

AMANDA DOS SANTOS

Semestre atual: 2024.1

DIRETORIA DO CAMPUS VARGINHA (11.64)

PORTAL DO DISCENTE > PROJETO DE PESQUISA

DADOS DO PROJETO DE PESQUISA

Código: PVC00613-2024**Título:** Algoritmos de aprendizado de máquina para predição de potenciais locais de criadouros do *Aedes aegypti***Tipo:** INTERNO (Projeto Novo)**Natureza do Projeto:** Projeto de Desenvolvimento Científico e Tecnológico**Tipo de Pesquisa:** Pesquisa Aplicada**Situação:** EM EXECUÇÃO**Unidade de Lotação do Coordenador:** DEPARTAMENTO DE SISTEMAS DE INFORMAÇÃO - VG (11.64.12)**Unidade de Execução:** DEPARTAMENTO DE SISTEMAS DE INFORMAÇÃO - VG (11.64.12)**Centro:** CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS (11.00)**Palavra-Chave:** *Aedes aegypti*, aprendizado de máquina, previsão de arbovírus**E-mail:** eduardogomes@cefetmg.br**Editais:** EDITAL DPPG Nº 24/2024 - PIBIC/PIBIC-Af/PIBITI CNPq**Cota:** PIBIC/PIBIC-Af/PIBITI CNPq 2024-2025 (01/07/2024 a 31/08/2025)

HISTÓRICO DE EDITAIS/COTAS

Edital	Cota	Período da Cota
EDITAL DPPG Nº 24/2024 - PIBIC/PIBIC-Af/PIBITI CNPq	PIBIC/PIBIC-Af/PIBITI CNPq 2024-2025	01/07/2024 a 31/08/2025

OBJETIVOS DE DESENVOLVIMENTO SUSTENTÁVEL



ÁREA DE CONHECIMENTO

Grande Área: Ciências Exatas e da Terra**Área:** Ciência da Computação**Subárea:** Metodologia e Técnicas da Computação**Especialidade:** Sistemas de Informação

GRUPO E LINHA DE PESQUISA

Grupo de Pesquisa: Núcleo de Excelência em Empreendedorismo e Inovação Tecnológica - CEFET/MG**Linha de Pesquisa:** Análise de Dados e Inteligência Artificial

CORPO DO PROJETO

Resumo

Todos os anos, o mosquito *Aedes aegypti* infecta milhões de pessoas com doenças graves como dengue, zika, chikungunya e febre amarela urbana. A identificação e remoção de criadouros é uma estratégia fundamental para prevenir essas doenças, mas o processo tradicional é demorado e demanda muitos recursos humanos, haja vista que é realizado de forma manual e com nenhum ou pouco apoio de tecnologias. Tal tarefa é dificultada também em razão da extensão do território mineiro. Contudo, ao analisar dados históricos de infecções, pode-se identificar padrões e regiões com alta probabilidade de novos criadouros. Para consecução dessa tarefa, modelos de aprendizado de máquina poderão ser úteis. O objetivo da presente proposta é desenvolver modelos preditivos usando aprendizado de máquina para identificar áreas potenciais de novos criadouros.

Introdução/Justificativa

(incluindo os benefícios esperados no processo ensino-aprendizagem e o retorno para os cursos e para os professores da instituição em geral)

Todos os anos, o mosquito *Aedes aegypti* infecta milhões de pessoas com doenças graves como dengue, zika, chikungunya e febre amarela urbana, sendo a mais conhecida a dengue. Apesar do recente desenvolvimento de vacinas contra a dengue, permanecem outras doenças transmitidas por esse vetor. Recentemente a febre do Mayaro vem ganhando notoriedade, e alguns especialistas trabalham com a hipótese dele ser transmitido também pelo mosquito *Aedes aegypti*.

Paul et al. (2018) advogam que a prevenção das arboviroses depende da priorização da eliminação do vetor, ou seja, dos focos de *Aedes aegypti*, e de estratégias de predição que possam fornecer aos gestores de saúde informações adequadas para prevenir surtos de arboviroses. De acordo com Ribeiro et al. (2021), a principal estratégia dos programas de controle do *Aedes aegypti* baseia-se no LIRAA (Levantamento de Índice Rápido para o *Ae. Aegypti*), um inquérito amostral no qual o Índice de Infestação Predial obtido é utilizado para priorizar áreas de intervenção. Contudo, Ribeiro et al. (2021) analisaram o desempenho do LIRAA quanto à sua sensibilidade na previsão de epidemias de dengue em municípios do Estado do Rio de Janeiro, em anos considerados epidêmicos e constataram um baixo desempenho do LIRAA na previsão de epidemias de dengue. Anteriormente, McCormack-Gelles et al. (2020) também demonstraram que os índices aferidos pelo LIRAA não apresentaram confiabilidade para prever a escala ou a localização da transmissão imediata ou posterior da dengue em Fortaleza (Ceará). Assim, investir em novas estratégias para predição torna-se relevante (Ribeiro et al., 2021). Lima et al. (2022) realizaram uma revisão sistemática da literatura para identificar modelos de predição de arbovírus, bem como modelos para a dinâmica de seus vetores transmissores. Os autores identificaram uma lacuna existente na construção de modelos espaço-temporais e concluíram que existe um grande potencial para explorar modelos com abordagem espaço-temporal. No Brasil, Silva et al. (2021) investigaram modelos de previsão para prever a distribuição espaço-temporal de casos a partir de dados de notificação primária de saúde e variáveis climáticas (velocidade do vento, temperatura e pluviosidade) em Recife (PE), demonstrando a eficácia da análise espaço-temporal sobre perceptrons multicamadas e resultados de máquinas de vetores de suporte. Contudo, Silva et al. (2021) sugeriram em estudos futuros aplicar essas e outras técnicas para localização de potenciais criadouros do mosquito *Aedes aegypti*. Apoiando em Lima et al. (2022), pode-se afirmar que modelos de aprendizagem de máquina podem ser uma ferramenta importante no combate às

doenças transmitidas por arbovírus, pois contêm informações espaciais de interesse epidemiológico que poderão direcionar de forma mais eficaz os recursos humanos e financeiros, especialmente em países mais vulneráveis.

Objetivos

O objetivo da presente proposta é desenvolver modelos preditivos usando aprendizado de máquina para identificar áreas potenciais de novos criadouros. São objetivos específicos:

- Coletar dados sobre áreas com histórico de residentes contaminados;
- Explorar diferentes algoritmos de aprendizado de máquina e treinar modelos com dados coletados;
- Utilizar os modelos treinados para prever áreas com alto risco de criadouros, gerando mapas ou relatórios.

Método Científico

O presente estudo utilizará a linguagem python e será realizado seguindo quatro etapas. A primeira etapa do projeto envolve a coleta de dados históricos. Serão utilizados dados epidemiológicos obtidos de sistemas de notificação de saúde, dados climáticos e socioeconômicos e demográficos obtidos utilizando a biblioteca requests.

A segunda etapa envolve o pré-processamento de dados, que utilizará as bibliotecas pandas, numpy e scikit-learn, visando tratar os valores faltantes, remover outliers, normalizar os dados e unificar dados de diferentes fontes.

Na terceira etapa, os dados serão separados em conjuntos de treino e teste e serão avaliados diferentes algoritmos de aprendizado de máquina utilizando a biblioteca Pycaret.

Por fim, a última etapa envolve a predição e geração de mapas ou dashboards. Nessa etapa serão utilizadas as bibliotecas geopandas, folium, streamlit e reportlab.

Referências

LIMA, Clarisse Lins de, et al. Temporal and spatiotemporal arboviruses forecasting by machine learning: a systematic review. *Frontiers in Public Health*, v. 10, p. 900077, 2022.

MACCORMACK-GELLES, Benjamin et al. Evaluation of the usefulness of Aedes aegypti rapid larval surveys to anticipate seasonal dengue transmission between 2012–2015 in Fortaleza, Brazil. *Acta tropica*, v. 205, p. 105391, 2020.

PAUL, Kishor Kumar et al. Risk factors for the presence of dengue vector mosquitoes, and determinants of their prevalence and larval site selection in Dhaka, Bangladesh. *PloS one*, v. 13, n. 6, p. e0199457, 2018.

RIBEIRO, Mário Sérgio et al. Índices larvais de Aedes aegypti e incidência de dengue: um estudo ecológico no Estado do Rio de Janeiro, Brasil. *Cadernos de saúde pública*, v. 37, p. e00263320, 2021.

SILVA, C. C. D. et al. Forecasting Dengue, Chikungunya and Zika cases in Recife, Brazil: a spatio-temporal approach based on climate conditions, health notifications and machine learning. *Research, Society and Development*, v. 10, n. 12, 2021.

MEMBROS DO PROJETO

Nome	Categoria	CH Dedicada	Função
EDUARDO GOMES CARVALHO	DOCENTE	Não informada	Coordenador(a)
LAZARO EDUARDO DA SILVA	DOCENTE	Não informada	Coordenador(a) Adjunto(a)

CRONOGRAMA DE ATIVIDADES

Atividade	2024				2025							
	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago
COLETA DE DADOS HISTÓRICOS												
PRÉ-PROCESSAMENTO DE DADOS												
SELEÇÃO E TREINAMENTO DE MODELOS												
PREDIÇÃO E VISUALIZAÇÃO												
PLANOS DE TRABALHO												

Título	Tipo da Bolsa	Situação	
Plano Bolsista 01	PIBITI - CNPq (IT)	APROVADO	
Plano Voluntário 01	VOLUNTÁRIO (IC)	APROVADO	