**Planejamento Urbano Sustentável com Base em Previsões de Dengue: Uma Abordagem com Séries Temporais**

PROJETO APLICADO IV

VALDINEY ATÍLIO PEDRO – 10424616

PATRICIA CORREA FRANÇA – 10423533

MARIANA SIMÕES RUBIO – 10424388

**UNIVERSIDADE PRESBITERIANA MACKENZIE**

São Paulo

2025

**SUMÁRIO**

[1. INTRODUÇÃO 5](#_Toc209769904)

[1.1 Contexto do Trabalho 5](#_Toc209769905)

[1.2 Motivação e Justificativa 5](#_Toc209769906)

[1.3. Objetivo Geral: 5](#_Toc209769907)

[1.4. Descrição da Base de dados 6](#_Toc209769908)

[2. Referencial Teórico 6](#_Toc209769909)

[2.1. Modelos ARIMA e SARIMA 6](#_Toc209769910)

[2.2. Algoritmos de Aprendizado de Máquina 6](#_Toc209769911)

[2.3. Redes Neurais Recorrentes (LSTM e GRU) 6](#_Toc209769912)

[2.4. Modelos Bayesianos (InfoDengue) 7](#_Toc209769913)

[2.5. Conceitos fundamentais que embasam a solução incluem: 7](#_Toc209769914)

[3. PIPELINE DA SOLUÇÃO 7](#_Toc209769915)

[3.1. Coleta de Dados 7](#_Toc209769916)

[3.2. Pré-processamento 7](#_Toc209769917)

[3.3. Análise Exploratória 7](#_Toc209769918)

[3.4. Modelagem 7](#_Toc209769919)

[3.5. Validação e Otimização 8](#_Toc209769920)

[3.6. Deploy e Automação 8](#_Toc209769921)

[3.7. Documentação e Comunicação 8](#_Toc209769922)

[4.CRONOGRAMA 9](#_Toc209769923)

[5. REFERÊNCIAS BIBLIOGRAFICAS 10](#_Toc209769924)

[5.1. Anexos 10](#_Toc209769925)

# 1. INTRODUÇÃO

## 1.1 Contexto do Trabalho

A urbanização acelerada e muitas vezes desordenada tem contribuído para o aumento da vulnerabilidade das cidades brasileiras frente a doenças transmitidas por vetores, como a dengue. A presença de áreas com infraestrutura precária, saneamento insuficiente e descarte inadequado de resíduos favorece a proliferação do mosquito *Aedes aegypti*, tornando a dengue um problema recorrente de saúde pública urbana.

Este projeto propõe o uso de técnicas de previsão baseadas em séries temporais para antecipar surtos de dengue em municípios brasileiros, contribuindo para o planejamento urbano sustentável e para a tomada de decisões mais eficazes em políticas públicas locais.

## 1.2 Motivação e Justificativa

A escolha do tema está diretamente relacionada ao **ODS 11 – Cidades e Comunidades Sustentáveis**, que visa tornar os assentamentos humanos inclusivos, seguros, resilientes e sustentáveis. A previsão de surtos de dengue pode:

* Apoiar ações preventivas em áreas urbanas vulneráveis
* Otimizar a alocação de recursos municipais
* Reduzir impactos sobre a saúde da população e a infraestrutura urbana
* Promover cidades mais resilientes frente a riscos sanitários

Além disso, o uso de dados abertos e confiáveis do sistema **InfoDengue**, mantido pela Fiocruz e pelo Ministério da Saúde, permite análises robustas e aplicáveis à realidade brasileira.. Objetivo Geral e Objetivos Específicos

## 1.3. Objetivo Geral:

Desenvolver modelos preditivos baseados em séries temporais para estimar o número de casos prováveis de dengue em municípios brasileiros. O projeto utilizará técnicas como ARIMA, Prophet e Redes Neurais Recorrentes (LSTM), com o intuito de gerar alertas antecipados e apoiar estratégias de planejamento urbano sustentável.

## 1.4. Descrição da Base de dados

A base de dados será extraída do Sistema InfoDengue ([https://info.dengue.mat.br](https://info.dengue.mat.br/)), que reúne informações atualizadas sobre casos prováveis de dengue, zika e chikungunya no Brasil. Os dados estão organizados por município e por semana epidemiológica, permitindo análises temporais detalhadas.

* **Estrutura**: registros semanais
* **Período de coleta**: histórico de 2010 até o primeiro semestre de 2025
* **Variáveis disponíveis**: casos prováveis, incidência por 100 mil habitantes, alertas de risco, indicadores de transmissão

**2. Referencial Teórico**

A modelagem de séries temporais para previsão de dengue envolve diferentes abordagens.

## 2.1. Modelos ARIMA e SARIMA

Os modelos ARIMA e SARIMA são amplamente utilizados por sua capacidade de decompor séries em componentes de tendência e sazonalidade, oferecendo boa interpretabilidade. No entanto, exigem estacionaridade e apresentam limitações diante de padrões não lineares (HYNDMAN; ATHANASOPOULOS, 2018).

## 2.2. Algoritmos de Aprendizado de Máquina

Random Forest e XGBoost permitem incorporar variáveis exógenas (temperatura, precipitação, densidade populacional), aumentando a robustez contra outliers. Contudo, demandam engenharia de atributos e podem ignorar a sequência temporal implícita (HYNDMAN; ATHANASOPOULOS, 2018).

## 2.3. Redes Neurais Recorrentes (LSTM e GRU)

LSTM e GRU são eficazes na captura de dependências de longo prazo e padrões complexos. Estudos demonstram ganhos de acurácia em grandes volumes de dados, embora exijam alto poder computacional e cuidados para evitar sobreajuste (LOPES et al., 2019).

## 2.4. Modelos Bayesianos (InfoDengue)

Ferramentas como o InfoDengue combinam dados epidemiológicos e climáticos em modelos bayesianos, promovendo a detecção precoce de surtos. Apesar da eficácia, requerem curadoria contínua das variáveis de entrada (BRASIL, 2025).

## 2.5. Conceitos fundamentais que embasam a solução incluem:

* Estacionaridade e diferenciação
* Decomposição aditiva/multiplicativa
* Validação temporal com rolling window
* Métricas de avaliação como MAE, RMSE e MAPE

## 3. PIPELINE DA SOLUÇÃO

* 1. **Coleta de Dados**
  + Extração de séries semanais via API do InfoDengue (2010–2025).
  + Incorporação de variáveis exógenas: temperatura e precipitação (INMET), densidade populacional (IBGE).

## Pré-processamento

* + Tratamento de valores faltantes e outliers.
  + Aplicação de diferenciação para garantir estacionaridade.
  + Normalização das variáveis e criação de lags (1–4 semanas) e janelas móveis (mínimo, máximo, média).

## Análise Exploratória

* + Decomposição da série em tendência, sazonalidade e resíduos.
  + Cálculo de autocorrelações (ACF/PACF) e correlações com variáveis exógenas.
  + Visualização espacial por município com mapas interativos.

## Modelagem

* + Implementação de ARIMA/SARIMA como baseline.
  + Treinamento de XGBoost com variáveis exógenas e atributos temporais.
  + Desenvolvimento de rede LSTM para capturar padrões complexos e dependências de longo prazo.

## Validação e Otimização

* + Validação cruzada com janela deslizante (rolling window cross-validation).
  + Otimização de hiperparâmetros via grid search e Bayesian optimization.

## Deploy e Automação

* + Encapsulamento do pipeline em container Docker.
  + Exposição via API RESTful e dashboard interativo com visualizações dinâmicas.

## Documentação e Comunicação

* + Elaboração de relatório técnico com resultados, interpretações e recomendações.
  + Apresentação final destacando a contribuição ao ODS 11 e a aplicabilidade social da solução.

# 4.CRONOGRAMA



# 5. REFERÊNCIAS BIBLIOGRAFICAS

**BRASIL.** Ministério da Saúde. *InfoDengue*. Disponível em:<https://info.dengue.mat.br>. Acesso em: 29 ago. 2025.

**ORGANIZAÇÃO MUNDIAL DA SAÚDE.** *Dengue and severe dengue*. Disponível em:<https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>. Acesso em: 29 ago. 2025.

**HYNDMAN, R. J.; ATHANASOPOULOS, G.** *Forecasting: principles and practice*. Melbourne: OTexts, 2018.

**LOPES, F. M.** *et al.* Time series analysis of dengue incidence in Brazil using ARIMA models. *Revista de Saúde Pública*, São Paulo, v. 53, n. 1, p. 1–8, 2019.

## 5.1. Anexos

https://github.com/valdineyatilio/ProjetoAplicado-IV/tree/main