

VALERIA: um aplicativo para auxiliar no diagnóstico diferencial de arboviroses

Sebastião Rogerio da Silva Neto*, Thomás Tabosa de Oliveira*, Igor Vitor Teixeira*, Élisson da Silva Rocha*, Kayo Henrique de Carvalho Monteiro*, †Vanderson de Souza Sampaio, Patricia Takako Endo*

srsn@ecom.poli.br, tto@ecom.poli.br, ivt@ecom.poli.br, esr2@ecom.poli.br, khcm@ecom.poli.br
vandersons@gmail.com, patricia.endo@upe.br

*Universidade de Pernambuco – Recife – Pernambuco – Brasil

†Instituto Todos pela Saúde – São Paulo – São Paulo – Brasil

Abstract

Brazil is one of the countries with the highest incidence of Neglected Tropical Diseases (NTDs), especially in the North and Northwest regions. Arboviruses, such as Dengue and Chikungunya, transmitted by mosquitoes, are the most common in the country. Arbovirus infection can cause persistent symptoms and negatively impact patients' quality of life, resulting in economic challenges for public health. Accurate diagnoses are essential, but the financial limitation for large-scale laboratory testing is a barrier. In this context, VALERIA is presented as a solution through the use of machine learning models to assist in the classification of arboviruses, relying solely on patients' clinical information, offering targeted treatments, and promoting positive social impact in the Brazilian territory.

Keywords: Neglected Tropical Diseases, Quality of life, Clinical information, Machine learning

1 INTRODUÇÃO

As Doenças Tropicais Negligenciadas (DTNs) são um conjunto de doenças que apresentam maior prevalência em países tropicais e em desenvolvimento, estando associadas a fatores como pobreza e baixo Índice de Desenvolvimento Humano (IDH) nas regiões [2, 9, 12]. O Brasil é um dos países com maior incidência de DTNs, sendo as regiões Norte e Noroeste as mais afetadas [10]. As arboviroses, que são parte das DTNs, são as mais comuns no país, incluindo Dengue, Chikungunya e Zika, transmitidas pelos mosquitos *Aedes aegypti* e *Aedes albopictus*. De acordo com a Organização Pan-Americana da Saúde (OPAS) [11], somente no ano de 2022, o Brasil notificou cerca de 2.363.490 casos de Dengue, o que corresponde a aproximadamente 84% dos casos notificados em todas as Américas.

A infecção por arbovírus pode causar sintomas persistentes por meses e até mesmo anos, como a artralgia associada à Chikungunya, impactando negativamente na qualidade de vida dos pacientes e, com consequências econômicas significativas devido à redução da produtividade dos pacientes [1]. A presença simultânea de diferentes arbovírus no Brasil demanda a implementação de novos métodos para o diagnóstico diferencial das arboviroses no âmbito da vigilância epidemiológica. No entanto, as restrições financeiras resultantes dos custos envolvidos na realização de exames laboratoriais de confirmação em larga escala representam um desafio significativo na busca para garantir a precisão das notificações de casos dessas doenças.

O método de diagnóstico mais comumente utilizado é baseado no quadro epidemiológico da região. Embora existam testes rápidos disponíveis, sua eficácia é limitada apenas aos três primeiros dias após o surgimento dos sinais e sintomas. Passado esse período, são necessários testes laboratoriais e sorológicos para confirmar o diagnóstico clínico dado pelo profissional de saúde no momento do atendimento. Porém, alguns testes sorológicos podem apresentar reação cruzada com outras doenças [8] e os testes laboratoriais necessitam de equipamentos específicos que, além de demandarem custo adicional (de implementação e operacional), podem não estar disponíveis em diversas regiões do Brasil.

Modelos de aprendizado de máquina se apresentam como uma solução acessível e de baixo custo de implementação que podem auxiliar profissionais da saúde a realizarem um diagnóstico clínico de arboviroses mais preciso no momento do atendimento do paciente. Os modelos de aprendizado de máquina são capazes de reconhecer padrões a partir de uma base de dados, o que torna possível utilizá-los apenas com dados clínicos e sociodemográficos (sem a necessidade de exames complementares). Além disso, após a fase de treinamento, esses modelos podem realizar a classificação requerendo baixo poder computacional, podendo ser implementados em tablets, celulares ou sistemas web, por exemplo.

Com base nessa perspectiva, apresentamos a VALERIA (acrônimo para Assistente Virtual para Processos de Aprendizagem com Arboviroses, do inglês *Virtual Assistant for Learning pRocesses In Arbovirus*), um aplicativo que utiliza

In: XXII Workshop de Ferramentas e Aplicações (WFA 2023), Ribeirão Preto, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2023.
© 2023 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

um modelo de aprendizado de máquina para auxiliar na classificação de Dengue e Chikungunya, utilizando somente informações clínicas do paciente, como sinais, sintomas, considerando também comorbidades prévias e a quantos dias o paciente está sentindo os sinais e sintomas. Nossa solução está relacionada ao ODS 3 (Objetivo de Desenvolvimento Sustentável 3) da Organização das Nações Unidas (ONU), que busca assegurar uma vida saudável e promover o bem-estar para todas as idades, incluindo o combate a doenças infecciosas como as arboviroses.

Com as restrições financeiras resultantes dos custos envolvidos na realização de exames laboratoriais de confirmação em larga escala representam um desafio significativo na busca para garantir a precisão das notificações de casos dessas doenças. A VALERIA pode auxiliar diretamente o profissional de saúde para um diagnóstico mais assertivo. Além disso, seus benefícios ultrapassam o aspecto clínico, pois a VALERIA possui um grande potencial para gerar políticas públicas de impacto social em todo território brasileiro, especialmente através do Sistema Único de Saúde (SUS).

2 A VALERIA

A VALERIA¹ é um aplicativo inovador de suporte à decisão que utiliza um modelo de aprendizado de máquina para auxiliar profissionais da saúde no diagnóstico diferencial, inicialmente, das seguintes arboviroses: Dengue e Chikungunya. Sua primeira proposta foi apresentada no trabalho de Silva Neto et al. [3] e seu primeiro protótipo apresentado por Tabosa et al. [7]. O aplicativo conta com um modelo que foi desenvolvido a partir de uma metodologia científica rigorosa, incluindo *benchmarking* de vários modelos de aprendizado de máquina, técnicas de seleção de atributos e otimização de hiperparâmetros para realizar a classificação das arboviroses [13], além de apresentar resultados de forma mais compreensível para o profissional de saúde. A Figura 1 apresenta a tela inicial da VALERIA, onde encontram-se uma descrição sobre o aplicativo, a equipe e os parceiros envolvidos no projeto, e a lista de publicações. A VALERIA é de uso exclusivo do profissional de saúde, onde primeiramente deve realizar um cadastro e ser autenticado no sistema².



Figure 1. Página inicial da VALERIA

¹<https://valeria.dotlabbrasil.com.br/>

²O cadastro foi temporariamente desabilitado enquanto a VALERIA está em fase de validação

Na área de auxílio ao diagnóstico (Figura 2), o usuário pode preencher um formulário, com as informações necessárias para realização do diagnóstico, utilizando um conjunto de onze sinais e sintomas (febre, mialgia, cefaleia, exantema, náusea, dor nas costas, conjutivite, artrite, artralgia, petéquias, dor retroorbital), duas comorbidades prévias (diabetes e hipertensão) e a quantidade de dias que o paciente apresenta os sintomas. A partir dessas informações, o modelo de aprendizado de máquina realiza a classificação, informando as probabilidades do paciente possuir alguma das arboviroses: Dengue, Chikungunya, ou se o resultado é inconclusivo. Para cada uma dessas possíveis saídas é apresentada a taxa de confiança do modelo. É válido ressaltar que os sinais, sintomas e as comorbidades foram selecionadas através de uma extensa bateria de experimentos e com auxílio de colaboradores especialistas em arboviroses.

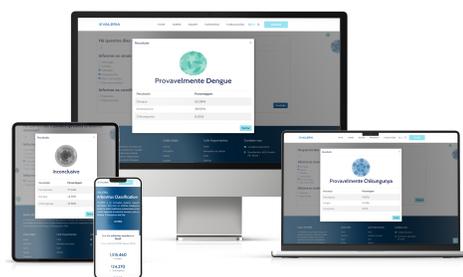


Figure 2. Auxílio ao diagnóstico de arboviroses

O aplicativo apresenta-se como grande possibilidade de inovação para SUS, uma vez que pode proporcionar agilidade no diagnóstico, integração com o sistema legado e melhor utilização dos recursos disponíveis. Essa ferramenta tecnológica pode fortalecer a capacidade de resposta do SUS e contribuir para o controle e prevenção das arboviroses no Brasil.

2.1 Principais funcionalidades da solução

A VALERIA possui três módulos principais: auxílio diagnóstico diferencial, resultado explicável e retreino (Figura 3). O módulo de auxílio ao diagnóstico diferencial oferece uma interface gráfica acessível por computador, smartphone ou tablet, permitindo aos profissionais de saúde inserirem informações do paciente para obter o diagnóstico mais provável de Dengue, Chikungunya ou um resultado inconclusivo. O resultado apresentado é interpretado pelo módulo de resultado explicável. Por fim, módulo de retreino realiza treinamentos e avaliações contínuas dos modelos, utilizando periodicamente dados coletados em bancos de dados públicos.

Módulo de auxílio ao diagnóstico diferencial: O módulo de auxílio ao diagnóstico diferencial é projetado com uma interface gráfica e contar com autenticação do profissional de saúde, exigindo dados de login e senha para garantir o acesso seguro a VALERIA. Esse modelo é capaz de calcular

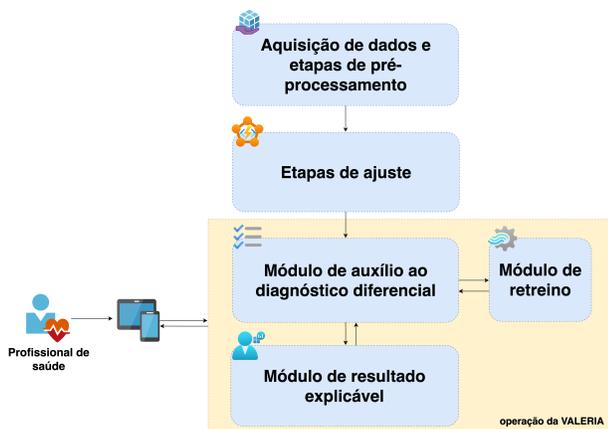


Figure 3. Funcionalidades do aplicativo VALERIA.

a probabilidade de o paciente estar infectado com um determinado arbovírus. Dessa forma, o sistema oferece uma avaliação mais precisa e rápida, auxiliando o profissional no processo de tomada de decisão para o diagnóstico clínico.

Módulo de resultado explicável: Este módulo complementa o módulo de auxílio ao diagnóstico diferencial ao detalhar o resultado obtido. Utilizando técnicas de interpretabilidade, o sistema explica de forma clara e compreensível os principais atributos que influenciaram o modelo em suas conclusões. Essa abordagem é fundamental para permitir que o profissional de saúde tome decisões mais transparentes e assertivas ao avaliar a classificação fornecida. Além de fornecer informações mais detalhadas sobre o diagnóstico, o sistema tem o potencial de sugerir condutas terapêuticas baseadas em diretrizes clínicas específicas para arboviroses.

Módulo de retreino O módulo de retreino desempenha um papel crucial no aprimoramento contínuo dos modelos de aprendizado de máquina, como o próprio nome sugere. À medida que novos dados são incorporados, possibilita o refinamento do modelo em produção, gerando resultados mais precisos e atualizados. A aplicação dos princípios de MLOps tem se mostrado essencial em diversos setores críticos da sociedade, é fundamental adotar esses princípios no treinamento e avaliação contínua dos modelos de aprendizado de máquina. Com base nessa abordagem, nossa proposta é treinar e avaliar os modelos de forma contínua, utilizando uma janela fixa de casos diagnosticados clinicamente. Para isso, novos casos de Dengue e Chikungunya seriam regularmente coletados de bancos de dados públicos, permitindo aprimorar continuamente o modelo antes de colocá-lo novamente em produção.

Essa abordagem contínua de treinamento e aperfeiçoamento do modelo com dados atualizados garantirá que o sistema esteja sempre alinhado com as mudanças e tendências na área de arboviroses. Ao adotar essa prática, estaremos assegurando a eficácia e a confiabilidade do aplicativo VALERIA ao longo do tempo, contribuindo para diagnósticos

mais precisos e, consequentemente, para um tratamento mais eficiente e seguro dos pacientes.

2.2 Diferenciais da solução frente a outras existentes no mercado

Após uma extensa revisão da literatura [6], constatamos que não há ferramentas semelhantes que ofereçam suporte específico para o diagnóstico clínico de arboviroses, como Dengue e Chikungunya. A VALERIA se destaca por oferecer funcionalidades de explicabilidade dos resultados e retreino de modelo, que não são comumente encontradas em outras ferramentas similares. A capacidade de explicar os resultados obtidos, destacando os sinais, sintomas e comorbidades prévias que influenciaram no diagnóstico, oferece uma compreensão mais completa aos profissionais de saúde. O módulo de retreino permite que a solução se mantenha atualizada, incorporando novos registros para treinamento e aprimorando continuamente o modelo de aprendizado de máquina utilizado. Esses diferenciais posicionam a VALERIA como uma solução inovadora e abrangente para o diagnóstico clínico de arboviroses a ser implementada no SUS.

3 RESULTADOS OBTIDOS E ESPERADOS

3.1 Das instituições e parceiros envolvidos

Este projeto está sendo desenvolvido por pesquisadores do Programa de Pós-Graduação em Engenharia da Computação (PPGEC) da Universidade de Pernambuco (UPE) em parceria com pesquisadores do Programa de Pós-Graduação em Medicina Tropical da Universidade do Estado do Amazonas (UEA), da Fundação de Medicina Tropical Dr. Heitor Vieira Dourado (FMT-HVD), da Fundação de Vigilância em Saúde Dra. Rosemary Costa Pinto (FVS-RCP/AM), e do Instituto Todos pela Saúde (ITpS). Vale ressaltar a importância destas parcerias para o desenvolvimento do projeto, considerando a interdisciplinaridade do problema em questão.

3.2 Da relevância e impacto social

Com potencial para influenciar políticas públicas de impacto social, a VALERIA assume relevância especial devido ao papel fundamental do SUS na garantia do direito à saúde dos brasileiros, atendendo milhões de pessoas e sendo um pilar essencial para a construção de uma sociedade mais saudável e equitativa. Ao direcionar adequadamente os tratamentos com base nos diagnósticos, a VALERIA também contribui significativamente para melhorar a qualidade de vida dos pacientes. Em um próximo *spin-off*, planeja-se ter a VALERIA de modo *offline*, e neste caso, os profissionais de saúde poderão utilizar nossa solução em áreas remotas e afastadas dos grandes centros urbanos. Essa expansão tornará possível levar um diagnóstico clínico mais preciso a populações vulneráveis, promovendo assim a inclusão e equidade no acesso aos serviços de saúde.

3.3 Da executabilidade do projeto

A VALERIA possui um modelo de classificação multiclasse que considera três classes: Dengue, Chikungunya e resultado inconclusivo, e utiliza dados clínicos dos pacientes. Esse modelo foi registrado no Instituto Nacional da Propriedade Industrial (INPI)³ com o número de registro BR512021002710-8.

Este projeto possui recursos advindo dos seguintes editais:

- Edital CNPq/AWS 032/2019 - Acesso às Plataformas de Computação em Nuvem da AWS (Cloud Credits for Research), no valor de USD \$ 11.226 (em créditos em serviço de computação em nuvem da AWS);
- Edital 006/2019 referente ao Programa de Apoio à Pesquisa - Universal Amazonas da FAPEAM, no valor de R\$ 62.212,00;
- Edital de Apoio a Pós-Graduação Stricto Sensu da UPE 2020 - Modalidade Auxílio para Projetos de Pesquisa (APQ), no valor de R\$ 5.000,00;
- Edital 20/2019 referente a Concessão de Bolsas de Pós-Graduação Stricto Sensu da FACEPE - Bolsa de Mestrado e de Doutorado;
- Edital CNPq/AWS 064/2022 - Acesso às Plataformas de Computação em Nuvem da AWS (Cloud Credits for Research), no valor de USD \$ 99.110,30 (em créditos em serviço de computação em nuvem da AWS).

3.4 Dos resultados já alcançados e esperados, potencial e/ou cronograma para execução ou continuidade do projeto

Os resultados obtidos até o momento demonstram uma sensibilidade em torno de 80% dos casos, o que evidencia a eficácia da VALERIA como uma ferramenta suporte a decisão clínica. Diversos artigos sobre a VALERIA já foram publicados em revistas e congressos científicos, destacando: Silva Neto et al. [3–6] e Tabosa et al [7, 13].

O cronograma a seguir apresenta resumidamente as próximas atividades do desenvolvimento e validação da VALERIA: 2023: (i) Criação da versão *offline* da VALERIA; (ii) Validação do módulo de resultado explicável. 2024: (i) Implementação e testes do módulo de retreino; (ii) Teste e validação em ambiente real.

4 Agradecimentos

Este trabalho foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) e pela Universidade de Pernambuco (UPE).

References

- [1] Antoine Bertolotti, Marême Thioune, Sylvie Abel, Gilda Belrose, Isabelle Calmont, Raymond Césaire, Minerva Cervantes, Laurence Fagour, Émilie Javelle, Catherine Lebris, et al. 2020. Prevalence of chronic chikungunya and associated risks factors in the French West Indies (La Martinique): A prospective cohort study. *PLoS neglected tropical diseases* 14, 3 (2020), e0007327.
- [2] Juliana Rodrigues da Conceição, Cecília Gatti Petrilli Lopes, Elizabeth Igne Ferreira, Sabrina Epiphany, and Jeanine Giarolla. 2022. Neglected tropical diseases and systemic racism especially in Brazil: from socio-economic aspects to the development of new drugs. *Acta Tropica* (2022), 106654.
- [3] Sebastião Rogério da Silva Neto, Thomás Tabosa de Oliveira, Vanderson de Souza Sampaio, Theo Lynn, and Patricia Takako Endo. 2020. Platform for monitoring and clinical diagnosis of arboviruses using computational models. In *2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*. IEEE, 1–3.
- [4] Sebastião Rogério Da Silva Neto, Thomás Tabosa, Leonides Medeiros Neto, Igor Vitor Teixeira, Sara Sadok, Vanderson De Souza Sampaio, and Patricia Takako Endo. 2023. Binary Models for Arboviruses Classification Using Machine Learning: A Benchmarking Evaluation. (2023).
- [5] Sebastião Rogério da Silva Neto, Thomás Tabosa de Oliveira, Igor Vitor Teixeira, Leonides Medeiros Neto, Vanderson Souza Sampaio, Theo Lynn, and Patricia Takako Endo. 2022. Arboviral disease record data-dengue and chikungunya, brazil, 2013–2020. *Scientific Data* 9, 1 (2022), 198.
- [6] Sebastião Rogério da Silva Neto, Thomás Tabosa Oliveira, Igor Vitor Teixeira, Samuel Benjamin Aguiar de Oliveira, Vanderson Souza Sampaio, Theo Lynn, and Patricia Takako Endo. 2022. Machine learning and deep learning techniques to support clinical diagnosis of arboviral diseases: A systematic review. *PLOS Neglected Tropical Diseases* 16, 1 (2022), e0010061.
- [7] Thomás Tabosa de Oliveira, Sebastião Rogério da Silva Neto, Igor Vitor Teixeira, Patricia Takako Endo, and Vanderson Souza Sampaio. 2021. VALERIA: Uma Plataforma para Auxiliar o Diagnóstico e o Monitoramento de Arbovíroses. In *Anais Estendidos do XXVII Simpósio Brasileiro de Sistemas Multimídia e Web*. SBC, 103–106.
- [8] Tamara Nunes Lima-Camara. 2016. Arbovíroses emergentes e novos desafios para a saúde pública no Brasil. *Revista de Saúde Pública* 50 (2016).
- [9] José Angelo L Lindoso and Ana Angélica BP Lindoso. 2009. Neglected tropical diseases in Brazil. *Revista do Instituto de Medicina Tropical de São Paulo* 51 (2009), 247–253.
- [10] Francisco Rogerlândio Martins-Melo, Alberto Novaes Ramos Jr, Carlos Henrique Alencar, and Jorg Heukelbach. 2016. Mortality from neglected tropical diseases in Brazil, 2000–2011. *Bulletin of the World Health Organization* 94, 2 (2016), 103.
- [11] OPAS. 2022. Casos Reportados de Dengue en las Américas. <https://www3.paho.org/data/index.php/es/temas/indicadores-dengue/dengue-nacional/9-dengue-pais-ano.html>
- [12] World Health Organization et al. 2017. *Towards universal coverage for preventive chemotherapy for neglected tropical diseases: guidance for assessing “who is being left behind and why”: working draft for further piloting during 2018-2019*. Technical Report. World Health Organization.
- [13] Thomás Tabosa de Oliveira, Sebastião Rogério da Silva Neto, Igor Vitor Teixeira, Samuel Benjamin Aguiar de Oliveira, Maria Gabriela de Almeida Rodrigues, Vanderson Souza Sampaio, and Patricia Takako Endo. 2022. A Comparative Study of Machine Learning Techniques for Multi-Class Classification of Arboviral Diseases. *Frontiers in Tropical Diseases* 2 (2022), 769968.

³<https://www.gov.br/inpi/pt-br>