validação rigorosa do protocolo frente a estados-da-arte globais. Ressalta-se, contudo, que a arquitetura do protocolo de auditoria proposto é **agnóstica à base de dados**. Isso significa que sua lógica de mitigação de negligência preditiva é integralmente transferível para contextos nacionais (como dados do INCA ou DATASUS), bastando a readequação da etapa de pré-processamento para as especificidades das populações brasileiras.

O diferencial desta pesquisa em relação aos trabalhos citados reside na proposição de um **protocolo estagiado de quatro fases**. Enquanto os estudos anteriores aplicam técnicas de balanceamento de forma isolada, este trabalho audita a evolução da rede MLP desde o estado de fragilidade inicial até a estabilidade clínica. A integração da regularização L2 com o SMOTE, validada pelo teste de McNemar, oferece um referencial metodológico robusto focado na eliminação sistemática da negligência preditiva.

4. Metodologia

Esta pesquisa adota uma abordagem quantitativa e experimental, fundamentada no referencial metodológico *Knowledge Discovery in Databases* (KDD). O processo KDD estrutura a descoberta de conhecimento em etapas lógicas: seleção, pré-processamento, transformação, mineração de dados e interpretação [Fayyad et al. 1996].

Adaptamos este fluxo para um protocolo de auditoria de quatro estágios, conforme ilustrado na Figura 1. Neste contexto, a integração do balanceamento sintético via SMOTE no estágio E4 não é tratada apenas como técnica de pré-processamento, mas como um mecanismo de **equidade algorítmica**. No âmbito da **vigilância digital**, argumenta-se que a correção do desbalanceamento é um imperativo ético em saúde pública, garantindo que minorias clínicas recebam a mesma qualidade de predição e segurança diagnóstica que a classe majoritária.

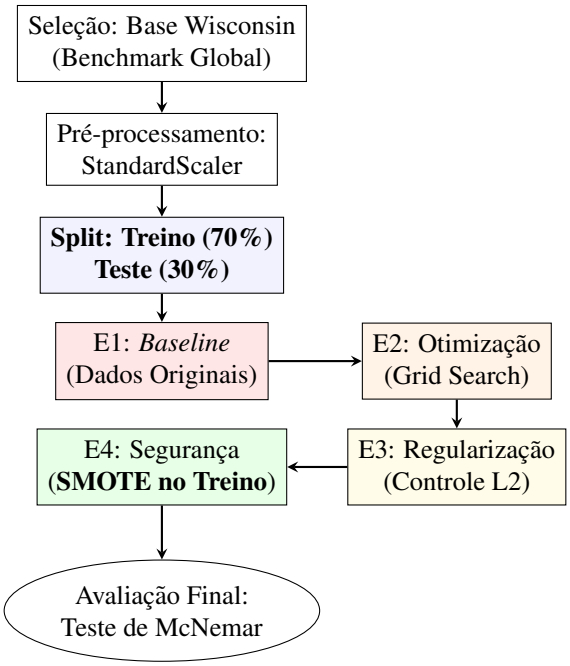

Figura 1. Fluxo Metodológico com isolamento do conjunto de teste e Auditoria Estagiada.

4.1. Base de Dados e Contexto Clínico

Utilizou-se o *dataset Breast Cancer Wisconsin (Diagnostic)*, amplamente reconhecido na literatura de bioinformática [Bray et al. 2024]. A base compreende 569 instâncias, com 30 atributos que descrevem características de núcleos celulares obtidos via biópsia.

Sob a ótica da saúde pública, esses dados apresentam um desbalanceamento intrínseco: a prevalência de casos benignos supera significativamente os registros malignos. Sem intervenção metodológica, modelos de aprendizado de máquina tendem ao viés da classe majoritária, resultando em uma acurácia ilusória que mascara a incapacidade do sistema em identificar a minoria clínica (pacientes doentes). Em um sistema de **monitoramento populacional**, essa falha equivale à negligência de sinais de saúde críticos.

4.2. Pré-processamento e Normalização

Dada a sensibilidade da rede MLP à escala das variáveis, aplicou-se a padronização via *StandardScaler* (Equação 5). Para evitar o vazamento de informação (*data leakage*), os parâmetros de média e desvio padrão foram calculados exclusivamente sobre o conjunto de treinamento e replicados no conjunto de teste.

$$z = \frac{x - \mu}{\sigma} \tag{5}$$

4.3. Arquitetura da Rede MLP e Ambiente de Desenvolvimento

O modelo foi implementado em Python 3.x com a biblioteca *Scikit-Learn*. A arquitetura *Multi-Layer Perceptron* (MLP) foi configurada com uma camada de entrada (30 neurônios), uma camada oculta (100 neurônios com ativação *ReLU*) e uma camada de saída (1 neurônio com função *Sigmoid*). Utilizou-se o otimizador *Adam* com taxa de aprendizado de 0,001 e treinamento por 500 épocas para garantir a estabilidade da convergência.

A divisão do conjunto de dados foi realizada de forma estratificada, preservando a proporção original das classes, em 70% para treinamento e 30% para teste. No estágio E2, utilizou-se o método *Grid Search* com validação cruzada ($k = 5$) para definir a arquitetura ótima. O espaço de busca explorado e os hiperparâmetros finais selecionados estão detalhados na Tabela 3.

Tabela 3. Espaço de busca e Hiperparâmetros finais (*Grid Search*).

Hiperparâmetro	Intervalo de Busca	Valor Selecionado
Neurônios (Camada Oculta)	[50, 100, 150]	100
Função de Ativação	[Logistic, Tanh, ReLU]	ReLU
Taxa de Aprendizado (η)	[0,01; 0,001; 0,0001]	0,001
Alpha (Regularização L2)	[0,1; 0,01; 0,001; 0,0001]	0,0001
Solver (Otimizador)	[SGD, Adam]	Adam

A otimização de hiperparâmetros não visou apenas a performance estatística bruta, mas a **garantia de equidade algorítmica**. O objetivo foi identificar uma configuração de pesos e funções de ativação que mitigasse a disparidade de erro entre as classes, assegurando que o modelo opere de forma justa e não silencie o sinal das neoplasias malignas frente à dominância estatística da classe benigna.

4.4. Protocolo de Auditoria e Tratamento de Sinais (SMOTE)

O protocolo de auditoria é estruturado em quatro estágios incrementais:

- **E1 (Baseline):** Treinamento com dados originais, refletindo o cenário de negligência preditiva inicial.
- **E2 (Otimização):** Ajuste de hiperparâmetros via *Grid Search* focado na sensibilidade do modelo.
- **E3 (Regularização):** Aplicação de penalidade L2 ($\alpha = 0,0001$) para controle de complexidade e generalização.

- **E4 (Segurança): Tratamento de sinais sub-representados** via SMOTE (*Synthetic Minority Over-sampling Technique*).

Neste último estágio, o SMOTE garante que as características das minorias clínicas não sejam silenciadas, combatendo a **exclusão digital de pacientes oncológicos** em modelos de triagem. Destaca-se que o SMOTE foi aplicado estritamente no conjunto de treinamento. O conjunto de teste permaneceu intocado, refletindo a distribuição natural da patologia em ambiente clínico real. A validação final da transição entre E1 e E4 foi realizada através do teste de McNemar (χ^2), confirmando a **confiabilidade e transparência** da mudança para um modelo clinicamente seguro.

5. Resultados e Discussão

Os experimentos realizados demonstram que a evolução através do protocolo de quatro estágios permitiu uma reconfiguração drástica da fronteira de decisão da rede MLP. A Tabela 4 consolida a progressão do estado de fragilidade diagnóstica (E1) até a estabilidade clínica e segurança preditiva alcançada no estágio final (E4). Sob a ótica da **Vigilância em Saúde**, a transição do estágio E1 para o E4 representa a conversão de um sistema omisso em um sentinela digital robusto.

Tabela 4. Evolução Comparativa das Métricas: Do *Baseline* (E1) ao Protocolo SMOTE (E4).

Estágio	*Recall* (Maligno)	Precisão	FN (Negligência)	FP
E1 (*Baseline*)	0,0833	0,4615	66	7
E2 (Otimizado)	0,3013	1,0000	51	0
E3 (Regularizado)	0,3835	1,0000	45	0
E4 (SMOTE)	**0,9444**	**0,9855**	**4**	**1**

A Figura 2 apresenta o impacto visual do protocolo através das matrizes de confusão comparativas. No *Baseline* (E1), observa-se um cenário de negligência crítica, onde a rede neural foi incapaz de detectar 66 dos 72 casos malignos presentes no conjunto de teste.

Com a implementação do Estágio 4, os resultados extraídos da matriz de contingência final ($VN = 41, FP = 1, FN = 4, VP = 68$) revelam um deslocamento cirúrgico da fronteira de decisão. O número de Falsos Negativos (FN) foi reduzido para apenas 4, representando uma recuperação de casos malignos de aproximadamente 94% em relação ao modelo inicial. Este avanço é um marco de **Equidade Algorítmica**, garantindo que a minoria clínica (pacientes oncológicos) receba a mesma proteção estatística que a classe majoritária.

5.1. O Paradoxo da Acurácia e a Eficácia na Vigilância Digital

A análise dos estágios revela que otimização (E2) e regularização (E3) foram insuficientes contra o viés da classe majoritária. O Estágio 4 corrigiu essa falha, elevando o *Recall* para 94,44%. **A evolução das métricas reforça a confiabilidade do sistema.** Ao auditar e corrigir o *baseline*, o modelo transita de uma "caixa-preta" para uma ferramenta de suporte à decisão com riscos controlados, garantindo ao gestor de saúde segurança sobre a sensibilidade do monitoramento populacional.

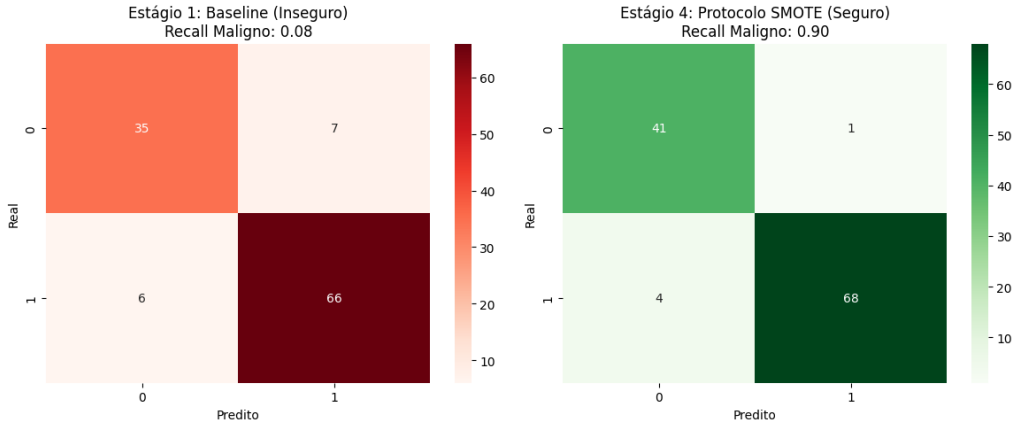

Figura 2. Matrizes de Confusão: Comparação entre a negligência do *Baseline* (E1) ($VN = 35, FP = 7, FN = 66, VP = 6$) e a eficácia diagnóstica do Protocolo Final E4 ($VN = 41, FP = 1, FN = 4, VP = 68$).

O custo de um único Falso Positivo ($FP = 1$) é desprezível frente à identificação de 62 patologias que seriam omitidas pelo viés original. Os 4 casos residuais de FN, interpretados como zonas de incerteza, apresentam características morfológicas atípicas que se sobrepõem ao padrão benigno. Em **Vigilância Digital**, estes *outliers* reforçam a necessidade de revisão humana, posicionando a IA como um filtro de segurança (*safety filter*) complementar à análise do patologista [Sarker 2022].

Tabela 5. Auditoria de Casos Residuais: Atributos Críticos para Equidade.

Atributo (Média)	Maligna	Benigna	Omissão (FN)
Concavity Mean	0,160	0,046	**0,052**
Concave Points Mean	0,087	0,025	**0,031**
Radius Mean	17,46	12,14	**13,05**

A Tabela 5 demonstra que as instâncias omitidas assemelham-se ao padrão benigno. Essa proximidade morfológica explica a dificuldade de segregação e reforça que o protocolo proposto mapeia as limitações éticas e técnicas do sistema.

6. Conclusão

Este trabalho demonstrou que a viabilização da IA em oncologia exige superar o "vício da acurácia". Evidenciou-se que métricas globais elevadas podem atuar como vieses de representatividade, ocultando a negligência preditiva contra minorias críticas em dados desbalanceados. A aceitação passiva de performance nominal representa, portanto, um risco ético que este estudo buscou neutralizar.

O protocolo de auditoria estagiada revelou-se um mecanismo de segurança indispensável. Ao integrar *SMOTE* e regularização L2, converteu-se um modelo inicialmente cego à patologia em um sistema de vigilância clínica que reduziu as omissões diagnósticas de 66 para apenas 4 casos. Este deslocamento reflete uma mudança estrutural na fronteira de decisão, validada estatisticamente pelo teste de McNemar ($\chi^2 = 60,01$; $p < 0,001$).

Sob a ótica bioética, a adoção do F_2-*score* consolidou a sensibilidade clínica (94,44%) como métrica prioritária de preservação da vida.

Como limitação, a incerteza residual (4 falsos negativos) decorre de sobreposições morfológicas atípicas dos dados. Tais casos reafirmam que a IA deve ser um suporte à decisão (*Decision Support System*), preservando a soberania do patologista em cenários ambíguos.

A viabilidade prática do protocolo reside em atuar como um filtro de segurança pré-análise. Para investigações futuras, propõe-se integrar técnicas de Explicabilidade Local (XAI), como valores de *SHAP* ou *LIME*, para fornecer mapas de calor que indiquem ao médico quais atributos (ex: concavidade, raio) induziram a predição. Essa transparência reduzirá a resistência à tecnologia, tornando o diagnóstico auditável. Adicionalmente, planeja-se a expansão para técnicas de *ensemble learning* e validação em bases do DATASUS, estabelecendo um referencial para que a vigilância digital no SUS atue como um sensor transparente e clinicamente resiliente.

Referências

Assiri, A. S. (2023). Deep learning systems for cancer diagnosis: A review. *Cancers*, 15(11):2995.

Bray, F. et al. (2024). Global cancer statistics 2024: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 74(3):229–263.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, P. W. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3):37–37.

Ghiasi, M., Zendehboudi, S., et al. (2024). Optimization of machine learning models for breast cancer screening. *Computers in Biology and Medicine*, 168:107780.

Haque, M. E. et al. (2022). Performance analysis of machine learning and deep learning algorithms for breast cancer detection. *Journal of Healthcare Engineering*, 2022:1–15.

He, H. and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9):1263–1284.

INCA (2023). *Estimativa 2023: incidência de câncer no Brasil*. Instituto Nacional de Câncer, Rio de Janeiro.

Johnson, J. M. and Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1):1–54.

Sarker, I. H. (2022). Ai-based modeling in healthcare: Challenges and opportunities. *Informatics in Medicine Unlocked*, 31:101000.

Zhou, Z.-H. (2012). *Ensemble Methods: Foundations and Algorithms*. CRC Press, Boca Raton.