

Monitoramento de epidemia de dengue na Amazônia usando Redes Neurais Artificiais

Wilson Rogério Soares e Silva, Carlos Renato Lisboa Francês

LTS-Laboratório de Tecnologias Sociais / PPGEE-Programa de Pós-Graduação em Engenharia Elétrica / ITEC- Instituto de Tecnologia/ UFPA-Universidade Federal do Pará.

Av. Augusto Corrêa, 01 Prédio do Laboratório de Eng. Elétrica, sala 26 Belém, PA, Brasil. CEP: 66050-000

rogerio.soares@ifpa.edu.br, rfrances@ufpa.br
<http://www.lprad.ufpa.br>

Resumo A dengue é uma doença infecciosa viral que afeta mais de 100 países no mundo. Segundo dados disponibilizados pelo Ministério da Saúde (2010), dos 409.073 casos notificados na região Norte, 106.433 ocorreram no Estado do Pará. Esse trabalho propõe um estudo de Redes Neurais Artificiais para prever novos casos de dengue na Amazônia. Para tanto, desenvolveu-se um sistema que usa base de dados públicos de casos da doença de ocorrência semanal na região metropolitana de Belém. E implementa um módulo de emissão de alertas, visando a detecção de um aumento repentino de novos casos da doença, contribuindo para tomada de decisão dos órgãos de saúde pública e suas respectivas ações de controle das epidemias nos municípios. Os resultados demonstraram que o modelo de Redes Neurais Artificiais, para o cenário em estudo, obteve um bom desempenho na previsão epidemiológica, alcançando acurácia satisfatória.

Keywords: Previsão, Dengue, RNA, Amazônia

1 Introdução

O aumento alarmante dos casos de dengue, zica e chikungunya nos últimos anos vem preocupando diversos países em todo o mundo. A dengue é uma doença infecciosa viral que afeta mais de 100 países no mundo. Dados da Organização Mundial da Saúde (OMS) mostram mais de 50 milhões de indivíduos infectados anualmente, com 500 mil casos de dengue hemorrágica e 20 mil mortos [1]. Ainda de acordo com a OMS, cerca de 2,5 bilhões de pessoas (aproximadamente metade da população mundial) estão atualmente em risco de contrair dengue.

Todos os anos, estima-se que cerca de 50 milhões de pessoas sofrem com sintomas do vírus da dengue, com aproximadamente 55 mil indivíduos que necessitam de hospitalização [5]. A OMS ao reconhecer os perigos dessa doença,

incluiu a dengue em sua campanha de 2014 para o controle de doenças transmitidas por vetores. Dos 96 milhões de casos estimados globalmente de dengue por ano, 70% ocorrem na Ásia, sendo que a Índia contribui em 34% de todos os casos [6] (Figura 01).

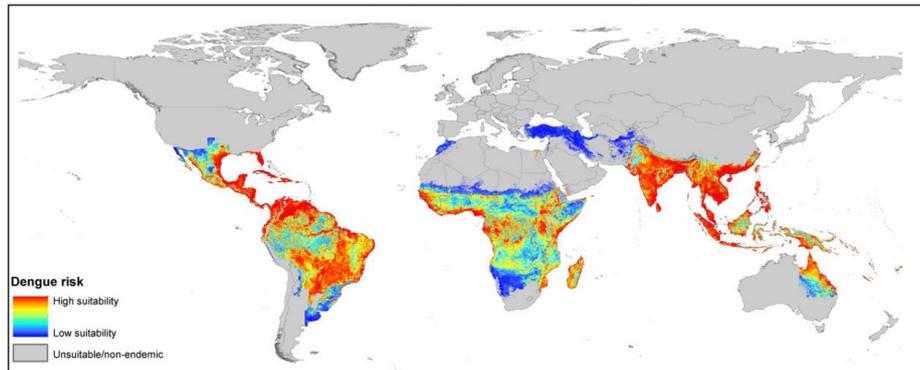


Figura 1. Distribuição global de dengue no mundo. Fonte: [16]

Segundo [7], o Brasil possui um dos mais abrangentes sistemas de vigilância de dengue, com mais de 200 milhões de pessoas sob observação e dados mensalmente ao longo de dez anos ou mais em 5570 municípios. O país já dedica recursos significativos para o controle de vetores, incluindo os transmissores da dengue, além de ter uma comunidade de pesquisadores dedicada a estudar o avanço de epidemias, suas causas e consequências.

O Ministério da Saúde (2010) afirma que dos 409.073 casos notificados na região Norte no período de 2000 a 2010, 106.433 ocorreram no Estado do Pará, o restante distribuído nos demais seis Estados, ou seja, cerca de 20% em um dos sete estados existentes. Os primeiros casos notificados de dengue ocorreram em 1995, nos municípios de Redenção e Rondon do Pará, região Sudeste do estado. Desde então, a dengue permanece presente na maioria dos municípios, atualmente com 32 municípios prioritários para o Programa Nacional de Controle da Dengue, dentre os quais estão Altamira, Belém, Cametá, Conceição do Araguaia, Itaituba, Marabá, Monte Alegre, Santarém, Soure e Tucuruí[3].

Diante do contexto exposto, é imperioso que tomadores de decisão no âmbito de políticas públicas possuam minimamente dados e mecanismos de apoio à eficiência de ações em saúde pública em geral e, em particular, em doenças epidêmicas como a dengue, zica e chicungunya.

Assim esse trabalho propõe um estudo de Redes Neurais Artificiais (RNAs) para prever casos de dengue na Amazônia. É realizado um estudo, que usa bases de dados públicos do SINAN (Sistema de Informação de Agravos de Notificação) de casos semanais de dengue na região metropolitana de Belém. Posteriormente, foram submetidos a modelo ANN para predição de novos casos epidemiológicos

e emissão de alertas, que poderão ser usados por órgãos competentes de saúde na tomada de decisão e ações de controle das epidemias em seus municípios. A partir da análise de dados, a pesquisa elaborada se desenvolve em 6 seções. A seção 2 apresenta os trabalhos correlatos. A seção 3 apresenta os Materiais e métodos. A seção 4 Resultados. A seção 5 a Discussão e a seção 6 Considerações finais.

2 Trabalhos correlatos

A detecção precoce de surtos da doença e ações de controle rápido são essenciais para prevenir e conter a propagação de doenças infecciosas e reduzir a morbidade. Assim senso, diversas abordagens são encontradas na bibliografia focadas na propagação de diferentes doenças.

Ao longo das últimas décadas, numerosos métodos de previsão foram propostos no campo da previsão epidemiológica. Tais métodos podem ser classificados em diferentes categorias, como deterministas *versus* probabilísticos, métodos comparativos *versus* métodos generativos [20].

Conforme o trabalho de [21], os casos diários de dengue nas 457 aldeias urbanas da cidade de Kaohsiung, Taiwan, de 2009 a 2012, foram utilizados para desenvolvimento e avaliação de modelos de previsão. Esta pesquisa tem como objetivo prever os casos confirmados por dengue usando Redes Neurais Artificiais (RNAs). Houve no total 2,997 casos confirmados de dengue durante esse período, um modelo de regressão foi ajustado aos incidentes diários ocorridos nas aldeias nos últimos 30 dias. O modelo foi usado para prever as probabilidades de incidência do surto de dengue para as aldeias no dia seguinte.

A pesquisa de [23] tem como objetivo prever os casos confirmados por dengue usando Redes Neurais Artificiais (RNAs). Em seu estudo usaram dados fornecidos pela Agência Nacional de Meio Ambiente de Singapura (NEA), foram utilizados para modelar o comportamento dos casos de dengue com base nos parâmetros físicos da temperatura, da umidade relativa e da precipitação total. Os resultados mostraram que as quatro características importantes, como a temperatura média, a umidade relativa média, a precipitação total e o número total de casos confirmados por dengue foram muito efetivas na predição do número de casos confirmados por dengue.

Nesse contexto o trabalho de [22] tem como objetivo auxiliar a Secretaria de Saúde do Estado do Rio de Janeiro a reduzir os riscos de epidemia de dengue no Estado. O aumento dos casos epidemiológicos vêm exigindo atenção dos gestores públicos de saúde, seja em ações, em serviços de saúde para a sua prevenção e/ou no seu controle.

Por conta da magnitude desse problema, é crucial adotar medidas que impeçam e/ou se antecipem os surtos da doença. Neste contexto, este trabalho apresenta mais uma ferramenta como alternativa de apoio à solução do problema

3 Materiais e métodos

3.1 Área de estudo

O estudo foi realizado na cidade de Belém, capital do estado do Pará/Brasil, que é a segunda cidade mais populosa da região Norte. Com área territorial 1.059,458 km², população estimada em 1.446.042 pessoas, densidade populacional de 1.315,26 km². Ocupando a 22^o posição entre os IDH por Capital. Localizada ao Nordeste da maior floresta tropical do mundo, Belém é altamente afetado por chuvas, devido seu clima equatorial influenciada pela Amazônia que contribui para o aumento de casos da doença. Segundo o Ministério da Saúde, Belém, no período de 2007 e 2012 manteve casos registrados de dengue superiores a 1500 casos por ano [13] [14].

3.2 Dados coletados

Esse trabalho utilizou técnicas de pré-processamento para escolha dos dados estudados. Essa etapa possui fundamental importância no processo de descoberta de conhecimento permitindo organizar as funções de pré-processamento em: extração de dados, seleção de atributos, limpeza dos dados, discretização, transformação de variáveis, etc [24].

A base de dados usada foi proveniente de dados públicos disponível no Portal SINAN (Sistema de Informação de Agravos de Notificação) [15] que um órgão responsável por prover informações relacionados a saúde pública, vinculado ao Ministério da Saúde. A escolha da base de dados se justifica por conter uma série histórica de casos confirmados de dengue em todos os municípios do território nacional, mostrando informações de casos confirmados de dengue em todas as cidades do estado do Pará.

Na primeira etapa os dados foram extraídos do SINAM, usando o algoritmo python apresentado em [25], que permite selecionar e visualizar os dados pertinentes para estudo. Essa extração gerou um dataset contendo a série histórica de notificação de casos de dengue de cada cidade do estado do Pará.

Posteriormente, na segunda etapa foram selecionados os dados correspondentes a região metropolitana de Belém-Pará. A série histórica registra dados ao longo de 144 instâncias, referente as semanas de abril de 2013 a dez de 2015, com um total de 2.421 casos de dengue notificados nesse período Figura 2.

Na etapa final, foi analisado se os atributos selecionados continham informações ausentes, errôneas ou inconsistentes, o que poderia comprometer a qualidade do modelo computacional usado na previsão.

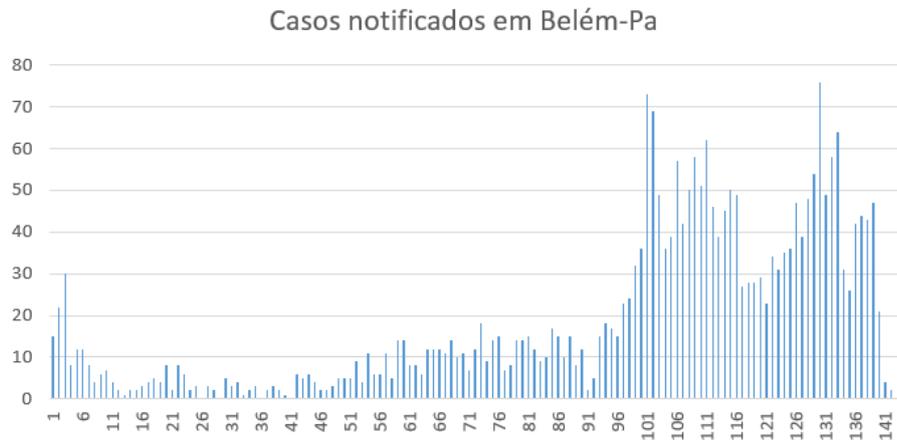


Figura 2. Casos notificados de dengue na região Metropolitana de Belém-pa, abril de 2013 a dez de 2015. Com total de 144 instâncias. Fonte [15]

3.3 Modelo da Rede Neural Artificial

As Redes Neurais Artificiais (RNA) são um modelo matemático ou computacional que tenta simular a estrutura e os aspectos funcionais de uma rede neural biológica. Consiste num grupo interconectado de neurônios artificiais. A arquitetura Perceptron Multicamadas (PMC) é a mais famosa RNA. Nela os neurônios são agrupados em camadas (uma camada de entrada, uma de saída e uma ou mais camadas escondidas), e existem apenas conexões para frente (feedforward) [17] [11].

O modelo de ANN usada nesse trabalho é do tipo Multi Layer Perceptron (MLP), usa como topologia uma camada de entrada, uma camada escondida, com 5 neurônios e uma camada saída. Para função de ativação foi usada a tangente hiperbólica e para função de treinamento foi utilizado Backpropagation Levenberg-Marquardt. As entradas da ANN foram normalizadas com valores máximos e mínimos, assumindo entre 0 a 1. figura 3.

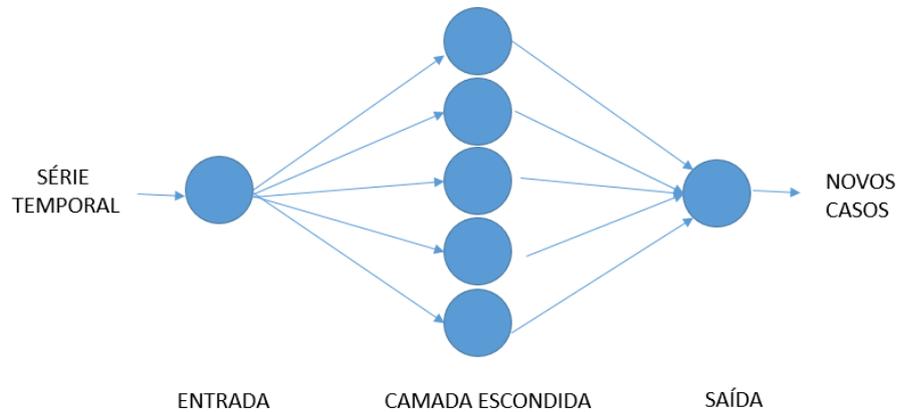


Figura 3. Modelo com 1 entrada, 05 neurônios na camada escondida e 1 saída

Dentre as diversas técnicas utilizadas para validar o desempenho obtido com RNAs, podemos citar: comparação gráfica, diagramas de dispersão, raiz do erro médio quadrático, erro percentual médio absoluto (MAPE), entre outros [4]. Utilizou-se nesse artigo o erro médio de porcentagem absoluta (MAPE) [12], para validar a capacidade de previsão do modelo, com janela de tempo igual a 5 semanas, o que permite ações mais planejadas pelos órgãos de saúde pública. O modelo com menor valor MAPE indica o melhor ajuste [8]. O MAPE é representado pela fórmula:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$$

N é o número total de semanas de dados; Y_t e \hat{Y}_t são, respectivamente, o número observado e previsto de casos de dengue em meses t [12].

3.4 Desenvolvimento do Sistema de monitoramento e alertas de epidemia de dengue

Foi desenvolvido um sistema de monitoramento e alerta contendo uma arquitetura dividida em três módulos: um módulo para importar dados públicos, um módulo que implementa o modelo de ANN e um módulo que emite alertas sobre o avanço da epidemia de dengue, informando assim os órgãos públicos de saúde responsáveis, conforme arquitetura mostrada na figura 4.

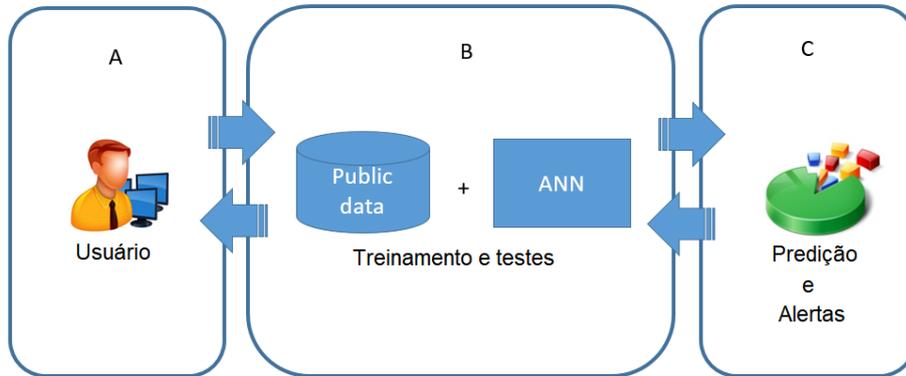


Figura 4. Arquitetura do sistema desenvolvido. Em A o módulo para importação das bases públicas. Em B o modelo ajustado de RNAs para prever novos casos. Em C os gráficos e alertas do sistema com novos avanços da doença.

Em A o usuário importa as bases públicas de diversas fontes: município, estado ou federação. Tais dados são disponibilizados em bases públicas em diversos formatos, essa arquitetura usa extensão CSV (Comma Separated Values).

Em B o módulo de previsão, usa um modelo de ANN ajustado anteriormente. Esse modelo é capaz de receber como entradas para os neurônios, o histórico de casos mensais das epidemias, aprender o comportamento da série temporal e realizar a predição de novos casos de dengue.

Em C são mostrados as previsões de novos casos. Essa previsão informa os gestores de saúde sobre o aumento de novos casos, emitindo assim, alertas sobre possíveis avanços de doenças, podendo usar tais informações para tomada de decisão e ações de controle das epidemias em seu município.

O sistema foi desenvolvido usando o Matlab R2015a [18], do tipo Gráfica User Interface (GUI). O programa foi utilizado por sua familiarização no meio acadêmico e fácil aprendizagem, visto a ser útil em alterações e melhorias futuras do sistemas.

4 Resultados

4.1 Avaliação da RNA

Foram realizados diversos experimentos com intuito de encontrar o modelo para prever novos casos de dengue. Variaram-se seus parâmetros, buscando avaliar o melhor desempenho da rede: janelas de tempo de 6 meses, números de camadas ocultas, número de neurônios das camadas intermediárias (oculta).

Todas as arquiteturas foram testadas com a mesma configuração: a função de ativação escolhida foi tangente hiperbólica, o algoritmo de treinamento usado foi backpropagation Levenberg-Marquardt e o critério de parada do algoritmo foi o tempo máximo de épocas determinadas, que nesse caso foi de 100.

Para comparar os resultados obtidos utilizou um conjunto de 32 execuções para cada RNA testada e o modelo que obteve o menor erro percentual médio absoluto (MAPE), conforme tabela a seguir.

Tabela 1. Desempenho do modelo neural: neurônios, camadas e seus MAPEs para treinamento e validação

NEURÔNIOS	CAMADAS	MAPE TREINAMENTO	MAPE VALIDAÇÃO
2	1	0.0072	0.0479
5	2	0.0031	0.0347
10	3	0.0025	0.0408
20	4	0.0093	0.1209
30	3	0.0029	0.0816
40	2	0.0072	0.0465

Na avaliação, a camada com 05 neurônios e 02 camadas, obteve o menor valor dentre os outros modelos avaliados que corresponde a erro percentual médio absoluto (MAPE) de 0.0340. O modelo convergiu para o resultado em 15 épocas, conforme figura 5. .

Com relação às janelas de tempo de previsão, a janela de médio prazo (6 meses) permitem que órgãos de saúde públicos possam realizar um planejamento de médio prazo em alguma ação epidemiológica já em curso.

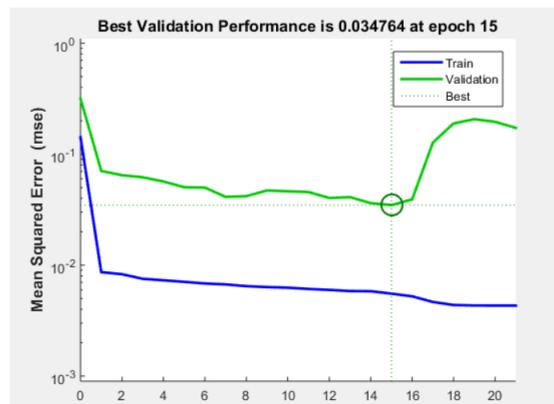


Figura 5. Resultado em 15 épocas

4.2 Predição para novos casos

Observa-se, na figura 6, a série temporal semanal referente aos casos reais (destacado na linha azul) e a prevista (na linha vermelha), foram plotados, a partir, da predição realizada pela ANN.

Foram usadas para treinamento as instâncias que correspondem ao intervalo de 0 a 120 instâncias (120 semanas e um total de 1793 casos notificados) obtendo MAPE de 0.0031.

O modelo ajustado foi então submetido a novas instâncias, na fase de testes, com intuito de prever novos casos de dengue, no período de 121 a 140 instâncias (19 semanas e um total de 848 novos casos). O modelo em questão demonstrou capacidade de previsão aceitável para novos casos, conseguindo acompanhar a curva real ao longo de toda série temporal e MAPE 0.0347, conforme figura 6.

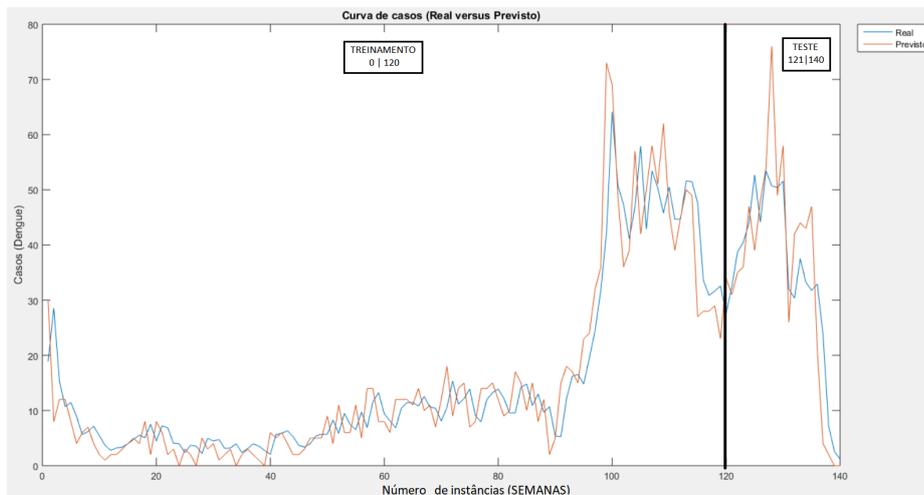


Figura 6. Predição de casos de dengue entre os anos de 2013 e 2015. A linha azul é o caso real representando as séries temporais e a linha em vermelho é a previsão do modelo de rede neural. Notam-se os números de instâncias para treinamento de 0 a 120 e as demais para testes de 121 a 140

4.3 Alerta para novos casos

Foi implementado um módulo de alerta baseado na quantidade de casos previstos semanais na fase de testes. Considerando a distribuição semanal desses novos casos, quando o número extrapolar um valor determinado, que nesse caso foi de 50 casos previstos semanais. Na figura 7 percebe-se que semanas 09, 10, 12 e 13 alcançam valores maiores que 50 casos previstos, o que dispara um alerta no sistema aos agentes públicos de saúde.

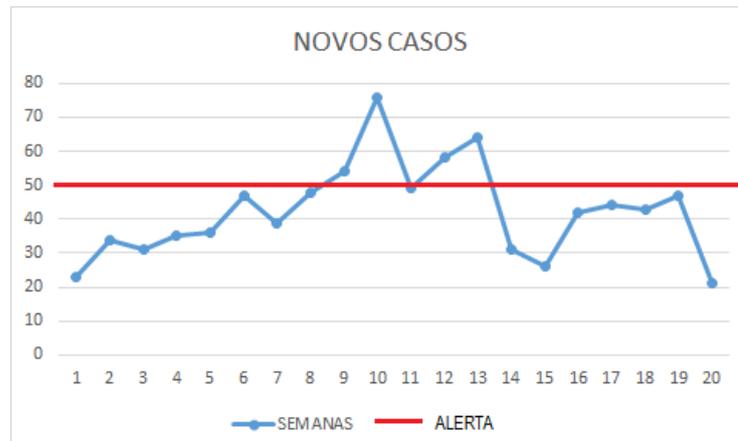


Figura 7. Visualização de novos casos de dengue previstos no sistema. Observa-se o aumento de casos nas semanas 09,10,12 e 13, sendo então emitido um alerta quando ultrapassem 50 casos semanais.

5 Discussão

Este estudo investigou uso de Redes Neurais Artificiais na previsão de epidemias de dengue. Os resultados indicaram que o modelo de rede neural, do estudo, consegue prever novos casos usando séries temporais. O modelo com 05 neurônios na camada escondida obteve a melhor desempenho de previsão para um período de 19 semanas, já o modelo com 40 neurônios na camada escondida foi o menos eficiente.

Após análises dos resultados, constatou-se que o modelo de Rede Neural Artificial, desenvolvido no estudo, é eficaz na previsão de novos casos de dengue. Na literatura, as Redes Neurais Artificiais são constantemente usadas para prever epidemias em especial a de dengue em diversos países [9]. No entanto, o estudo atual obteve melhores resultados quando comparado ao modelo preditivo desenvolvido pelo autor [9].

A previsão de epidemias usando séries temporais, desenvolvido no estudo atual é consistente com trabalhos anteriores. Como é o caso do trabalho de [10] que aponta casos diários de dengue nas 457 aldeias urbanas da cidade de Kaohsiung, Taiwan, de 2009 a 2012, totalizando com 2,997 casos confirmados, realizando previsão com cenário de curto prazo de até 30 dias. No entanto, o estudo atual aponta casos semanais, totalizando 2,641 casos confirmados, com possibilidade de prever cenários em médios prazos, de até 6 meses, o que permite ações mais planejadas dos órgãos de saúde, o que não é discutido pelo autor.

O sistema monitoramento e alertas, desenvolvido nesse estudo, é um passo chave na adição de valor aos dados epidemiológicos coletados por sistemas de vigilância. Outros trabalhos usam sistemas de monitoramento para prever surtos de dengue, em [2] implementa uma ferramenta para prever casos de dengue. No

entanto, o estudo atual consegue ter mais funcionalidades, pois é implementado um módulo de emissão de alertas, o que não é visualizado na proposta do autor [2].

6 Considerações finais

Este trabalho demonstra o estudo de Redes Neurais Artificiais (RNA) na previsão de novos casos de dengue na região metropolitana de Belém-Pa-Brasil. O estudo se mostra eficiente para previsões de novos casos, usando séries temporais, podendo prever com erros médios de 0.0347, servindo para a tomada de decisão dos gestores de saúde públicos.

Para pesquisas futuras, pretende-se comparar os resultados já obtidos, com outros modelos computacionais de aprendizagem de máquina. Da mesma forma realizar experimentos para prever novos casos em curto prazo de 2 meses e longos prazos de 12 meses. Possibilitando a Rede Neural Artificial prever novos casos em curto, médio e longos prazos.

A pesquisa realizada, bem como o sistema de monitoramento desenvolvido, mostra-se útil na detecção precoce e alerta no aumento de novos casos de dengue, que podem ser usados por órgãos de saúde pública na tomada de decisão e ações de controle das epidemias em seus municípios.

Referências

1. Santos, Charles Sodre Schramm dos: Grid Incidence of Dengue in the State From Rio de Janeiro through Artificial Neural Networks. IEEE Press, New York (2013)
2. Frank I. Eljorje, Denmar S. Clarite, Bobby D. Gerardo and Yungcheol Byun: Grid Tracking and Prediction of Dengue Outbreak Using Cloud-Based Services and Artificial Neural Network. intelligent personal assistant, (2016)
3. Correa, Jessica Ariana: Possible relationships between meteorological elements and the temporal space epidemiology of dengue and malaria in Para state. (2013)
4. Faria, Elisângela Lopes de: Prediction of the Brazilian Stock Market using Artificial Neural Networks. (2008)
5. Rasheed, S. B., Butlin, R. K., and Boots, M. (2013). A review of dengue as an emerging disease in Pakistan. *Public Health*, 127(1):11–17.
6. Mukherji, S. and Kaushik, S. K. (2015). Dengue: A runaway epidemic and a bewildered. public health worker. *Medical Journal Armed Forces India*, 71(1):3–4.
7. Brady, O. J., Smith, D. L., Scott, T. W., and Hay, S. I. (2015). Dengue disease outbreak definitions are implicitly variable. *Epidemics*, 11:92–102.
8. Saberian, F., Zamani, A., Gooya, M. M., Hemmati, P., Shoorehdeli, M. A., and Teshnehlab, M. (2014). Prediction of Seasonal Influenza Epidemics in Tehran Using Artificial Neural Networks. *IEEE, (Iccc):1921–1923*.
9. Aburas, H. M., Cetiner, B. G., and Sari, M. (2010). Dengue confirmed-cases prediction: A neural network model. *Expert Systems with Applications*, 37(6):4256–4260.
10. Chan, T.-C., Hu, T.-H., and Hwang, J.-S. (2015). Daily forecast of dengue fever incidents for urban villages in a city. *International Journal of Health Geographics*, 14(1):9.

11. Hani, M. Aburas, Gultekin, Cetiner, Murat, Sari Erratum. Dengue confirmed-cases prediction: A neural network model. Expert Systems with Applications, P. (2010). Expert Systems with Applications, Volume 38, Issue 10, 15 September 2011, Pages 13495-13496
12. Dung, Phunga, Cunrui, Huang, Shannon, Rutherforda, Cordia, Chua, Xiaoming Wangb, Minh, Nguyenb, Nga, Huy Nguyenc, Cuong, Do Manh. Identification of the prediction model for dengue incidence in Can Tho city, a Mekong Delta area in Vietnam. Acta Tropica P. (2015).
13. IBGE, <http://cidades.ibge.gov.br/xtras/perfil.php?codmun=150140>
14. Ministério da Saúde, DATASUS 2007 - 2012., http://cidades.ibge.gov.br/painel/saude.php?codmun=150140&lang=_ES
15. Sistema de Informação de Agravos de Notificação (SINAN). <http://portalsinan.saude.gov.br/sinan-dengue-chikungunya>
16. scientificmalaysian. <http://magazine.scientificmalaysian.com/issue-12-2016/dengue-vaccine-dilemma-route-prevention-yet/>.
17. Rodrigues, Leopoldo Marchiori, Mestria, Mário. Métodos de Classificação baseado em Redes Bayesianas e Neural para Reconhecimento de Atividade Humana. p. (2015). XII Congresso Brasileiro de Inteligência Computacional – CBIC 2015.
18. Mathworks, <https://www.mathworks.com/help/nnet/gs/neural-network-time-series-prediction-and-modeling.html>. acessado em 01 de março de 2017.
19. Girond, Florian and Randrianasolo, Laurence and Randriamampionona, Lea and Rakotomanana, Fanjasoa and Randrianariveolosia, Milijaona and Ratsitorahina, Maherisoa and Brou. Analysing trends and forecasting malaria epidemics in Madagascar using a sentinel surveillance network: a web-based application. p. (2015). Malaria Journal.
20. Tabataba, Farzaneh Sadat and Chakraborty, Prithwish and Ramakrishnan, Naren and Venkatramanan, Srinivasan and Chen, Jiangzhuo and Lewis, Bryan and Marathe, Madhav. A Framework for Evaluating Epidemic Forecasts.
21. Chan, Ta-Chien and Hu, Tsuey-Hwa and Hwang, Jing-Shiang. Daily forecast of dengue fever incidents for urban villages in a city. International Journal of Health Geographics.
22. Charles Sodré Schramm dos Santos. Previsão da Incidência de Dengue no Estado do Rio de Janeiro através de Redes Neurais Artificiais. Monografia de Final de Curso Especialização em Business Intelligence.
23. Hani, M. Aburas, Murati, Sara. Dengue confirmed-cases prediction: A neural network model. Expert Systems with Applications.
24. Universidade Estado Maringá, <http://www.din.uem.br/~gpea/linhas-de-pesquisa/mineracao-de-dados/pre-processamento/pre-processamento-em-data-mining/>. acessado em 01 de outubro de 2017.
25. Python, <https://www.python.org/>. acessado em 01 de agosto de 2017.