

Estimativa da Ocupação de Leitos para Tratamento da COVID-19 Usando Dados Temporais

Thiago M. Ventura, Raphael S. R. Gomes, Gabriel S. G. Pedroso,
Daniel A. Vecchiato, Rebeca L. Rezende

¹ Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Caixa Postal 78060-900 – Cuiabá – MT – Brasil

{thiago, raphael, daniel}@ic.ufmt.br
{gabriel.de.s.g.pedroso, beca.rezendel}@gmail.com

Abstract. *Efficient management of the quantity and types of hospital beds is a critical challenge in public healthcare, as highlighted during the COVID-19 pandemic. This study proposes an approach based on artificial neural networks of the Multi-Layer Perceptron (MLP) type and time series analysis, using data collected in the state of Mato Grosso between 2020 and 2022. The model was trained with hospitalization data from the previous two weeks to predict the demand for clinical and complementary beds, achieving a root mean square error (RMSE) of 45.83 for a two-week forecast horizon. The results demonstrate its applicability as a decision-support tool for healthcare managers, contributing to a more efficient allocation of hospital resources.*

Resumo. *O gerenciamento eficiente da quantidade e dos tipos de leitos hospitalares é um desafio crítico na saúde pública, evidenciado na pandemia de COVID-19. Este estudo propõe uma abordagem baseada em redes neurais artificiais do tipo Multi Layer Perceptron (MLP) e análise de séries temporais, utilizando dados coletados no estado de Mato Grosso entre 2020 e 2022. O modelo foi treinado com dados de internações das duas semanas anteriores para prever a demanda de leitos clínicos e complementares, alcançando um erro quadrático médio de 45,83 para um horizonte de duas semanas. Os resultados demonstram sua aplicabilidade como ferramenta de apoio à decisão para gestores, contribuindo para uma alocação mais eficiente de recursos hospitalares.*

1. Introdução

A recente pandemia de COVID-19 causou sérios problemas de ocupação nos leitos hospitalares. Em diversos locais, a superlotação dos leitos comprometeu o atendimento de novos pacientes, gerando graves consequências [Marcilio et al. 2022]. A expansão da capacidade de leitos para tratar pacientes da pandemia foi essencial para garantir um atendimento adequado da população. Essa expansão foi alcançada por meio de reservas de leitos especificamente para os pacientes com COVID-19 e da disponibilização de novos leitos.

Entretanto, reservar leitos a mais do que o necessário pode prejudicar pacientes com outras enfermidades. Ao mesmo tempo, a criação de muitos leitos pode resultar em desperdício de recursos que poderiam ser direcionados para outros tipos de gastos na saúde beneficiando a população. Esse dilema acontece tanto com o crescimento da

pandemia quanto com seus momentos de diminuição de casos. Portanto, é fundamental desenvolver um método preciso para estimar a necessidade de leitos frente a uma doença que afeta uma grande parte da população.

Alguns trabalhos já foram elaborados nesse contexto. Em [Runge et al. 2022] foi modelado dados de transmissão do vírus para a cidade de Chicago, nos Estados Unidos da América, de março a agosto de 2020, para possíveis ocupações para setembro de 2020 a maio de 2021 trazendo diferentes variáveis para delimitar um limiar de Unidades de Terapia Intensivas (UTI), como taxa de transmissão ou mitigação. O trabalho de [Ritter et al. 2021] foi utilizado um modelo estatístico simples, no entanto, aplicaram parâmetros específicos a uma região para prever a necessidade futura de UTI, como taxa de UTI, atraso entre infecção e admissão na UTI, além da duração da estadia da UTI, com aplicação a dados de 3 cidades muito afetadas da Europa: Madri, na Espanha, Berlim na Alemanha e Lombardia na Itália.

Em [Almeida et al. 2021] foi utilizado um modelo comportamental da mesma forma que [Runge et al. 2022], com diferença na escolha de variáveis para a modelagem, e utilizaram dados abertos do Ministério da Saúde do Brasil. Já [Bekker et al. 2023] foi modelado a ocupação hospitalar com teoria das filas trazendo o tempo de permanência dos pacientes como variável com a predição futura da ocupação baseada em pacientes hospitalizados e pacientes que serão admitidos no hospital.

Ampliando as abordagens anteriores, [Gomes et al. 2024] propuseram um modelo matemático para a alocação eficiente de leitos hospitalares, levando em conta diferentes tipos de leitos e uma taxa de ocupação ideal. O estudo aplicou técnicas de otimização com o objetivo de maximizar a taxa de sobrevivência dos pacientes, validando o modelo por meio de dados reais coletados no Hospital Universitário da Universidade Federal do Maranhão. Essa abordagem se diferencia das anteriores por buscar uma alocação mais precisa dos recursos hospitalares.

Em [Consoli et al. 2024] foi investigado a previsão de admissões hospitalares em hospitais brasileiros utilizando técnicas de aprendizado de máquina, baseando-se em dados do BRATECA, um dataset de saúde da atenção terciária. O objetivo do estudo foi prever se um paciente seria internado logo após a entrada no hospital, atingindo uma acurácia de 88% na classificação de internações. Essa abordagem inovadora oferece uma previsão em tempo real da necessidade imediata de leitos.

Diferente dessas abordagens, nosso estudo emprega redes neurais artificiais e análise de séries temporais para prever a demanda de leitos hospitalares ao longo do tempo. Essa metodologia permite um planejamento estratégico e proativo da ocupação hospitalar, promovendo uma otimização mais eficiente dos recursos de saúde e adaptando-se dinamicamente as variações na demanda.

Neste trabalho, foi desenvolvida uma nova metodologia para estimar a necessidade de leitos em uma determinada região. O foco está na previsão da quantidade de leitos necessária para as semanas subsequentes, utilizando os dados de internações das duas semanas anteriores. Para isso, foram utilizados dados sobre a utilização de leitos no estado de Mato Grosso, no Brasil, durante o período da pandemia causada pela COVID-19. Este artigo está estruturado da seguinte forma: a Seção 2 apresenta a metodologia usada para conduzir nossos experimentos, a Seção 3 discute os resultados dos testes, a

Seção 4 as respectivas discussões sobre os resultados e, por fim, a Seção 5 conclui este artigo com potenciais trabalhos futuros.

2. Materiais e Métodos

2.1. Dados

Os dados utilizados neste trabalho contêm informações sobre a admissão e alta dos pacientes. Há informações sobre o dia que o paciente começou a ter apoio de um leito, para cada um dos hospitais de cada município do estado de Mato Grosso, Brasil, tanto públicos quanto particulares, no período de 2020 a 2022. Os leitos são categorizados em dois tipos principais: Clínico (Enfermaria e Isolamento) ou Complementar (Isolamento, Intensivo e Semi Intensivo). Além disso, os dados abrangem a evolução clínica dos pacientes (Cura/Alta, Óbito, Óbito por Outras Causas; e Transferência ou Regulação) e a região de saúde abrangida pelo município, que auxilia na integração de ações e políticas públicas de saúde.

É importante ressaltar que os dados analisados neste estudo não incluem informações pessoais identificáveis dos pacientes internados. A análise foi conduzida com dados agregados e anonimizados, assegurando que a privacidade e a confidencialidade dos pacientes fossem mantidas. A coleta e o tratamento dos dados foram realizados respeitando diretrizes de privacidade e boas práticas sugeridas na literatura, como a limitação do escopo de coleta, o controle de acesso e o uso de protocolos seguros para compartilhamento, conforme discutido por Wacksman (2021)[Wacksman 2021].

2.2. Transformação

Para a análise, foram calculadas as internações semanais para cada tipo de leito, inicialmente agregados nas categorias Clínico e Complementar. Esse processo foi aplicado a cada registro de internação na base de dados. Vale salientar ainda que caso a internação passasse de 2 semanas, era marcado como demanda por leito (separado por cada tipo de leito). A Tabela 1 possui uma amostra da saída dos processamentos descritos, em que 50% dos dados foram utilizados para teste.

Tabela 1. Saída dos processamentos temporais por tipos de leitos

	semana 1 clinico	semana 2 clinico	semana 1 complementar	semana 2 complementar	demanda clinico	demanda complementar
1	1	1	1	1	1	1
2	1	1	1	1	1	1
...
153	245	274	101	111	296	128
...

2.3. Rede Neural Artificial

Em [Braga et al. 2021] foi demonstrado resultados consistentes com Redes Neurais Artificiais ao utilizar dados do estado do Pará, Brasil, para estimativas de curto prazo do número acumulado de casos e mortes, o número diário de casos e mortes, e o número

de leitos para avaliar o impacto da doença ao longo do tempo. Assim, para este trabalho, foi escolhido o tipo de Rede Neural Artificial *Multi Layer Perceptron* (MLP), visto que o algoritmo também já foi utilizado para modelagem da propagação da COVID-19 [Car et al. 2020].

Para melhor direcionamento de espaço de busca dos pesos, foi realizada *Grid Search* sobre o número de neurônios, a função de ativação, o otimizador de parâmetros, além do valor da força do termo de regularização L2 e a taxa de aprendizagem. Essa abordagem permitiu identificar a melhor combinação de hiperparâmetros para o MLP, aprimorando a precisão das estimativas de demanda por leitos. Foram exploradas diferentes configurações para o número de neurônios, testando as seguintes opções: 15, 30 e 45 neurônios em uma única camada; 30 e 15, 60 e 20, e 90 e 10 neurônios em duas camadas; e 20, 10 e 5 neurônios em três camadas. A grid search foi escolhida em vista de não poder mudar vários hiperparâmetros ao mesmo tempo e então observar o impacto de cada mudança de parâmetros citados anteriormente [Ali et al. 2023].

A busca pela função de ativação se deu entre tangente hiperbólica ou *Rectified Linear Unit* (ReLU), enquanto pelo otimizador ocorreu entre *Stochastic Gradient Descent* (SGD) e Adam. Por fim, a busca pelo termo de regularização L2 ficou entre 0,0001 e 0,05 e a taxa de aprendizagem entre permanecer constante com um valor de 0,0001 ou ser modificada ao longo do treinamento caso a função de perda não trouxesse queda significativa em uma tolerância de 0,0001 em 2 épocas seguidas quando o otimizador no *Grid Search* for o SGD.

Ao fim da busca, o melhor modelo teve função de ativação ReLU, 2 camadas com 60 e 20 neurônios, respectivamente, taxa de aprendizagem constante e otimizador Adam. Com esses parâmetros foi utilizado o modelo para a retroalimentação para estimativas futuras, com a finalização do treinamento em 500 iterações ou 10 iterações sem reduzir o valor de *loss* em 0,0001.

2.4. Retroalimentação para Estimativas Futuras

Após identificar o melhor modelo, e com base no processo definido na seção 2.2, foram feitas as previsões para as quatro semanas subsequentes para cada semana do conjunto de dados transformados. As estimativas realizadas foram reutilizadas como entrada para as estimativas posteriores, a fim realizar previsões ao longo do tempo.

2.5. Métricas para Avaliação

Para avaliação do modelo treinado foram utilizadas as métricas Coeficiente de Determinação R^2 e *Root Mean Squared Error* (RMSE). A métrica R^2 é dada pela subtração de 1 e a razão da soma quadrada residual (SS_{res}) de leitos e leitos previstos, e soma quadrada total (SS_{tot}) de leitos e leitos, conforme Equação 1. Essa métrica foi utilizada para poder determinar o quanto da variância das estimativas de leitos do modelo pode explicar a variância da demanda de leitos.

A métrica *RMSE* é dada pela raiz quadrada da média dos erros quadrados da quantidade de leitos, conforme Equação 2. Essa métrica foi utilizada para penalizar as diferenças entre as demandas e as estimativas previstas e ter uma visão do erro das previsões em relação ao acompanhamento e sugestão de leitos necessários para um

hospital.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2} \quad (2)$$

3. Resultados

Os primeiros testes foram para estimar o uso de leitos para a semana seguinte. O número de semanas anteriores usadas como entrada do modelo foi variado nesses testes. A Tabela 2 mostra os resultados das estimativas usando medidas R^2 e $RMSE$.

Também avaliamos a necessidade de leitos para as semanas subsequentes. Além da semana seguinte, conforme mencionado anteriormente, as estimativas foram feitas para duas, três e quatro semanas adiante, representando previsões de até um mês. Nesses casos, a variação no número de semanas como entrada para o modelo também foi testada. A Figura 1 apresenta os resultados das estimativas para todas essas combinações considerando o Erro Médio Absoluto (EMA).

Tabela 2. Estimativas de leitos variando de acordo com o número de semanas como entrada do modelo

Semanas de entrada	R^2	$RMSE$
1	0,94	83,36
2	0,98	45,83
3	0,97	56,09
4	0,94	66,52
5	0,94	64,21

4. Discussões

Como esperado, é possível notar que o pior desempenho foi quando somente uma semana de entrada foi utilizada, com R^2 de 0,94 e $RMSE$ de 83,36. Já o melhor desempenho ocorreu pela solução de semanas de entradas proposto inicialmente neste trabalho (duas semanas) com um desempenho significativamente melhor com R^2 DE 0,98 e $RMSE$ de 45,83.

Apesar do melhor resultado com duas semanas de entrada, a partir de três semanas como entrada, foi observado uma deterioração das estimativas, com R^2 de 0,97 e $RMSE$

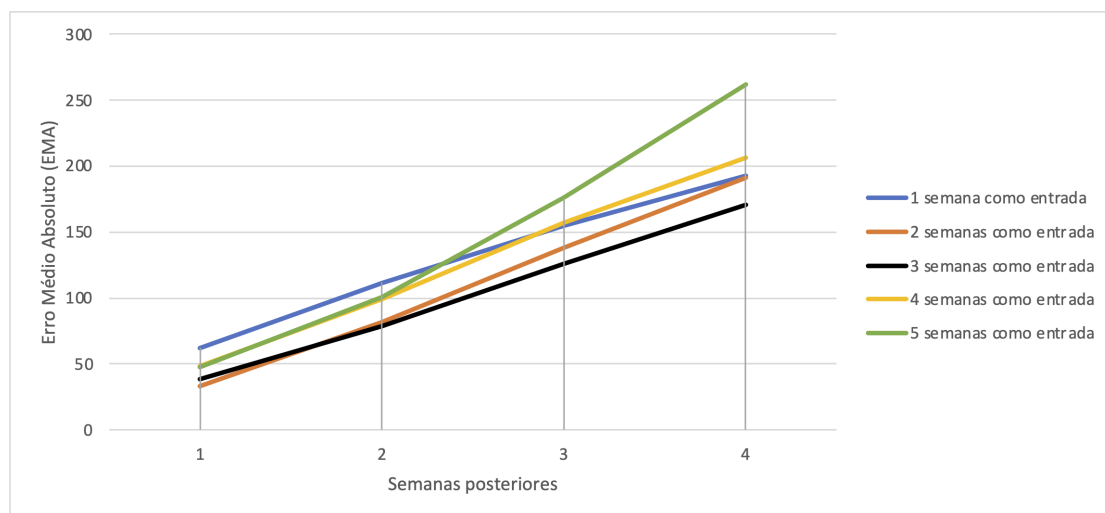


Figura 1. Estimativas de leitos para 1, 2, 3 e 4 semanas à frente.

de 56,09, em que para cinco semanas de entrada o desempenho foi ligeiramente superior que com quatro semanas de entrada, com mesmo R^2 de 0,94, mas $RMSE$ 64,21 e 66,56 respectivamente.

A inclusão de estimativa para duas, três e cinco semanas à frente do conjunto de dados revelou diferenças significativas no impacto da montagem de quantidades diferentes de semanas de entrada para a previsão de saída. Apesar do pior desempenho com uma semana de entrada, as estimativas para duas semanas até um mês demonstraram menor variação da estimativa de leitos precisos em relação a outras quantidades de semanas de entrada. Com duas semanas de entrada, o desempenho foi melhor, mas até a previsão de quatro semanas o resultado teve tendência a piorar, e com possível superação de quatro semanas de entrada para as estimativas que exijam mais de um mês.

Três semanas de entradas apresentaram a segunda menor variação da estimativa de leitos, começando com o segundo melhor desempenho e encerrando com o melhor desempenho entre as possibilidades de semanas de entrada. Por fim, cinco semanas de entrada tiveram maior variação da estimativa de leitos, acima da quantidade prevista com outras quantidades de semanas, e quatro semanas encerraram com o segundo pior desempenho entre as cinco possibilidades de entrada.

Os resultados indicam que a abordagem proposta pode ser uma ferramenta valiosa para gestores de saúde, auxiliando na alocação eficiente de recursos hospitalares e na mitigação de impactos durante crises sanitárias. O principal diferencial deste estudo está na integração inovadora de Redes Neurais Artificiais com análise de séries temporais, permitindo uma previsão contínua e dinâmica da demanda por leitos. Essa abordagem se destaca em relação a estudos anteriores, como os de [Gomes et al. 2024], que focam na otimização matemática para alocação de recursos, e [Consoli et al. 2024], que utilizam aprendizado de máquina para prever admissões hospitalares. Nossa proposta avança ao antecipar flutuações na demanda com maior precisão, oferecendo um suporte preditivo que se adapta ao comportamento temporal das internações.

Além disso, os resultados reforçam a relevância de técnicas de aprendizado de máquina aplicadas ao contexto hospitalar, contribuindo para uma tomada de decisão mais

ágil e eficiente. A combinação de previsão temporal e redes neurais fornece uma ferramenta poderosa para gestores de saúde, otimizando a alocação de leitos em um cenário de alta incerteza.

Os resultados obtidos podem ser utilizados para apoiar a tomada de decisões por gestores das redes de saúde. No entanto, a precisão das estimativas pode ser melhorada com a inclusão de variáveis adicionais além dos dados temporais, como a influência de outras doenças nos leitos; ou ainda a taxa de transmissão ou taxa de UTI, trazidas na seção de introdução através de outros trabalhos similares na literatura. Outro ponto importante foram os períodos em que 100% dos leitos estavam ocupados, não representando fielmente qual era a demanda necessária naquele período, uma vez que por estar em lotação não eram registrados novas internações.

5. Considerações Finais

Neste trabalho, foi desenvolvida uma nova metodologia para estimar a necessidade de leitos hospitalares durante uma pandemia, utilizando dados de internações em Mato Grosso, Brasil, no contexto da COVID-19. A metodologia empregou uma rede neural do tipo MLP com otimização dos hiperparâmetros por meio de Grid Search e aproveitando de dados temporais semanais, inclusive com retroalimentação das estimativas.

Os resultados demonstraram que o uso de duas semanas como entrada para o modelo apresentou o melhor desempenho em termos de R^2 e RMSE para previsões de uma semana até um mês à frente, evidenciando a robustez da abordagem. Embora o desempenho tenha se deteriorado ao aumentar o número de semanas de entrada, o modelo ainda se mostrou útil para apoiar decisões estratégicas em períodos críticos.

A precisão das estimativas pode ser aprimorada com a inclusão de variáveis adicionais, como taxas de transmissão da COVID-19, notificações de doenças emergentes, influência de outras enfermidades nos leitos e variáveis externas que afetam diretamente a dinâmica hospitalar. Além disso, o uso de indicadores mais específicos, como a taxa de ocupação de UTI ou o tempo médio de internação, pode proporcionar uma visão mais abrangente e alinhada às diversas dinâmicas do sistema de saúde. Esses aprimoramentos são recomendados para trabalhos futuros.

A inclusão desses fatores pode fortalecer o modelo, aprimorar a assertividade das estimativas e ampliar sua aplicabilidade em diferentes cenários. Além de melhorar o desempenho preditivo, esse tipo de enriquecimento de base permite antecipar surtos com maior precisão, o que oferece aos gestores informações mais confiáveis para planejar ações de resposta rápida, distribuir recursos com maior efetividade e reduzir os impactos em períodos de sobrecarga do sistema de saúde.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Os autores também agradecem o apoio da Secretaria de Estado de Saúde de Mato Grosso pela disponibilização dos dados por meio do Convênio nº 126/2022.

Referências

- Ali, Y. A., Awwad, E. M., Al-Razgan, M., and Maarouf, A. (2023). Hyperparameter search for machine learning algorithms for optimizing the computational complexity. *Processes*, 11(2).
- Almeida, J. F. d. F., Conceição, S. V., Pinto, L. R., Horta, C. J. G., Magalhães, V. S., and Campos, F. C. C. d. (2021). Estimating brazilian states' demands for intensive care unit and clinical hospital beds during the covid-19 pandemic: Development of a predictive model. *Sao Paulo Medical Journal*, 139(2):178–185.
- Bekker, R., uit het Broek, M., and Koole, G. (2023). Modeling covid-19 hospital admissions and occupancy in the netherlands. *European Journal of Operational Research*, 304(1):207–218. The role of Operational Research in future epidemics/pandemics.
- Braga, M. d. B., Fernandes, R. d. S., Souza, Jr, G. N. d., Rocha, J. E. C. d., Dolácio, C. J. F., Tavares, Jr, I. d. S., Pinheiro, R. R., Noronha, F. N., Rodrigues, L. L. S., Ramos, R. T. J., Carneiro, A. R., Brito, S. R. d., Diniz, H. A. C., Botelho, M. d. N., and Vallinoto, A. C. R. (2021). Artificial neural networks for short-term forecasting of cases, deaths, and hospital beds occupancy in the covid-19 pandemic at the brazilian amazon. *PLOS ONE*, 16(3):1–27.
- Car, Z., Baressi Šegota, S., Anđelić, N., Lorencin, I., and Mrzljak, V. (2020). Modeling the spread of covid-19 infection using a multilayer perceptron. *Computational and Mathematical Methods in Medicine*, 2020(1):5714714.
- Consoli, B., Viera, R., Bordini, R. H., and Manssour, I. H. (2024). Predicting inpatient admissions in brazilian hospitals. In *Anais do 24º Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 284–295. Sociedade Brasileira de Computação.
- Gomes, C. E. V., Clímaco, G., de Carvalho, J. A. F. N., and Quintanilha, D. B. P. (2024). Modelagem matemática para a alocação de leitos de uti com diferentes tipos de leitos e taxa de ocupação. In *Anais do 24º Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*, pages 226–237. Sociedade Brasileira de Computação.
- Marcilio, I., Lazar Neto, F., Lazzeri Cortez, A., Miethke-Morais, A., Dutilh Novaes, H. M., Possolo de Sousa, H., de Carvalho, C. R. R., Shafferman Levin, A. S., Ferreira, J. C., Gouveia, N., and Group, H. C.-. S. (2022). Mortality over time among covid-19 patients hospitalized during the first surge of the pandemic: A large cohort study. *PLOS ONE*, 17(9):1–12.
- Ritter, M., Ott, D. V. M., Paul, F., Haynes, J.-D., and Ritter, K. (2021). Covid-19: A simple statistical model for predicting intensive care unit load in exponential phases of the disease. *Scientific Reports*, 11:5018.
- Runge, M., Richardson, R. A. K., Clay, P. A., Bell, A., Holden, T. M., Singam, M., Tsu-boyama, N., Arevalo, P., Fornoff, J., Patrick, S., Ezike, N. O., and Gerardin, J. (2022). Modeling robust covid-19 intensive care unit occupancy thresholds for imposing mitigation to prevent exceeding capacities. *PLOS Global Public Health*, 2(5):1–17.
- Wacksman, J. (2021). Digitalization of contact tracing: balancing data privacy with public health benefit. *Ethics and Information Technology*, 23:855–861.