

TB-MT Dashboard: Uma abordagem utilizando estatística espaço-temporal associada à aprendizagem de máquina no desenvolvimento de um Painel Inteligente

Anthony M. P. Oliveira¹, Vinicius S. Ferrari², Vinícius B. Bastos³,
Bruno S. Oliveira⁴, Josilene D. Alves⁵ André S. Abade⁶

¹EpiGeo Araguaia - Grupo de Pesquisa em Epidemiologia e Geoprocessamento do Araguaia
Universidade Federal de Mato Grosso (UFMT) – Campus Araguaia
Unidade II – Barra do Garças – MT – Brasil

²Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
Campus Barra do Garças – MT – Brasil

anthonymuniz12@gmail.com

Abstract. *This paper outlines an Intelligent Panel project for Tuberculosis Monitoring in Mato Grosso, utilizing data analysis and machine learning. With a focus on time series, the project aims to refine SINAN data and develop a predictive model to analyze the likelihood of patient abandonment or mortality in Tuberculosis cases in the state. The objective is to contribute to epidemiological control, providing valuable insights for healthcare professionals to enhance treatment administration and improve patient outcomes. The project also aims to offer an accessible interface for healthcare practitioners and researchers alike.*

Resumo. *Este trabalho descreve um projeto de Painel Inteligente para Monitoramento da Tuberculose em Mato Grosso, utilizando análise de dados e aprendizado de máquina. Com enfoque em séries temporais, o projeto busca refinar dados do SINAN e desenvolver um modelo preditivo para analisar a probabilidade de abandono ou óbito de pacientes com Tuberculose no estado. O objetivo é contribuir para o controle epidemiológico, fornecendo insights valiosos para os profissionais de saúde aprimorarem a administração do tratamento e melhorarem os resultados dos pacientes. O projeto também visa oferecer uma interface acessível tanto para profissionais da saúde quanto para pesquisadores.*

1. Introdução

As novas estratégias globais do setor de saúde para 2022-2030, propõem uma visão comum para acabar com as epidemias e promover a cobertura universal de saúde, os cuidados de saúde primários e a segurança sanitária em um mundo no qual todas as pessoas têm acesso a serviços de saúde de qualidade e baseados em evidências. É preciso destacar a necessidade de otimizar sistemas, setores e parcerias para promover sinergias com os cuidados de saúde primários, a governança da saúde, o financiamento, a força de trabalho, os produtos e a prestação de serviços, ao mesmo tempo que promove respostas multissetoriais aos determinantes sociais e estruturais da saúde. Adicionalmente, recolher, analisar e utilizar dados, com desagregação por sexo, idade e/ou características relevantes

da população para monitorar e avaliar o progresso para orientar as ações, a inovação, a investigação e o desenvolvimento para promover soluções com alto impacto nos processos de tomada de decisão e a formulação de estratégias de saúde pública [Organization 2022].

Neste contexto, gerar dados de alta qualidade e utilizar análises de dados para impulsionar ações, inclusive em níveis descentralizados, aplicando a ciência de dados como força motriz desse processo [Pacheco et al. 2019], pode ser um diferencial determinante para alcançar as metas propostas. Dados precisos, oportunos e granulares são essenciais para o planejamento estratégico nacional, a atribuição de recursos, a prestação de serviços de saúde, a defesa e a responsabilização. Esforços conjuntos para fortalecer os sistemas de informação estratégica, incluindo estimativas de carga populacional e desagregação relevante de dados por sexo, idade e outras características relevantes da população, são importantes para orientar implementação do programa e melhoria da qualidade.

A ciência de dados emerge como uma ferramenta poderosa no combate a diversas doenças globais, incluindo a tuberculose (TB). Sua relevância para alcançar os objetivos da estratégia END-TB 2030, que visa erradicar a TB como problema de saúde pública, é multifacetada e impactante. Por meio da análise de grandes volumes de dados, modelagem preditiva e aprendizado de máquina, a ciência de dados fornece *insights* cruciais que podem informar e orientar políticas de saúde pública, estratégias de prevenção, diagnóstico precoce, e tratamento eficaz da TB [Hripcsak et al. 2015, World Health Organization 2023].

A TB é uma doença evitável e geralmente curável. No entanto, em 2022, ela foi a segunda principal causa de morte no mundo devido a um único agente infeccioso, depois da doença por coronavírus (COVID-19), e causou quase o dobro de mortes que o HIV/SIDA. Mais de 10 milhões de pessoas continuam a adoecer com TB todos os anos. É necessária uma ação urgente para acabar com a epidemia global de TB até 2030, um objetivo adotado por todos os Estados-Membros das Nações Unidas (ONU) e pela Organização Mundial da Saúde (OMS) [World Health Organization 2023].

Nesta perspectiva, este trabalho está relacionado à complexidade e à persistência dos desafios enfrentados no controle da TB, principalmente quanto aos desfechos negativos da doença. Outro agravante que pode ser mencionado são as limitações dos profissionais e gestores de saúde quanto ao uso da epidemiologia, técnicas estatísticas, modelagens matemáticas e tecnologias computacionais. Além disso, tem-se a crescente necessidade de políticas públicas que sejam baseadas em evidências científicas. Diante dessas questões tem-se a necessidade de inovações tecnológicas que possam ser aliadas ao diagnóstico, tratamento e monitoramento da TB.

Diante dos desafios mencionados, é crucial intensificar os esforços para otimizar tempo, investimentos financeiros e recursos humanos no SUS. Este trabalho adota uma abordagem transversal, visando integrar diversas especialidades e áreas do conhecimento para contribuir principalmente com a redução dos desfechos negativos da TB. O TB-MT DashBoard propõe uma abordagem inovadora e tecnológica ao incorporar técnicas avançadas de inteligência artificial, análise de dados e modelagem matemática, que podem auxiliar na previsão de desfechos negativos, refletindo na melhoria do cuidado ao paciente e na redução da morbidade e mortalidade associadas à TB. Destaca-se que o TB-MT DashBoard é uma ferramenta abrangente, que engloba desde pacientes até gestores públicos, sendo uma iniciativa relevante para o enfrentamento da TB no Estado de Mato

Grosso, com aplicabilidade direta e benéfica ao SUS.

2. Trabalhos Relacionados

Nesta seção, as principais abordagens que utilizam ciência de dados, modelagem preditiva e series temporais são apresentadas. Embora poucas destes estudos utilizem essas tecnologias para predição de séries temporais em TB, ampliamos nosso escopo para investigações que contribuem de alguma forma para o preenchimento dessa lacuna do conhecimento. Assim, foi possível traçar uma estratégia com base nas singularidades, fragilidades e avanços já relatados.

A proposta de [Sulaiman and Yahaya 2013] aborda o desenvolvimento de um painel para análise de doenças cardiovasculares, com base em um esquema estelar. A pesquisa teve como objetivo integrar dados médicos dispersos em um único armazém de dados, identificando especificações de *design* para essa finalidade. O estudo também incluiu a implementação do armazém de dados utilizando dois tipos de bancos de dados e enfatizou a importância da integração de dados na análise de informações médicas.

A abordagem proposta por [Silva et al. 2014], apresenta um modelo preditivo com variáveis explicativas as características dos indivíduos com TB para definir o tratamento mais adequado. O estudo explora desde os critérios de extração de características até a construção customizada de um modelo preditivo para designar o melhor tratamento.

O estudo de [Yermalovich 2020] aborda técnicas de visualização para *dashboards*, fundamentado em um entendimento avançado dos papéis na segurança da informação. Destaca-se a importância de compreender esses papéis na concepção de ferramentas de visualização, enfatizando que uma representação visual detalhada pode acelerar a resposta a incidentes de segurança. A adaptação específica a cada papel permite uma representação mais ergonômica dos dados. O artigo continua com a análise de outros aspectos relacionados à segurança da informação e visualização de dados.

Os estudos conduzidos por [Kenigsberg et al 2022], concentra-se no desenvolvimento de um painel para a visualização em tempo quase real de dados de vigilância de segurança da vacina COVID-19. Esse *dashboard* interno foi projetado para proporcionar uma visualização facilitada e revisões rápidas de grandes conjuntos de dados semanais automatizados de segurança de vacinas. O objetivo principal é apoiar a vigilância ativa e a pesquisa de segurança de vacinas, utilizando Análise de Ciclo Rápido para monitoramento em tempo quase real das vacinas COVID-19. A ferramenta permite revisões rápidas de dados, contribuindo assim para a tomada de decisões em saúde pública.

3. Materiais e Método

Um fluxo de trabalho par ao método proposto é descrito na Figura 1 e as subseções a seguir apresentam cada detalhe envolvido no processo de construção da abordagem proposta neste estudo.

É preciso destacar que todos os algoritmos desenvolvidos para a coleta, tratamento dos dados, modelagem preditiva e visualização de dados foram implementados na linguagem Python. A escolha foi motivada pela simplicidade, concisão e pela ampla gama de ferramentas que oferece, facilitando assim o manuseio eficiente dos dados.

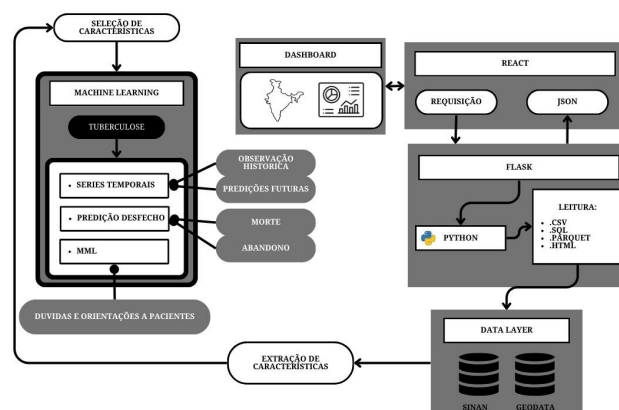


Figura 1. Fluxo de trabalho do método para construção do Painel Inteligente usando Ciência de Dados e Modelagem Preditiva para apoio aos processos de tomada de decisão em TB.

3.1. Tipo de Estudo e Cenário

Este é um estudo exploratório de séries temporais, aplicando modelagem preditiva, realizado no estado de Mato Grosso, localizado na região Centro-Oeste do Brasil caracterizado pelas coordenadas aproximadas de latitude 14° a 17° sul e longitude 53° a 58° oeste. O Mato Grosso, especialmente em áreas de vulnerabilidade socioeconômica, pode ter taxas significativas dessas condições, tornando-o um local de estudo importante.

O estado de Mato Grosso é um importante corredor de migração e tem significativa mobilidade populacional devido ao trabalho sazonal na agricultura e outras indústrias. Isso pode afetar os padrões de transmissão e a dinâmica da infecção por TB. Adicionalmente, o Mato Grosso é caracterizado por uma diversidade populacional significativa, incluindo áreas urbanas densas, comunidades rurais e povos indígenas. Esta diversidade permite a análise da TB em diferentes contextos sociais, econômicos e culturais.

3.2. Conjunto de Dados

A população do estudo consistiu em todos os casos de TB do estado de Mato Grosso notificados no Sistema de Informação de Agravos de Notificação (SINAN). Somente pessoas com diagnóstico de TB de acordo com a Classificação Internacional de Doenças (CID-10), foram incluídos no estudo. O SINAN é um sistema de informação brasileiro responsável por registrar e processar informações sobre doenças de notificação obrigatória em todo o país, mantido pelo Departamento de Informática do Sistema Único de Saúde (DATASUS)[Departamento de Informática do SUS (DATASUS) 2024]. Os dados populacionais foram extraídos do Censo Demográfico 2022, disponíveis no portal eletrônico do Instituto Brasileiro de Geografia e Estatística (IBGE)[Instituto Brasileiro de Geografia e Estatística (IBGE) 2022].

3.3. Análise dos Dados

Antes de começar a análise de dados, foi essencial compreender as variáveis disponíveis no SINAN para remover redundâncias e garantir a integridade dos dados. Devido à

complexidade do arranjo de dados do SINAN, optou-se por extrair informações usando técnicas de Raspagem de Dados (*Web Scraping*) com Python. Essa abordagem, definida por [Barbosa and Cavalcanti 2020] como a raspagem de dados diretamente da web, na qual informações relevantes são extraídas de sites por meio de bots, sendo posteriormente submetidas à análise. Esse método revelou-se crucial para identificar padrões e tomar decisões com maior precisão.

Após o pré-processamento com os dados devidamente tratados, foi computada a taxa de incidência TB mensal de acordo com o número de casos e a população, estratificada por alguns indicadores sociodemográficos, como por exemplo, sexo, faixa etária, raça e grau de escolaridade. Para o cálculo da taxa de incidência da TB, utilizou-se a seguinte fórmula:

$$\frac{\text{Numero de novos casos confirmados de TB/HIV de residentes do estado de Mato Grosso}}{\text{Numero total de residentes no Estado de Mato Grosso durante o periodo}} \times 100.000 \text{ Pessoas} \quad (1)$$

3.4. Análise Espaço-Temporal

A análise espaço-temporal é uma abordagem na qual dados são analisados considerando tanto a dimensão espacial quanto a dimensão temporal. Isso significa que além de considerar onde os eventos ocorrem (dimensão espacial), também leva em conta quando esses eventos acontecem (dimensão temporal). Essa análise permite entender como os eventos se distribuem no espaço e como essa distribuição muda ao longo do tempo.

Para realizar essas análises, os dados foram agregados para corresponder à granularidade espacial almejada, com agregação de dados de nível de município para nível estadual. Adicionalmente, os dados foram normalizados para facilitar a comparação entre diferentes regiões ou períodos de tempo.

Com os dados pré-processados, é realizada a análise espacial e temporal. Para análise espacial, utiliza-se a biblioteca *geopandas* para manipular dados geoespaciais, incluindo arquivos *shapefile* que contêm informações sobre os limites geográficos das regiões de interesse. Análises espaciais são realizadas para identificar padrões e tendências na distribuição da incidência da tuberculose ao longo do tempo e do espaço. Isso pode incluir cálculos de densidade, análise de *clusters* espaciais e outros métodos.

Com os resultados da análise em mãos, passa-se para a etapa de visualização de dados. Utilizando bibliotecas de visualização como *matplotlib*, *seaborn*, *folium* e *plotly*, são criados mapas e gráficos que representam os resultados das análises espaciais e temporais. Mapas de calor, mapas de coropléticos ou mapas de pontos são criados para visualizar a distribuição espacial da taxa de incidência da tuberculose, enquanto gráficos de linha, gráficos de barras ou gráficos de caixa são utilizados para visualizar variações temporais na incidência.

3.5. Série Temporal

A análise de séries temporais é uma ferramenta essencial em estudos epidemiológicos, permitindo a organização e análise de informações quantitativas ao longo do tempo. No contexto da TB, a análise de séries temporais pode ser aplicada para identificar padrões

de incidência da doença, avaliar tendências ao longo do tempo e prever cenários futuros com base em dados históricos [Antunes and Cardoso 2015].

A aplicação da análise de séries temporais em estudos epidemiológicos de acordo com [Antunes and Cardoso 2015] assume uma importância fundamental ao facultar a antecipação de cenários futuros na distribuição de doenças dentro da população. Adicionalmente, viabiliza a identificação de fatores capazes de modificar essa distribuição, seja para melhor ou para pior. Essa abordagem proporciona uma compreensão aprofundada da dinâmica das medidas de interesse em saúde, capacita a previsão de resultados e promove a identificação criteriosa dos fatores que exercem influência sobre os mesmos.

3.6. Modelagem Preditiva

A modelagem preditiva de séries temporais é ramo da estatística e da análise de dados que se concentra em usar informações históricas para prever eventos ou tendências futuras [Brockwell and Davis 2002]. Em nosso estudo exploramos abordagens desde a estatística clássica até o aprendizado de máquina, utilizando uma gama diversificada de métodos analíticos, como, por exemplo, suavização exponencial (*Simple Exponential Smoothing* [Brown 1963], *Double Exponential Smoothing* [Holt 1957], *Triple Exponential Smoothing* (Holt-Winters) [Chatfield 1978]), modelos autorregressivos (ARIMA) [Box and Pierce 1970] e modelos baseados em aprendizado de máquina (*Support Vector Regression* (SVR) [Cortes and Vapnik 1995], *Extreme Gradient Boosting* (XGBoost) [Chen and Guestrin 2016], *Long Short-Term Memory* (LSTM) [Hochreiter and Schmidhuber 1997], *Gated Recurrent Unit* (GRU) [Cho et al. 2014], *Convolutional Neural Network* (CNN) [LeCun et al. 1998]), para fazer previsões de coinfeção de TB/HIV e aumentar a robustez dos resultados. Se vários métodos independentes apontam para a mesma tendência ou previsão, há implicitamente um aumento na confiabilidade dos resultados alcançados.

Na implementação dos modelos, utilizamos a estratégia de otimização de parâmetros aplicando a técnica de *Grid Search* [Belete and Huchaiah 2022], intuindo melhorar o desempenho ponderando o espaço de parâmetros e a complexidade de cada modelo utilizado. Inicialmente, implementamos uma busca com um *grid* mais amplo e grosseiro, seguida de buscas mais refinadas em áreas promissoras do espaço de parâmetros para cada modelo.

3.7. Descrição do sistema de software e hardware

Os experimentos foram implementados em uma máquina Linux, com Ubuntu 22.04, processador AMD Ryzen™ 9 7950X3D, 2 GPU NVIDIA® Quadro RTX A5500 de 24GB, 128GB de RAM DDR5. Todos os modelos foram desenvolvidos utilizando as bibliotecas StatModels 0.14.1, TensorFlow 2.15, Scikit-Learn 0.20. Python 3.10.12 com as bibliotecas folium 0.16.0, geopandas 0.14.3 selenium 4.18.1 foram usados para implementação de algoritmos e normalização de dados.

4. Resultados e Discussão

Nesta seção serão apresentados os resultados alcançados nas seguintes etapas de desenvolvimento do projeto: Extração e Limpeza de Dados, Análise Exploratória, Modelagem

Preditiva e a Preparação de Visualização.

Na Tabela 1 são apresentados os resultados alcançados com extração, limpeza e normalização dos dados já aplicados agrupamentos sociodemográficos. É preciso destacar que não houve recorte de período de tempo nos resultados, sendo que os resultados correspondem ao período de 1999 à 2024.

Tabela 1. Dados Sociodemográficos - Extração e Análise Exploratória dos Dados (1999-2024)		
Idade		
Variáveis	Numero de Casos	%
01 a 04 anos	657	1.75
05 a 09 anos	486	1.29
10 a 14 anos	620	1.65
15 a 19 anos	1699	4.52
20 a 24 anos	3574	9.5
25 a 29 anos	3866	10.28
30 a 34 anos	3896	10.36
35 a 39 anos	3630	9.65
40 a 44 anos	3361	8.94
45 a 49 anos	3280	8.72
50 a 54 anos	2971	7.9
55 a 59 anos	2549	6.78
60 a 64 anos	2085	5.54
65 e +	4557	12.12
Menor 01 ano	382	1.02
Total	37613	100.0
Raça dos Paciente		
Variáveis	Numero de Casos	%
Amarela	360	0.96
Branca	8160	21.69
Indígena	2233	5.94
Parda	17741	47.17
Preta	4319	11.48
Ignorados/Em Branco	4800	12.76
Total	37613	100.0
Sexo		
Variáveis	Numero de Casos	%
Feminino	12527	33.3
Masculino	25080	66.68
Ignorado/Em Branco	6.0	0.02
Total	37613	100.0
Forma Clinica		
Variáveis	Numero de Casos	%
Extrapulmonar	3517	9.35
Pulmonar	33070	87.92
Pulmonar + Extrapulmonar	762	2.03
Ignorado	264	0.7
Total	37613	100.0
Escolaridade		
Variáveis	Numero de Casos	%
1ª a 4ª série incompleta do EF	5790	15.39
4ª série completa do EF	1990	5.29
5ª a 8ª série incompleta do EF	6688	17.78
Sem Escolaridade	2929	7.79
Educação superior completa	1506	4.0
Educação superior incompleta	520	1.38
Ensino fundamental completo	2061	5.48
Ensino médio completo	3281	8.72
Ensino médio incompleto	3270	8.69
Não se aplica	1121	2.98
Ignorado/Em Branco	8457	22.48
Total	37613	100.0
Agravos Associados a AIDS		
Variáveis	Numero de Casos	%
Não	20656.0	54.92
Sim	2575	6.85
Ignorado/Em Branco	14382	38.24
Total	37613	100.0

Na Figura 2A é apresentado o mapa construindo por meio da estatística de Moran que testa a autocorrelação local e detecta objetos espaciais com influência no indicador Moran Global. Já na Figura 2B é apresentado o mapa de calor (*hotspot*) destacando as áreas com maior incidência de TB no período avaliado. A Figura 2C representa um mapa com a taxa de incidência no estado. Por fim, a Figura 2D apresenta um mapa de clusters por número de casos notificados no estado.

Os resultados analisam séries temporais de notificações de TB com coinfeção HIV fazem um recorte temporal de 2012 a 2023. Na Figura 3, as séries temporais da taxa de incidência mensal são mostradas, estratificadas por sexo e população geral. A taxa de incidência da população total oscila ao longo do tempo, sem uma tendência clara de aumento ou diminuição. A taxa de incidência masculina parece ter uma leve tendência de aumento, com um pico por volta de 2022, enquanto a taxa de incidência feminina permanece estável.

Cada modelo foi avaliado usando uma divisão de dados em conjuntos de treinamento e validação, com cerca de 80% das observações para treinamento e 20% para validação. O conjunto de validação consistiu em 24 observações finais para cada estratificação por sexo e população total, garantindo um intervalo homogêneo entre as séries para facilitar comparações e análises. O desempenho de cada modelo conforme a

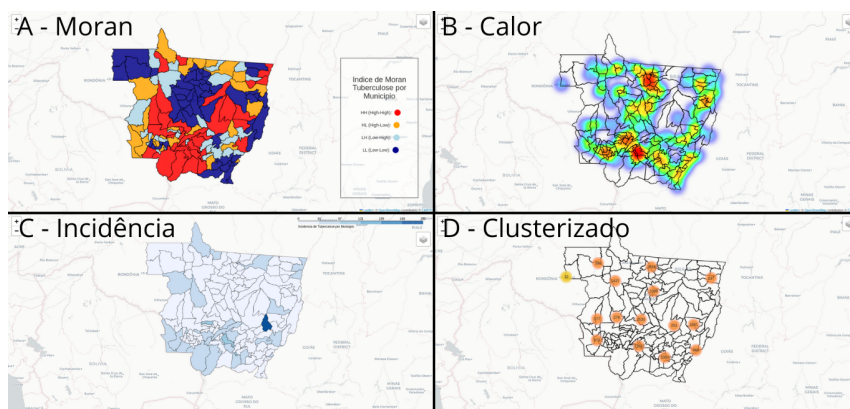


Figura 2. Análise espaço-temporal: A - Estatística de Moran; B - Mapa de Calor por número de casos; C - Mapa da taxa de incidência de TB; D - Mapa de clusters por número de casos

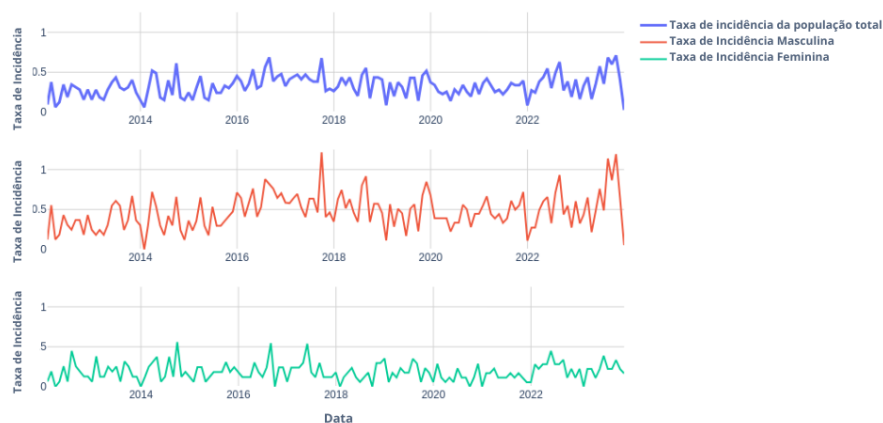


Figura 3. TB: Histórico da taxa de incidência de TB com HIV no estado de Mato Grosso - (2012 a 2023)

estratégia de treinamento pode ser visto na Figura 4A.

5. Conclusão

Em conclusão, os resultados ainda que parciais deste trabalho, demonstraram que a abordagem proposta é exitosa quanto a produção de conhecimento que podem ser utilizados para apoio ao processo de tomada de decisão no combate a TB. Os resultados enfatizam a importância da utilização de abordagens de modelagem avançadas na saúde pública para aprimorar as estratégias de prevenção e controle de doenças infecciosas, como a TB, contribuindo assim para esforços mais efetivos na redução da incidência dessas condições.

Referências

[Antunes and Cardoso 2015] Antunes, J. L. F. and Cardoso, M. R. A. (2015). Uso da análise de séries temporais em estudos epidemiológicos. *Epidemiol. Serv. Saúde*, 24(3).

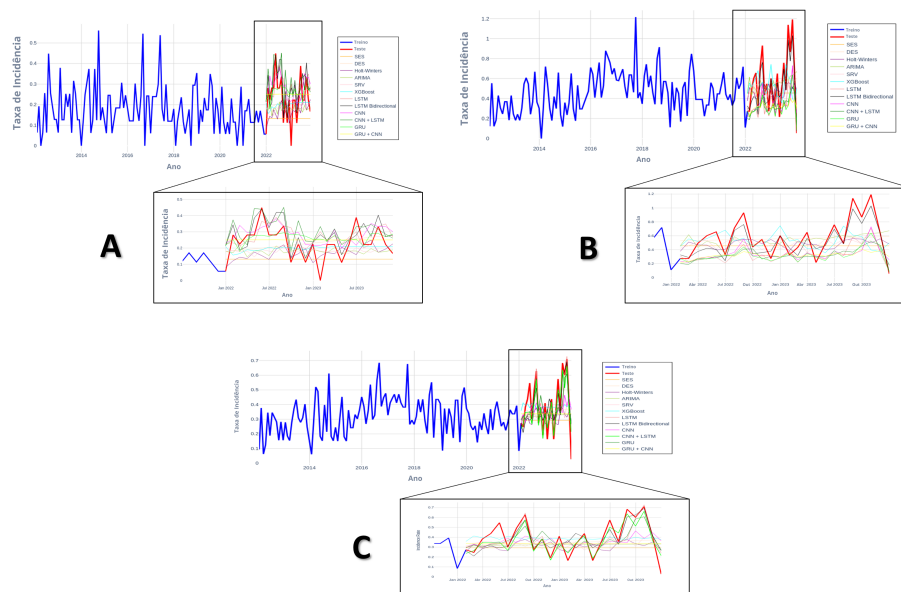


Figura 4. Visualização do treinamento dos modelos em séries temporais. A - Feminino, B - Masculino e C - População Total.

- [Barbosa and Cavalcanti 2020] Barbosa, A. B. G. and Cavalcanti, A. B. (2020). Web scraping e análise de dados. https://www.editorarealize.com.br/editora/anais/conapesc/2020/TRABALHO_EV138_MD4_SA24_ID1284_24112020001516.pdf. Acesso em: 29 fev. 2024.
- [Belete and Huchaiah 2022] Belete, D. M. and Huchaiah, M. D. (2022). Grid search in hyperparameter optimization of machine learning models for prediction of hiv/aids test results. *International Journal of Computers and Applications*, 44(9):875–886.
- [Box and Pierce 1970] Box, G. E. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332):1509–1526.
- [Brockwell and Davis 2002] Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer.
- [Brown 1963] Brown, R. G. (1963). *Smoothing, Forecasting and Prediction of Discrete Time Series*. Prentice-Hall.
- [Chatfield 1978] Chatfield, C. (1978). The holt-winters forecasting procedure. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 27(3):264–279.
- [Chen and Guestrin 2016] Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. ACM.
- [Cho et al. 2014] Cho et al., K. (2014). Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *Proceedings of the 2014 Conference on (EMNLP)*, pages 1724–1734. Association for Computational Linguistics.
- [Cortes and Vapnik 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

- [Departamento de Informática do SUS (DATASUS) 2024] Departamento de Informática do SUS (DATASUS) (2024). Sistema de informação de agravos de notificação (sinan). DataSUS - Ministério da Saúde.
- [Hochreiter and Schmidhuber 1997] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- [Holt 1957] Holt, C. C. (1957). Forecasting trends and seasonal by exponentially weighted averages. *ONR Research Memorandum, Carnegie Institute of Technology*.
- [Hripcsak et al. 2015] Hripcsak et al., G. (2015). Observational health data sciences and informatics (ohdsi): opportunities for observational researchers. *Studies in health technology and informatics*, 216:574.
- [Instituto Brasileiro de Geografia e Estatística (IBGE) 2022] Instituto Brasileiro de Geografia e Estatística (IBGE) (2022). Censo demográfico 2022. Instituto Brasileiro de Geografia e Estatística.
- [Kenigsberg et al 2022] Kenigsberg et al, T. A. (2022). Dashboard development for near real-time visualization of covid-19 vaccine safety surveillance data in the vaccine safety datalink. <https://doi.org/10.1016/j.vaccine.2022.04.010>. Acessado em: 28 fev. 2024.
- [LeCun et al. 1998] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- [Organization 2022] Organization, W. H. (2022). Global health sector strategies on hiv, viral hepatitis and sexually transmitted infections for the period 2022–2030. <https://iris.who.int/bitstream/handle/10665/360348/9789240053779-eng.pdf?sequence=1>.
- [Pacheco et al. 2019] Pacheco, M., Brister, B., Disconzi, S., and Salton, M. (2019). Ciência de dados: Enfoque no desafio do processamento. <https://www.redalyc.org/articulo.oa?id=560662202012>. Acesso em: 29 fev. 2024.
- [Silva et al. 2014] Silva, R. F. et al. (2014). Modelo preditivo ao abandono do tratamento da tuberculose. *Saúde em Debate*.
- [Sulaiman and Yahaya 2013] Sulaiman, N. S. and Yahaya, J. H. (2013). Development of dashboard visualization for cardiovascular disease based on star scheme. *Procedia Technology*, 11.
- [World Health Organization 2023] World Health Organization (2023). Global tuberculosis report 2023. <https://iris.who.int/bitstream/handle/10665/373828/9789240083851-eng.pdf?sequence=1>.
- [Yermalovich 2020] Yermalovich, P. (2020). Dashboard visualization techniques in information security. In *Proceedings of the IEEE*. IEEE.