

**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

<b>1. INTRODUÇÃO .....</b>	<b>5</b>
<b>1.1 Contexto do Trabalho .....</b>	<b>5</b>
<b>1.2 Motivação e Justificativa .....</b>	<b>5</b>
<b>1.3. Objetivo Geral: .....</b>	<b>5</b>
<b>1.4. Descrição da Base de dados.....</b>	<b>6</b>
<b>2. Referencial Teórico .....</b>	<b>6</b>
<b>2.1. Modelos ARIMA e SARIMA.....</b>	<b>6</b>
<b>2.2. Algoritmos de Aprendizado de Máquina.....</b>	<b>6</b>
<b>2.3. Redes Neurais Recorrentes (LSTM e GRU) .....</b>	<b>6</b>
<b>2.4. Modelos Bayesianos (InfoDengue).....</b>	<b>7</b>
<b>2.5. Conceitos fundamentais que embasam a solução incluem:.....</b>	<b>7</b>
<b>3. PIPELINE DA SOLUÇÃO .....</b>	<b>7</b>
<b>3.1. Coleta de Dados .....</b>	<b>7</b>
<b>3.2. Pré-processamento.....</b>	<b>7</b>
<b>3.3. Análise Exploratória .....</b>	<b>7</b>
<b>3.4. Modelagem .....</b>	<b>7</b>
<b>3.5. Validação e Otimização .....</b>	<b>8</b>
<b>3.6. Deploy e Automação.....</b>	<b>8</b>
<b>3.7. Documentação e Comunicação .....</b>	<b>8</b>
<b>4.CRONOGRAMA .....</b>	<b>9</b>
<b>5. REFERÊNCIAS BIBLIOGRAFICAS .....</b>	<b>10</b>
<b>5.1. Anexos .....</b>	<b>10</b>

# 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>), 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**

### **3.1. Coleta de Dados**

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

### **3.2. 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).

### **3.3. 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.

### **3.4. 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.

### **3.5. Validação e Otimização**

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

### **3.6. Deploy e Automação**

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

### **3.7. 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

Período	Atividade	Marco / Entregável
13/08 – 29/08	Definição de equipe, tema, proposta e descrição da base	Entrega 1 (29/08)
30/08 – 05/09	Coleta automatizada de dados InfoDengue, INMET e IBGE; documentação de metadados	Dados coletados e documentados
06/09 – 10/09	Pesquisa bibliográfica e redação do Referencial Teórico (versão inicial)	Rascunho do referencial teórico
11/09 – 17/09	Pré-processamento: limpeza, imputação, diferenciação, criação de lags	Dataset pré-processado
18/09 – 20/09	Implementação de modelos baseline (Naïve; Média Móvel; ARIMA/SARIMA)	Resultados iniciais de baseline
21/09 – 23/09	Treinamento e avaliação de XGBoost com variáveis exógenas	Resultados XGBoost
24/09 – 25/09	Ajustes finais no pipeline e preparação para Entrega 2	Pipeline refinado
26/09	Entrega 2: Referencial Teórico, Pipeline da Solução e Cronograma	Entrega 2
27/09 – 03/10	Implementação de LSTM: definição da arquitetura, treinamento inicial	Resultados iniciais de LSTM
04/10 – 10/10	Otimização de hiperparâmetros (grid search e Bayesian) e validação rolling window	Modelos otimizados e validados
11/10 – 17/10	Containerização Docker, API RESTful e desenvolvimento de dashboard	Ambiente de deploy e dashboard
18/10 – 24/10	Redação de relatório intermediário e notebooks executáveis	Notebook e relatório parcial
25/10 – 31/10	Revisão geral, ensaio de apresentação e ajustes para Entrega 3	Entrega 3 (31/10)
01/11 – 12/11	Compilação de artefatos finais no GitHub (códigos, dados, documentação)	Repositório organizado
13/11 – 19/11	Roteiro, gravação e edição do vídeo de apresentação	Vídeo pronto para avaliação
20/11 – 26/11	Ajustes finais em artigo, notebook e vídeo; revisão geral	Artefatos finais prontos
27/11 – 28/11	Buffer para imprevistos e submissão da entrega final	Entrega 4 (28/11)

## 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>