

Inteligência artificial e saúde materna: a experiência da Caren em Goiás

Artificial intelligence and maternal health: the Caren experience in Goiás

Inteligencia artificial y salud materna: la experiencia de Caren en Goiás

Matheus Saraiva Alcino¹, Pedro Manuel Rodrigues¹, Wanderson da Silva Marques¹, Carlos Augusto Gonçalves Tibiriça¹, Willian Farias Carvalho Oliveira¹, Diogo Antônio Leal¹

1 Cientista de dados, Gerência de Inovação, Secretaria de Estado de Saúde, Goiânia (GO), Brasil.

Autor correspondente: Matheus Saraiva Alcino
E-mail: matheus.saraiva@goias.gov.br

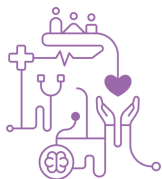
Resumo

Objetivo: A aplicação Caren tem como objetivo aprimorar a gestão dos cuidados neonatais em hospitais públicos de Goiás, utilizando inteligência artificial para prever níveis de atenção médica a recém-nascidos. **Método:** a aplicação emprega um modelo preditivo supervisionado, treinado com dados do Sistema Único de Saúde, com técnicas de integração e subamostragem para lidar com desbalanceamento. **Resultados:** os resultados mostram que o modelo escolhido, ao priorizar a revocação, demonstrou eficácia, destacando uma abordagem conservadora. A análise temporal indica a necessidade de precaução nas predições após o primeiro dia de vida. **Conclusão:** A Caren é uma ferramenta inovadora para a gestão eficiente dos recursos neonatais, sinalizando avanços na saúde materna.

Descritores: Caren; Inteligência Artificial em Saúde Materna; Otimização de cuidados neonatais.

Abstract

Objective: The Caren application aims to enhance neonatal care management in public hospitals in Goiás, using artificial intelligence to predict levels of medical attention for newborns. **Method:** The application employs a supervised predictive model, trained with



data from the Unified Health System, using integration and undersampling techniques to deal with imbalance. Results: The results show that the chosen model, prioritizing recall, demonstrated effectiveness, highlighting a conservative approach. Temporal analysis indicates the need for caution in predictions after the first day of life. Conclusion: Caren is an innovative tool for efficient management of neonatal resources, signaling advances in maternal health.

Keywords: Caren; Artificial Intelligence in Maternal Health; Optimization of Neonatal Care.

Resumen

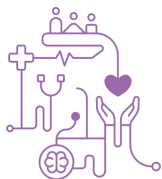
Objetivo: La aplicación Caren tiene como objetivo mejorar la gestión de los cuidados neonatales en hospitales públicos de Goiás, utilizando inteligencia artificial para prever niveles de atención médica a recién nacidos. Método: la aplicación emplea un modelo predictivo supervisado, entrenado con datos del Sistema Único de Salud, utilizando técnicas de integración y submuestreo para abordar el desbalance. Resultados: Los resultados muestran que el modelo elegido, al priorizar la recuperación, demostró eficacia, destacando un enfoque conservador. El análisis temporal indica la necesidad de precaución en las predicciones después del primer día de vida. Conclusión: Caren es una herramienta innovadora para la gestión eficiente de recursos neonatales, señalando avances en la salud materna.

Descriptor: Caren; Inteligencia Artificial en Salud Materna; Optimización de Cuidados Neonatales.

Introdução

Um recém-nascido tem seu estado clínico avaliado logo nos primeiros minutos de vida por meio de uma série de variáveis que podem ser determinantes para a tomada de decisão médica em quaisquer intervenções ou encaminhamentos a unidades de tratamento intensivo. Alguns destes métodos são, por exemplo, o índice Apgar⁽³⁾ e o boletim Silverman-Andersan - BSA⁽⁸⁾.

A avaliação do estado clínico de um recém-nascido através destes métodos pode ser crucial para a sua própria sobrevivência e, pode ser utilizada, como critério principal para determinar a priorização do atendimento em maternidades onde número de



neonatos é incompatível com a quantidade de médicos, leitos e outros recursos disponíveis.

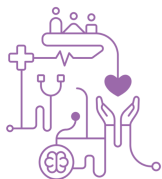
Neste contexto, considera-se como potencialmente útil o emprego de ferramentas do campo da inteligência artificial (IA), em especial os modelos supervisionados, para encontrar padrões não perceptíveis por humanos na tarefa de detectar anomalias e prevenir situações de risco. Portanto, este trabalho tem o objetivo de apresentar uma ferramenta baseada em inteligência artificial denominada Caren, que visa determinar de maneira automática e não viesada os níveis de atenção à recém-nascidos de unidades de saúde pública do estado de Goiás. É evidente que a proposta de uso deste tipo de ferramenta busca auxiliar na administração eficiente de recursos escassos, sobretudo a atenção de especialistas, de modo a otimizar tal gestão e garantir a segurança de um maior número de neonatos.

Método

A criação de um modelo preditivo supervisionado requer necessariamente uma base de dados de treinamento rotulada e tal rotulação é um desafio no caso da saúde pública brasileira. O modelo que a aplicação Caren utiliza é por natureza definido como um modelo de classificação binária que tem o objetivo de estimar as probabilidades de óbito de nascidos vivos e utilizá-las como níveis de atenção necessários a cada neonato.

O Departamento de Informática do Sistema Único de Saúde⁽¹⁾ (DATASUS) é responsável por manter as duas principais bases de dados necessárias para o treinamento do modelo Caren. A base de dados SINASC (Sistema de Informações sobre Nascidos Vivos) tem como principal finalidade registrar as Declarações de Nascido Vivo e suas variáveis. Já base de dados SIM (Sistema de Informações sobre Mortalidade) coleta, organiza e disponibiliza informações sobre as Declarações de Óbitos.

Para a rotulação dos dados no contexto do modelo Caren, os dados SIM foram subselecionados de modo que os registros de óbitos se restrinjam àqueles cuja diferença de temporal entre a data do óbito e a data de nascimento seja inferior à 6 dias. Esta é uma característica importante do modelo Caren. Este é um modelo binário treinado para



realizar predições de óbitos neonatais precoce, ou seja, óbitos ocorridos em até 6 dias após o nascimento.

A modelagem almejada pela Caren envolve dados naturalmente desbalanceados, exigindo cuidado na otimização da utilização dos dados de óbitos disponíveis na base SIM durante a integração com a base SINASC. A realização efetiva dessa tarefa, portanto, demandou a aplicação de técnicas avançadas de integração, conhecidas como técnicas de *linkage*⁽¹⁾.

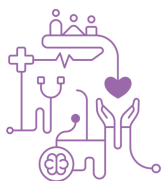
Em conjuntos de dados reais, é natural observar desequilíbrio entre as proporções das classes binárias. Contudo, apenas são considerados como desbalanceados pela literatura aqueles conjuntos que exibem um desequilíbrio significativo, como por exemplo, na ordem de 100 para 1, sendo ainda mais extremo, como no caso de detecção de fraudes financeiras, podendo atingir proporções de 100.000 para 1⁽²⁾.

Há duas abordagens principais para lidar com esse problema na literatura⁽²⁾. A primeira consiste no uso de técnicas de sobreamostragem, aplicadas aos dados da classe minoritária para aumentar sua representatividade. A segunda envolve técnicas de subamostragem aplicadas aos dados da classe majoritária, reduzindo sua prevalência. Ambas as abordagens têm como objetivo equilibrar os dados, proporcionando uma modelagem mais eficiente, independentemente do modelo utilizado.

Durante o treinamento do modelo Caren, ambas técnicas foram testadas experimentalmente e, de acordo com as métricas de desempenho escolhidas, a subamostragem da classe majoritária se mostrou mais apropriada no contexto da modelagem. Na Tabela 1 são mostradas as proporções dos dados de óbito e sobrevivência juntamente com a subamostragem utilizada.

Tabela 1: Estratégia de subamostragem de dados para treinamento do modelo Caren.

| Fonte de dados | Não óbito | Óbito | Proporção de óbitos |
|------------------------------|-----------|-------|---------------------|
| Dados originais após linkage | 1.130.439 | 6.483 | 0,0057 |
| Dados para subamostragem | 791.307 | 4.538 | 0,0057 |



| | | | |
|----------------------|---------|-------|--------|
| Dados de teste (30%) | 339.132 | 1.945 | 0,0057 |
| Dados de treino | 13.614 | 4.538 | 0,25 |

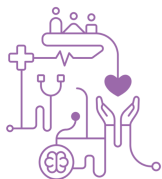
Como evidenciado na Tabela 1, os dados de treino do modelo representam uma amostra estratificada que inclui 4.538 casos de óbito (70% do total de óbitos) e aleatoriamente selecionou 13.614 casos de sobrevivência. Essa abordagem visa atenuar o desbalanceamento dos dados de treino. No entanto, é importante observar que o desbalanceamento original foi mantido para avaliar o poder preditivo do modelo.

Em relação às etapas de engenharia de recursos para tratamento e criação de novos preditores, foram empregadas algumas técnicas comuns de modelagem supervisionada, como preenchimento de dados faltantes pela média de subgrupos específicos e codificação de variáveis exógenas de natureza categórica. Uma vez que a codificação de variáveis categóricas produz, no caso dos dados de treinamento do modelo Caren uma matriz esparsa, o último processamento que compõe o encadeamento das etapas de modelagem se refere a remoção de preditores com baixa variabilidade, mantendo apenas aquelas com uma variância acima do limite definido (0,01).

Vários modelos de classificação foram empiricamente testados, incluindo Regressão Logística, Naive Bayes, Árvores Aleatórias e Gradient Boosting. Todos os modelos passaram por etapas semelhantes de modelagem e hiperparametrização, com avaliações de desempenho adequadas para dados desbalanceados. A análise da área sob a curva PR (precisão x revocação) e do F-score foi predominante na comparação e seleção do melhor modelo.

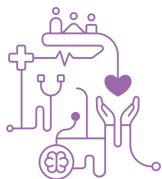
Entre todas as comparações, o modelo Histogram-based Gradient Boosting Classification Tree (HGBCT) se destacou, demonstrando superioridade nas métricas definidas. O modelo HGBCT mostra alta precisão em conjuntos de dados extensos, robustez em cenários de desequilíbrio nos dados e eficiência computacional tanto no ajuste quanto na realização de previsões⁽⁴⁾. O conjunto de preditores selecionados para predição de óbito é mostrado no Quadro 1.

Quadro 1: Preditores do modelo Caren.



| Preditor | Significado | Natureza |
|-------------------------------|---|------------|
| Apgar | Índice Apgar no quinto minuto de vida. O índice Apgar é um método de avaliação clínica de um recém-nascido feito pela equipe médica no momento de seu nascimento. | Discreta |
| Jovialidade da mãe | “baixa” se a idade da mãe é inferior à 30 anos; “média” se a idade da mãe está entre 31 e 39 anos; “alta” se idade da mãe está entre 40 e 50 anos; “muito alta” se a idade da mãe é superior a 50 anos. | Catagórica |
| Sexo do bebê | Masculino/Feminino. | Catagórica |
| Tipo de parto | Cesário/Vaginal. | Catagórica |
| Estado civil da mãe | Solteira/Casada/Viúva/Divorciada/ignorado. | Catagórica |
| Peso do recém-nascido | Peso em gramas registrado no nascimento. | Contínua |
| Total de filhos mortos da mãe | Total de abortos ou filhos com morte neonatal. | Discreta |
| Total de partos | Total de partos realizados pela mãe. | Discreta |
| Idade gestacional | Total de semanas desde o início da gestação até o parto. | Discreta |
| Escolaridade da mãe | Grau de escolaridade da mãe. | Catagórica |
| Consultas pré-natal | Total de consultas pré-natal realizadas pela mãe durante o período gestacional. | Discreta |
| Tipo de gravidez | Única/Dupla/Tripla ou mais. | Catagórica |
| Malformação | Existência de algum tipo de malformação no recém-nascido. | Catagórica |
| Hospital de nascimento | Unidade de saúde em que foi realizado o parto. | Catagórica |
| Classe Lubchenco | A classe Lubchenco é uma ferramenta utilizada na neonatologia para avaliar o crescimento fetal, comparando o peso e a idade gestacional do recém-nascido com padrões de referência estabelecidos para determinar se o bebê está adequado, pequeno ou grande para a idade gestacional ⁽⁶⁾ . | Catagórica |

Todas as etapas de modelagem, linkage, análises exploratórias, pré processamento de dados e ajuste de modelos foram realizados através da linguagem de programação



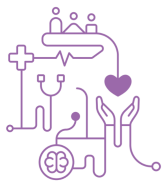
python e a construção do encadeamento das etapas de modelagem juntamente com o ajuste do modelo HGBCT foi realizado por meio da biblioteca scikit-learn⁽⁹⁾.

Resultados e discussão

A abordagem adotada pelo projeto Caren reflete a natureza crítica do problema de predição do estado de saúde de recém-nascidos. Considerando a importância vital de prever com precisão o risco de óbito neonatal, optou-se por um viés conservador durante o treinamento do modelo. Essa escolha foi motivada pela necessidade de minimizar o risco de falsos negativos, ou seja, casos em que o modelo não prevê o óbito de um recém-nascido que, na verdade, está em risco.

Nesse contexto, a priorização de falsos positivos em detrimento de falsos negativos é uma decisão deliberada e fundamentada. Embora os falsos positivos possam resultar em recomendações de atenção médica adicional para recém-nascidos que não necessitam de cuidados intensivos, as consequências desses erros são menos graves comparadas aos falsos negativos. Estes últimos podem levar a uma subestimação da gravidade da condição de um recém-nascido em risco, potencialmente resultando em intervenções tardias ou insuficientes.

Portanto, o modelo foi projetado para preferir erros por excesso de cautela, garantindo que todos os recém-nascidos que possam estar em risco recebam atenção médica adequada. Essa abordagem conservadora almeja salvar vidas, assegurando que nenhum recém-nascido com necessidade de cuidados intensivos seja negligenciado devido a uma previsão incorreta. Em suma, é preferível que o modelo sugira uma atenção médica adicional desnecessária do que correr o risco de não indicar cuidados vitais para aqueles que realmente precisam. No âmbito da modelagem, a ênfase recai na importância prática de um modelo com maior revocação em detrimento da precisão, no contexto específico do projeto Caren. Essa escolha estratégica é motivada pelo desequilíbrio acentuado nos dados, conforme evidenciado na Tabela 1. Diante desse desbalanceamento, a métrica selecionada para guiar a otimização dos parâmetros do modelo durante a fase de hiperparametrização foi a F-Score, conforme definida pela



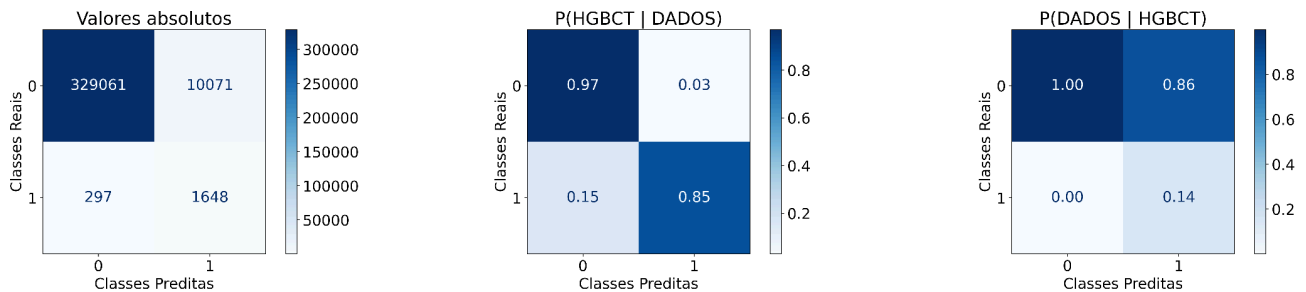
equação (1). Esta abordagem busca alcançar um equilíbrio sensato entre precisão e revocação, levando em consideração a natureza sensível do problema em questão.

$$F(\beta) = (1 + \beta^2) \cdot \frac{\text{precisão} \cdot \text{revocação}}{(\beta^2 \cdot \text{precisão}) + \text{revocação}} \quad (1)$$

em que $\beta \in [0, \infty)$. Matematicamente, a métrica F-score pode ser entendida como uma forma de cálculo de média geométrica entre duas grandezas, no caso precisão e revocação. Quanto maior for o valor de β , maior é a importância atribuída à revocação em detrimento da precisão. Desta forma, a busca exaustiva realizada durante a hiperparametrização foi feita com objetivo de maximizar a métrica F-score com $\beta = 2$.

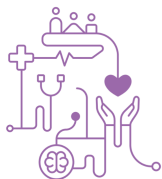
Conforme mencionado, o modelo hiperparametrizado teve o seu desempenho avaliado em um conjunto de dados de teste que manteve o desbalanceamento original dos dados. As matrizes de confusão com valores absolutos e normalizados são mostradas na Figura 1.

Figura 1: Matrizes de confusão do modelo.



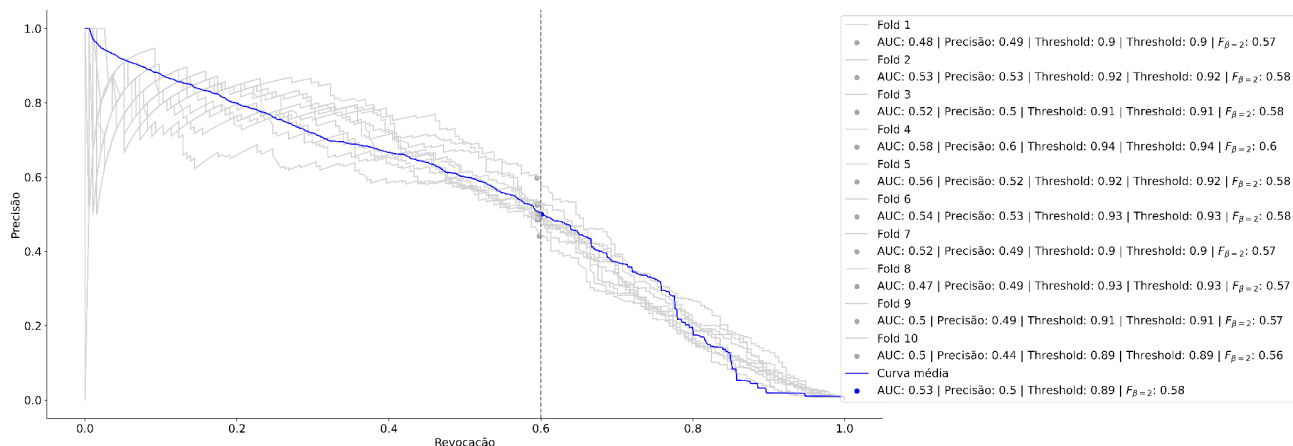
A matriz de confusão com valores absolutos mostrada na Figura 1 ilustra objetivamente o conservadorismo desejado do modelo. Observa-se que há 10.071 instâncias de sobrevivência erroneamente classificadas como óbito pelo modelo, caracterizando falsos positivos. No contexto do projeto Caren, esse tipo de erro é considerado menos crítico, uma vez que implica na atribuição indevida de atenção a recém-nascidos teoricamente em estado clínico adequado.

Este equívoco, denominado estatisticamente como erro tipo II, assume uma gravidade mais acentuada em cenários onde os recursos em unidades de saúde são escassos. Por outro lado, há 297 casos de falsos negativos. Este erro, denominado como erro tipo I, representa no contexto Caren os casos mais graves.



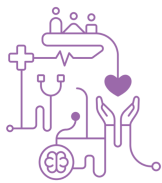
Vale lembrar que a geração de cada classe produzida pelo modelo segue a regra definida: 1 se a probabilidade estimada for maior ou igual a 0,5; 0 caso contrário. Entretanto, o produto a ser utilizado pelo modelo Caren são as próprias probabilidades de óbito estimadas, que na prática, serão interpretadas como níveis de atenção requeridos pelos especialistas. Sendo assim, a área sob a curva PR (precisão x revocação) é também uma métrica complementar que auxilia no processo de decisão médica. No Gráfico 1, são mostradas as curvas PR interpoladas durante a validação cruzada conduzida com os dados de teste.

Gráfico 1: Curvas PR geradas durante a validação cruzada conduzida nos dados de teste.



A curva média (destacada em azul) exibida no Gráfico 1 possui uma área sob a curva PR de 53%. Conforme ilustrado, para um nível de revocação de 60%, a precisão estimada para o modelo é de aproximadamente 50%. Se houvesse interesse na geração de classes pelo modelo, em vez de suas probabilidades, a função para tal geração deveria ser definida por: 1 se a probabilidade estimada for maior ou igual a 0,89; 0 caso contrário.

Naturalmente, ainda que o modelo fosse perfeito do ponto de vista preditivo, a sua utilização demandaria precauções específicas, sendo a principal delas a validade temporal das previsões produzidas. A dinâmica do estado clínico de um recém-nascido pode ser inconstante, uma vez que é influenciada por uma série de fatores clínicos que se estendem desde a gestação até o pós-parto. A complexidade desse cenário ressalta a necessidade de cautela na interpretação e aplicação das previsões do modelo ao longo do tempo. O debate em torno das práticas mais apropriadas para o monitoramento de



recém-nascidos, como evidenciado em alguns estudos^(5,7), destaca a importância de considerar abordagens robustas e dinâmicas na gestão da saúde neonatal.

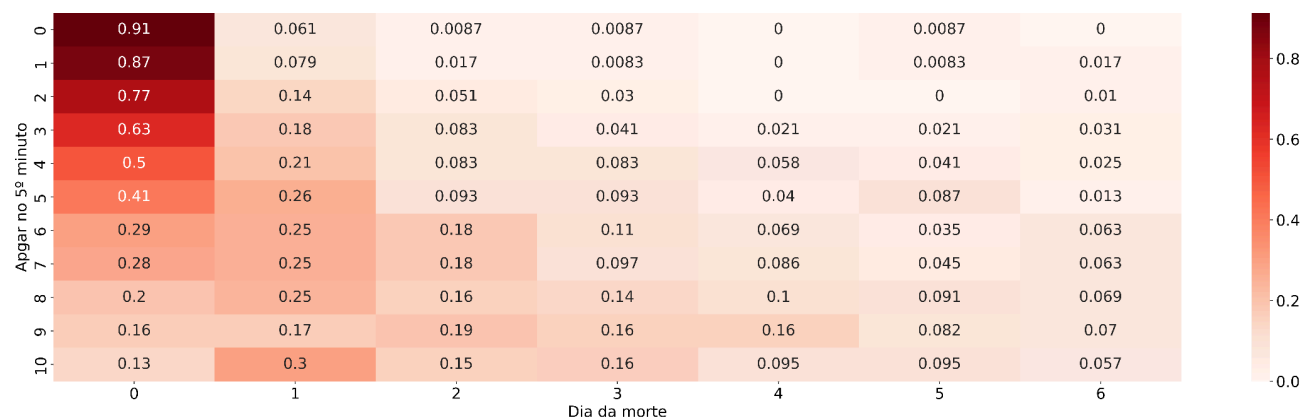
Neste sentido, a validade das previsões do modelo Caren foi avaliada por meio da relação probabilística entre o i -ésimo dia da morte de um recém-nascido e o índice Apgar no quinto minuto em um subconjunto de dados que contém apenas óbitos neonatais precoce. A escolha do índice Apgar para realizar tal análise se deve ao fato de que esta é uma das principais componentes do modelo Caren, além do fato de que este talvez seja o preditor mais confiável do ponto de vista prático, uma vez que é gerado pelos próprios especialistas ao analisar o neonato.

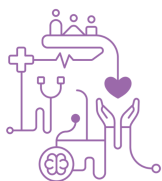
A distribuição de interesse neste contexto é denotada matematicamente pela distribuição de probabilidade condicional da equação (2).

$$P(D | A) = \frac{P(A, D)}{P(A)} \quad (2)$$

em que D representa a morte de um nascido vivo no i -ésimo dia após o seu nascimento e, A é o índice Apgar atribuído ao recém-nascido pelo médico no quinto minuto de vida. Por conveniência de notação, a distribuição a posteriori $P(D | A)$ é referenciada neste estudo como ${}_d q_0$, que é uma notação demográfica/atuarial comumente utilizada para tratar distribuições desta natureza. As distribuições de ${}_d q_0$ podem ser visualizadas através do mapa de calor da Figura 2.

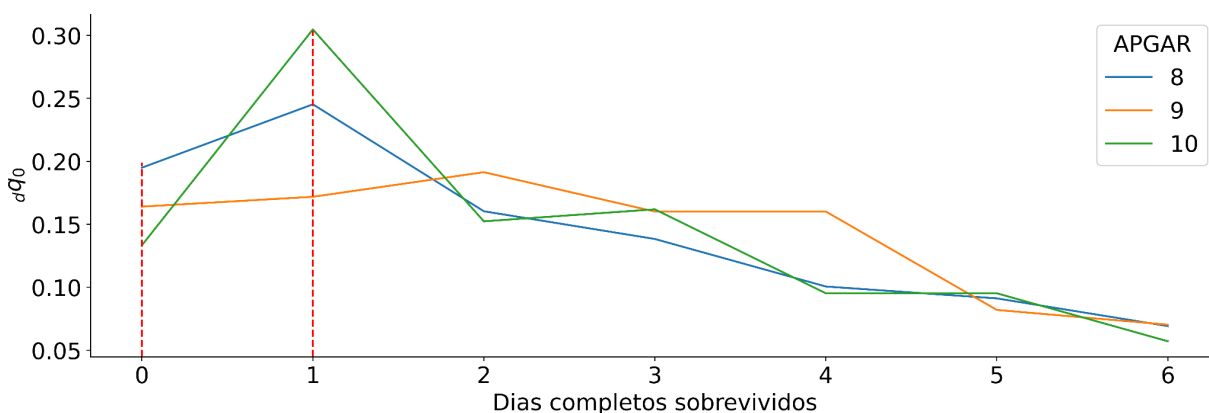
Figura 2: Distribuição a posteriori ${}_d q_0$.





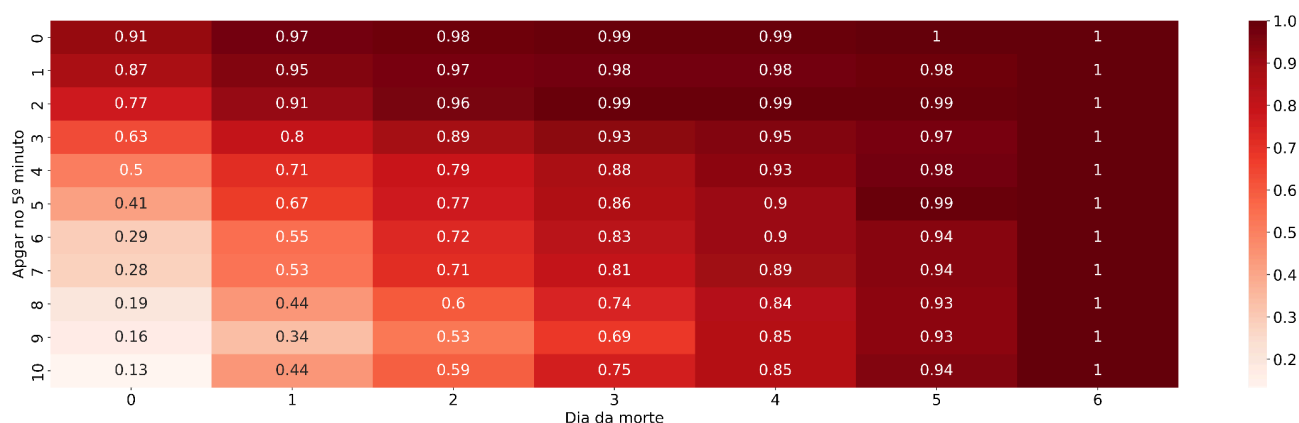
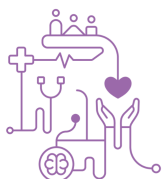
Nota-se que o primeiro dia de vida e os três subsequentes são decisivos para a sobrevivência de um neonato, obviamente ${}_d q_0$ é mais intensa nos nascidos vivos com menores valores de Apgar. É interessante observar que para os recém-nascidos com óbito em até 6 dias e com as três maiores atribuições de valores de Apgar (8, 9 e 10) a sobrevivência ao 1 dia de vida é mais desafiadora do que o dia do nascimento (dia 0), conforme mostrado no Gráfico 2.

Gráfico 2: Distribuição a posteriori ${}_d q_0$ para indivíduos com índice Apgar 8, 9 e 10.



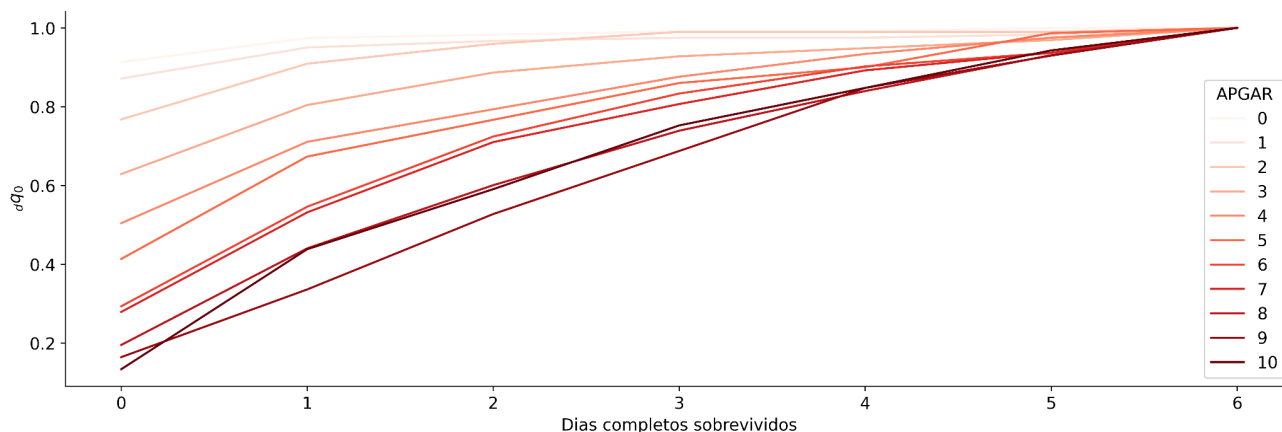
Considerando que os dados desta análise considera apenas os nascidos vivos que vieram a óbito dentro de 6 dias após seu nascimento, a análise da distribuição acumulada da posteriori encontrada implica em uma interpretação análoga à tábua de vida⁽¹⁰⁾, em que uma de suas premissas é a estimação das probabilidades de morte à exata idade de uma população fechada. A ideia de estudar a distribuição acumulada da posteriori de interesse é fazer com que cada linha do mapa de calor (índice Apgar) seja interpretada como uma população fechada e mutuamente exclusiva que inevitavelmente será extinta dentro de 6 dias. Assim, espera-se que a velocidade de tal extinção seja diretamente relacionada com o índice Apgar atribuído pelo médico. Na Figura 3, é mostrada tal distribuição.

Figura 3: Distribuição $P(d < D|A)$.

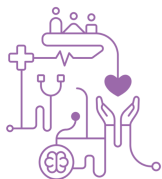


Como pode ser observado, cada população (nota de apagar) possui uma relação gradual entre si para o falecimento total dos nascidos vivos, de modo que a população de nascidos vivos com maiores índices de Apgar é aquela que leva mais tempo para atingir 100% de mortalidade. Este resultado demonstra, inclusive, o quão importante é o índice Apgar como variável preditora do modelo Caren. A velocidade de extinção de cada uma destes subgrupos de população pode ser observada através do Gráfico 3.

Gráfico 3: Velocidades de extinção de subgrupos de neonatos com diferentes Apgar: $P(d < D|A)$.



A taxa de variação das curvas das populações com maiores índices Apgar é mais agressiva e portanto, este é um indicativo de que as probabilidades estimadas pelo modelo Caren não devem ser consideradas por um período longo de tempo. Devido às distribuições e características encontradas, recomenda-se que os níveis de atenção estimados pelo modelo tenham validade apenas para o primeiro dia de vida, principalmente para nascidos vivos com maiores índices Apgar.



Conclusão

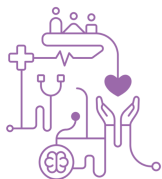
A aplicação Caren, desenvolvida pela Gerência de Inovação da Secretaria de Saúde de Goiás, representa uma ferramenta inovadora que integra tecnologia e gestão para otimizar os cuidados neonatais em unidades de saúde pública. Durante o desenvolvimento do modelo preditivo supervisionado, foram enfrentados desafios específicos relacionados à rotulação dos dados, integração de bases e desbalanceamento das classes. Após testes experimentais, adotaram-se técnicas de subamostragem da classe majoritária para equilibrar o conjunto de treinamento. O modelo escolhido, Histogram-based Gradient Boosting Classification Tree (HGBCT), demonstrou eficácia diante do desbalanceamento acentuado dos dados, avaliado por métricas como F-Score e curvas PR.

Ao priorizar a revocação em detrimento da precisão, buscou-se minimizar os erros mais críticos, como os falsos negativos, conforme revelado pelas matrizes de confusão. A análise da validade temporal das predições enfatizou a importância do índice Apgar como variável preditora, especialmente para o primeiro dia de vida. Essa abordagem ressalta a necessidade de interpretação cautelosa das probabilidades estimadas pelo modelo, direcionando os níveis de atenção principalmente para o período neonatal inicial.

Assim, a aplicação Caren não apenas representa um avanço na gestão eficiente dos cuidados neonatais, mas também destaca a importância de abordagens cautelosas e dinâmicas na interpretação e aplicação de modelos preditivos em saúde materna. A harmonia entre a democratização de recursos tecnológicos e a implementação de ferramentas de IA evidenciadas neste projeto sinaliza um caminho promissor para melhorar a qualidade dos cuidados neonatais e otimizar a administração de recursos públicos na área de saúde.

Referências

1. Brasil. Ministério da Saúde. DATASUS. Tabnet. Brasília, DF: Ministério da Saúde; 2022. Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>. Acesso em: janeiro de 2022.



2. Chawla, N, Bowyer, K, Hall, L, Kegelmeyer, W. "SMOTE: synthetic minority over-sampling technique". *Journal of artificial intelligence research* 2002; 16:321–357.
3. Cnattingius, S, Johansson, S, Razaz, N. "Apgar score and risk of neonatal death among preterm infants". *New England Journal of Medicine* 2020; 383(1):49–57.
4. Gudmann, A, Mucsi, L. "Pixel and object-based land cover mapping and change detection from 1986 to 2020 for Hungary using histogram-based gradient boosting classification tree classifier". *Geographica Pannonica* 2022; 26(3).
5. Haran, C, Van Driel, M, Mitchell, B, Brodribb, W. "Clinical guidelines for postpartum women and infants in primary care—a systematic review". *BMC pregnancy and childbirth* 2014; 14:1–9..
6. Lubchenco, L, Hansman, C, Dressler, M, Boyd, E. "Intrauterine growth as estimated from liveborn birth-weight data at 24 to 42 weeks of gestation". *Pediatrics* 1963; 32(5):793–800.
7. Moller, AB, Newby, H, Hanson, C, Morgan, A, El Arifeen, S, Chou, D, Diaz, T, Say, L, Askew, I, Moran, A. "Measures matter: a scoping review of maternal and newborn indicators". *PloS one* 2018; 13(10)..
8. Nussbaum, C, Lengauer, M, Puchwein-Schwepcke, A, Weiss, V, Spielberger, B, Genzel-Boroviczeny, O. "Noninvasive Ventilation in Preterm Infants: Factors Influencing Weaning Decisions and the Role of the Silverman-Andersen Score". *Children* 2022; 9(9):1292.
9. Pedregosa, F, Varoquaux, G, Gramfort, A, Michel, V, Thirion, B, Grisel, O, Blondel, M, Prettenhofer, P, Weiss, R, Dubourg, V, Vanderplas, J, Passos, A, Cournapeau, D, Brucher, M, Perrot, M, Duchesnay, E. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 2011; 12:2825–2830.
10. Preston, S, Heuveline, P, Guillot, M. *Demography: Measuring and Modeling Population Processes*. Wiley; 2000.
11. Winkler, W. "Matching and record linkage". *Wiley interdisciplinary reviews: Computational statistics* 2014; 6(5):313–325.