

Impacto de Mecanismos de Ausência de Dados na Performance de um Sistema de Monitoramento de COVID-19

Leonardo S. Cardoso¹, Afonso M. S. Lima¹,
Agma J. M. Traina¹, Elaine P. M. de Sousa¹

¹ Instituto de Ciências Matemáticas e de Computação (ICMC)
Universidade de São Paulo (USP)
São Carlos – SP – Brasil

{leo.silva.cardoso, afonso.matheus}@usp.br
{agma, parros}@icmc.usp.br

Abstract. *The integration of wearable technology into remote health monitoring has revolutionized patient oversight by enabling continuous data collection and early clinical intervention. However, these systems frequently suffer from missing data, which can introduce bias and impair clinical decision-making. This scientific initiation work investigates the impact of three missing mechanisms (i.e., MCAR, MAR, and MNAR) simulated on heart rate data used for wearable-based alerts in COVID-19 detection task, assessing how missing distributions and rates affect alert reliability. Our findings indicate that while the system maintains high precision across most scenarios, it suffers a substantial decline in recall, leading to missed COVID-19 alerts.*

Resumo. *A integração de dispositivos vestíveis ao monitoramento remoto de saúde tem transformado a supervisão de pacientes, viabilizando a coleta contínua de dados e a intervenção clínica precoce. No entanto, esses sistemas frequentemente apresentam dados ausentes, o que pode introduzir vies e comprometer a tomada de decisão clínica. Este trabalho de iniciação científica investiga o impacto de três mecanismos de ausência (MCAR, MAR e MNAR) simulados sobre dados de frequência cardíaca utilizados para geração de alertas na detecção de COVID-19, avaliando como diferentes distribuições e taxas de ausência afetam a confiabilidade dos alertas. Os resultados indicam que, embora o sistema mantenha alta precisão na maioria dos cenários, há uma queda substancial na revocação, resultando na supressão de alertas de COVID-19.*

1. Introdução

Nas últimas décadas, a crescente preocupação com a saúde individual impulsionou a adoção de dispositivos vestíveis (*wearables*) pela população em geral, tornando-os ferramentas cada vez mais presentes no cotidiano [Canali et al. 2022]. No entanto, a qualidade dos dados coletados por esses dispositivos é frequentemente comprometida pela perda

Agradecimentos Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - processo 001, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - processos 2023/18026-8, 2024/13328-9, 2025/27722-3 e 2026/12389-0.

de informação, tornando a falta de dados um problema recorrente em aplicações reais [Emmanuel et al. 2021, Lima and Sousa 2024].

Este trabalho de iniciação científica tem como objetivo avaliar o impacto de dados ausentes na geração de alertas vermelhos (i.e., anomalias) pelo algoritmo *NightSignal* [Alavi et al. 2022], um sistema projetado para a detecção de sinais fisiológicos precoces de infecções, como a COVID-19, fundamentado na análise da frequência cardíaca em repouso (*Resting Heart Rate* — RHR). Essa abordagem justifica-se com a finalidade de avaliação de como a perda de integridade da série temporal pode mascarar alterações fisiológicas críticas e comprometer a sensibilidade clínica do sistema. Para isso, foram considerados três mecanismos de ausência de dados: *Missing Completely at Random* (MCAR), *Missing at Random* (MAR) e *Missing Not at Random* (MNAR) [Emmanuel et al. 2021]. Para esse estudo, foram avaliados 20 pacientes que testaram positivo para COVID-19 e apresentaram sintomas da doença [Alavi et al. 2022].

Nossos resultados mostraram que o MCAR obteve o menor impacto negativo entre os mecanismos analisados, com uma pequena redução em todas as métricas avaliadas — Acurácia, Precisão e Revocação —, mesmo com o aumento da taxa de ausência. Já o MNAR foi o mecanismo que causou maior perda, tendo uma significativa redução da Acurácia e da Revocação, conforme se elevava a taxa de ausência. Além disso, obtivemos um resultado importante da Precisão: para todos os mecanismos avaliados, ela se manteve alta, indicando uma pequena presença de alertas falsos gerados pelo algoritmo sob condições de ausência.

2. Fundamentação

No contexto da pandemia de COVID-19, os dispositivos vestíveis ganharam relevância médica ao possibilitar a coleta contínua de sinais fisiológicos — como batimentos cardíacos e contagem de passos — que se mostraram úteis na detecção pré-sintomática da doença [Mishra et al. 2020]. Porém, em condições reais de uso, a ocorrência de registros ausentes em séries temporais de sensores vestíveis é frequente. Em aplicações críticas, como a detecção de infecções, compreender como essas ausências afetam o desempenho dos algoritmos de análise é essencial para avaliar sua confiabilidade em condições reais de uso, uma vez que ignorá-las ou omiti-las pode levar a análises tendenciosas ou imprecisas [Getzen et al. 2023, Benhamza et al. 2025].

Mecanismo de Ausência	Definição	Exemplo
<i>Missing Completely At Random</i> (MCAR)	A causa da ausência não está relacionada a outros valores observados ou não observados.	O dispositivo vestível falhou e o valor não foi gerado.
<i>Missing At Random</i> (MAR)	A causa da ausência depende de outros atributos observados.	O paciente removeu o sensor vestível durante as horas de sono (“frequência cardíaca” depende de “data/hora”).
<i>Missing Not at Random</i> (MNAR)	A causa da ausência está relacionada a informações específicas sobre o próprio valor ausente, que não estão presentes no conjunto de dados.	O sensor vestível perde leituras durante episódios de taquicardia (“frequência cardíaca” depende de si mesma).

Tabela 1. Mecanismos de dados ausentes [Lima 2025]

Na literatura, os mecanismos de ausência de dados são classificados em três categorias, cada uma refletindo um padrão distinto de ocorrência [Emmanuel et al. 2021].

O primeiro, denominado *Missing Completely At Random* (MCAR), ocorre quando a ausência de um registro é independente de qualquer outro valor do conjunto de dados. O segundo, *Missing At Random* (MAR), pressupõe que a probabilidade de ausência depende de outras variáveis observadas, mas não do próprio valor ausente. Por fim, o *Missing Not At Random* (MNAR) representa o cenário mais crítico: a ausência está diretamente relacionada ao valor que seria registrado. A distinção entre esses mecanismos é fundamental para a avaliação de robustez de algoritmos clínicos, uma vez que cada um introduz padrões de ausência com implicações distintas sobre a integridade dos dados analisados. A Tabela 1 sintetiza tais mecanismos.

Para avaliar o impacto da simulação da ausência de dados, foi utilizado um algoritmo de detecção precoce de anomalias fisiológicas, o *NightSignal* [Alavi et al. 2022]. Seu principal objetivo é identificar sinais de infecções, como a COVID-19, ainda em fase pré-sintomática. O *pipeline* de processamento utiliza séries temporais com resolução de um minuto, contendo registros de frequência cardíaca e contagem de passos. Há três níveis de alerta: os alertas verdes, que indicam normalidade fisiológica; os amarelos, aviso de elevação moderada; e os vermelhos, uma anomalia crítica sustentada.

Neste trabalho, analisamos especificamente como a ausência de dados impacta os alertas vermelhos gerados pelo *NightSignal*. Essa escolha se justifica pelo fato de que eles representam o produto final e mais crítico do algoritmo, sendo diretamente responsáveis por sinalizar ao usuário a possível ocorrência de um evento fisiológico anômalo. Considerando que o *NightSignal* opera sobre séries temporais contínuas de frequência cardíaca noturna, a ausência de dados em noites consecutivas compromete diretamente o cálculo da linha de base individual, que é o elemento central do mecanismo de detecção [Alavi et al. 2022]. Investigar esse impacto é, portanto, fundamental para compreender os limites de confiabilidade do algoritmo em cenários reais de uso, nos quais falhas de coleta são inevitáveis.

A metodologia de avaliação de impacto de ausência aplicada neste estudo baseia-se nos trabalhos de Lima (2025) e Lima (2026). Este trabalho diferencia-se por conduzir a análise em um novo conjunto de pacientes, em uma tarefa de mineração distinta da utilizada pelo trabalho anterior, fornecendo novas evidências acerca do impacto da ausência em outra tarefa de descoberta de conhecimento.

3. Metodologia

A Figura 1 sintetiza o *pipeline* utilizado. No estudo de Alavi *et al.* (2022), foram disponibilizados dados de 3318 participantes, dentre os quais 66 correspondiam a casos sintomáticos de COVID-19. Desse grupo, 37 utilizavam dispositivos vestíveis da marca Fitbit. Visando à delimitação do escopo experimental e à otimização de recursos computacionais, uma subamostra de 20 pacientes foi selecionada aleatoriamente a partir desse subconjunto para a realização dos experimentos deste trabalho (Figura 1-A). Para cada paciente, foram disponibilizados data, horário, batimento cardíaco e número de passos registrados pelo dispositivo. Neste trabalho, os dados foram agregados por minuto, computando a média de batimentos cardíacos e de passos. Foram considerados apenas minutos para os quais houve registro de batimentos. A ausência de passos nesses minutos foi indicada com valor zero. Desse modo, as observações da série temporal de cada paciente possuem o formato <data/horário, batimentos, data/horário, passos> (e.g. 2028-10-19 00:02:00,

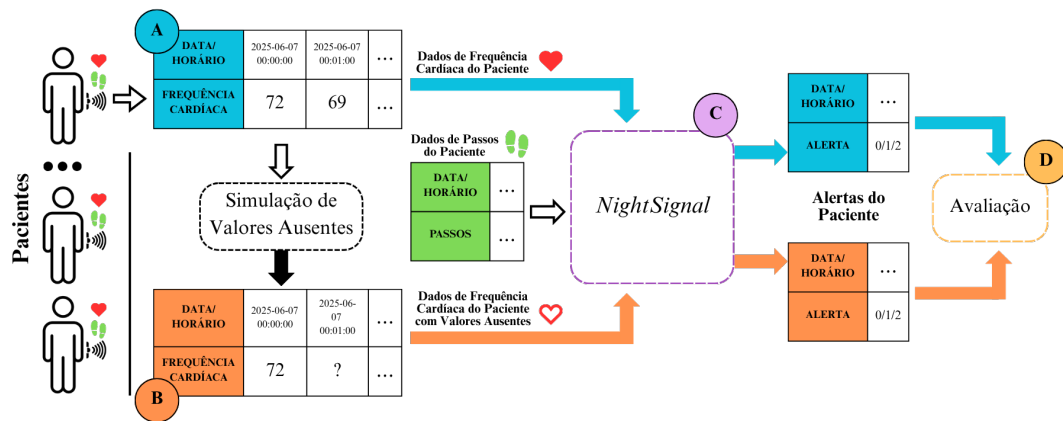


Figura 1. Pipeline para avaliação de impacto de valores ausentes

67, 2028-10-19 00:02:00, 0). As 20 séries temporais são a entrada do algoritmo *NightSignal*, que gera as séries de alertas e respectivas representações visuais (Figura 1-C).

Com o objetivo de avaliar o impacto de dados ausentes nos resultados do algoritmo analisado, foram simulados os diferentes mecanismos, MCAR, MNAR e MAR, por meio da *mdatagen*, uma biblioteca em *Python* para geração artificial de dados faltantes [Santos et al. 2019, Mangussi et al. 2025] (Figura 1-B). Motivados pela arquitetura do *NightSignal*, foram escolhidos os seguintes parâmetros: para o MAR, associa-se a ausência aos horários iniciais da janela noturna, período central para o cálculo da frequência cardíaca basal; para o MNAR, associa-se a ausência aos maiores valores de frequência cardíaca, o que pode impactar mais a tarefa de detecção de anomalias. Em cada um dos mecanismos, a taxa de ausência variou entre 5% e 25%, sendo restrita das 00:00 às 06:00, período no qual ocorre a avaliação dos batimentos em repouso pelo *NightSignal*. Essas taxas permitiram avaliar desde cenários menos extremos, até casos mais danosos ao conjunto de dados. Devido à natureza aleatória do mecanismo MCAR, os resultados de uma única simulação poderiam ser influenciados pela escolha eventual dos dados removidos. Para mitigar esse viés, cada taxa de ausência foi submetida a cinco iterações independentes (utilizando *seeds* distintas). Foram também testadas múltiplas execuções por taxa nos demais mecanismos, buscando refletir variações nos dados ausentes; contudo, não foram observadas diferenças significativas, razão pela qual a execução do *pipeline* foi padronizada em uma única iteração. O algoritmo *NightSignal* foi executado novamente para os cenários com dados ausentes simulados.

Para quantificar o impacto da perda de dados na confiabilidade do algoritmo, foi utilizada uma matriz de confusão (Figura 1-D). Nesta abordagem, os rótulos de alerta gerados pelo *NightSignal* a partir do conjunto de dados original (íntegro) foram definidos como o *ground truth*. Esses rótulos foram comparados com as previsões obtidas após a aplicação dos mecanismos de ausência de dados em diferentes taxas. A matriz permite observar não apenas se o algoritmo manteve o alerta correto, mas a natureza exata do erro introduzido pela ausência de informação. As classes consideradas foram os alertas verde (0), amarelo (1) e vermelho (2), conforme definido pela lógica de estados do algoritmo. A diagonal principal da matriz representa os casos de sucesso, onde o estado foi preservado mesmo com dados faltantes. Fora da diagonal, observam-se dois fenômenos críticos: a subestimação de risco, por alertas originalmente vermelhos que são reclassificados; ou

ainda a falsa detecção, por alertas verdes ou amarelos que tornam-se críticos.

Como o MCAR foi executado 5 vezes, foi gerada, para esse mecanismo, uma matriz de confusão acumulada. Esta matriz representa o somatório das contagens de todas as execuções para cada paciente, mecanismo e taxa, fornecendo uma visão macroscópica das transições de estado causadas pela perda aleatória de dados. A partir dessas matrizes, as métricas de desempenho Acurácia, Precisão e Revocação foram calculadas individualmente para cada paciente, execução e taxa. Dessa forma, é possível identificar qual tipo de falha (MCAR, MAR ou MNAR) é mais prejudicial à estabilidade do *NightSignal*.

4. Resultados

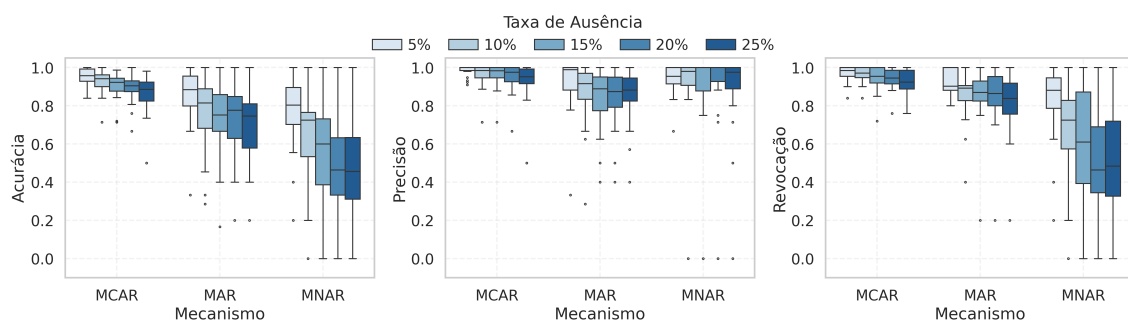


Figura 2. Distribuição das métricas de desempenho do *NightSignal* em função da taxa de ausência de dados para os mecanismos MCAR, MAR e MNAR.

Os resultados das métricas de Acurácia, Revocação e Precisão, em função dos alertas, foram organizados em *boxplots* agrupados por mecanismo de ausência e taxa, conforme apresentado na Figura 2. Os resultados mostram que o MCAR é o mecanismo que causa o menor impacto nas métricas de qualidade. Mesmo na taxa mais alta testada (25%), a mediana da Acurácia permanece acima de 0,88 e a da Revocação acima de 0,90, com *boxplots* relativamente estreitos indicando consistência entre os pacientes. Esse comportamento é esperado do ponto de vista estatístico: como o *NightSignal* computa uma média sobre centenas de leituras noturnas e em momentos com ausência de dados, a remoção aleatória de uma fração delas preserva a estimativa do ritmo cardíaco médio da noite, sem introduzir viés sistemático. Vale destacar que o MNAR com taxas mais baixas de ausência, como 5% e 10%, degrada mais os resultados da tarefa proposta do que o MCAR com 25%, gerando maior impacto tanto em Acurácia quanto em Revocação. Esse comportamento é uma evidência de que o mecanismo de ausência importa mais do que a taxa para a degradação do algoritmo.

Sobre o MAR, observa-se uma degradação progressiva e moderada à medida que a taxa aumenta. A variabilidade entre pacientes é consideravelmente maior em comparação ao MCAR, com *boxplots* mais largos e *outliers* que chegam a valores próximos de 0,17 em Acurácia, 0,30 em Precisão e 0,20 em Revocação. Isso sugere que a ausência concentrada nos horários iniciais da madrugada afeta os pacientes de maneira mais heterogênea dependendo de como cada um apresenta seus picos de RHR ao longo da janela noturna. Já o MNAR mostrou-se o mecanismo mais prejudicial ao desempenho do *NightSignal*. A Acurácia cai para mediana próxima de 0,46 na taxa de 25%, e a Revocação apresenta degradação mais acentuada entre os três mecanismos, com mediana em torno de 0,48 na

mesma taxa. Isso ocorre porque o MNAR remove preferencialmente os maiores valores de batimento cardíaco da janela noturna — exatamente os registros que elevam a média noturna acima da linha de base e acionam os alertas. Ao apagar o próprio sinal de infecção, o mecanismo força o algoritmo a subestimar sistematicamente o RHR das noites afetadas, suprimindo alertas que deveriam ser gerados.

Um resultado importante a ser analisado é o comportamento da Precisão. Nos três mecanismos e em todas as taxas avaliadas, a Precisão permanece alta, com medianas acima de 0,88 mesmo nos cenários mais adversos, e valores próximos de 1,0 para MCAR e para MNAR. Isso indica que, mesmo com a ausência de registros de batimentos cardíacos em diferentes taxas e por mecanismos distintos, os alertas gerados pelo *NightSignal* em condições de ausência de dados raramente são falsos positivos: quando o algoritmo emite um alerta, tende a estar correto quanto ao comportamento original. A degradação observada nas outras métricas ocorre fundamentalmente devido ao aumento dos falsos negativos — alertas reais que deixam de ser emitidos — e não pela geração de alertas espúrios. No contexto de um sistema de triagem de saúde, esse padrão tem implicações diretas: a ausência de dados pode tornar o algoritmo silencioso em momentos críticos, e não ruidoso, o que representa um risco clínico distinto e possivelmente mais grave do que o excesso de alarmes.

Por fim, vale ressaltar a presença de *outliers* em todos os mecanismos, especialmente nas taxas mais elevadas. Esses pontos atípicos correspondem, em sua maioria, a pacientes com poucos dias de alerta no período analisado: nesses casos, a perda ou o surgimento de um único alerta altera as métricas de forma desproporcional, resultando em valores muito baixos de Acurácia, Precisão ou Revocação. Dessa forma, os *outliers* representam problemas para o desempenho do *NightSignal*, uma vez que a ausência em seus dados compromete diretamente a detecção de anomalias associadas à COVID-19.

5. Conclusão

Este trabalho avaliou o impacto da ausência no algoritmo *NightSignal* frente a três mecanismos de ausência de dados — MCAR, MAR e MNAR — simulados sobre séries temporais de frequência cardíaca noturna de dispositivos Fitbit, em taxas de ausência variando de 5% a 25%. Os resultados revelam um comportamento consistente do algoritmo: a Precisão se mantém alta em praticamente todos os cenários avaliados, indicando que, quando o *NightSignal* emite um alerta, ele tende a estar correto. Contudo, essa estabilidade não se repete na Revocação, que apresenta quedas expressivas, principalmente pelo MNAR. Isso se deve ao fato de que esse mecanismo remove preferencialmente os picos de frequência cardíaca — exatamente o sinal que o algoritmo monitora para detectar infecções. Esse padrão revela um algoritmo conservador — que raramente gera falsos positivos, mas que falha em disparar alertas reais quando os dados de maior relevância clínica estão ausentes.

Este estudo apresenta limitações que abrem espaço para investigações futuras. Os experimentos foram conduzidos exclusivamente com dados de Fitbit no contexto de COVID-19, e o comportamento observado pode variar para outros dispositivos e condições clínicas. Como próximos passos, pretende-se avaliar o impacto de estratégias de imputação aplicadas antes da execução do *NightSignal*, bem como estender a análise a outros algoritmos de detecção baseados em dispositivos vestíveis.

Referências

- Alavi, A., Bogu, G. K., Wang, M., Rangan, E. S., Brooks, A. W., et al. (2022). Real-time alerting system for covid-19 and other stress events using wearable data. *Nature Medicine*, 28(1):175–184.
- Benhamza, K., Benselim, R., Naidja, H., and Seridi, H. (2025). A comprehensive survey of imputation methods in medical missing data analysis. *Applied Intelligence*, 55(11):778.
- Canali, S., Schiaffonati, V., and Aliverti, A. (2022). Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness. *PLOS Digital Health*, 1(10):e0000104.
- Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B., and Tabona, O. (2021). A survey on missing data in machine learning. *Journal of Big data*, 8:1–37.
- Getzen, E., Ungar, L., Mowery, D., Jiang, X., and Long, Q. (2023). Mining for equitable health: Assessing the impact of missing data in electronic health records. *Journal of biomedical informatics*, 139:104269.
- Lima, A. M. S. (2025). Assessing the impact of missing value mechanisms on anomaly detection in healthcare wearable data. In *Anais do XL Simpósio Brasileiro de Bancos de Dados (SBBD)*, pages 781–787, Fortaleza, CE, Brasil. Sociedade Brasileira de Computação.
- Lima, A. M. S. (2026). The impact of missing values on anomaly detection tasks in healthcare wearable data. *Journal of Information and Data Management*. Accepted for publication; to appear.
- Lima, A. S. and Sousa, E. (2024). Handling missing values in data streams: An overview. In *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados*, pages 750–756, Florianópolis, SC, Brasil. Sociedade Brasileira de Computação.
- Mangussi, A. D., Santos, M. S., Lopes, F. L., Pereira, R. C., Lorena, A. C., and Abreu, P. H. (2025). mdatagen: A python library for the artificial generation of missing data. *Neurocomputing*, 625:129478.
- Mishra, T., Wang, M., Metwally, A. A., Bogu, G. K., Brooks, A. W., Bahmani, A., Alavi, A., Celli, A., Higgs, E., Dagan-Rosenfeld, O., et al. (2020). Pre-symptomatic detection of covid-19 from smartwatch data. *Nature biomedical engineering*, 4(12):1208–1220.
- Santos, M. S., Pereira, R. C., Costa, A. F., Soares, J. P., Santos, J., and Abreu, P. H. (2019). Generating synthetic missing data: A review by missing mechanism. *IEEE Access*, 7:11651–11667.

Em relação ao uso de IA generativa. As ferramentas de IA generativa foram utilizadas exclusivamente para a edição do texto em português (gramática e estilo). Todo o conteúdo foi revisado, corrigido e aprovado pelos autores, que assumem total responsabilidade pelo manuscrito.