Development of a Recurrent Neural Network Model for Prediction of Dengue Importation

Sun-Young Kim, Kyung-Duk Min, Seohyun Lee, SaRang Park

Seoul National University, Seoul, Korea (the Republic of)

Objective

We aim to develop a prediction model for the number of imported cases of infectious disease by using the recurrent neural network (RNN) with the Elman algorithm [1], a type of artificial neural network (ANN) algorithm. We have targeted to predict the number of imported dengue cases in South Korea as the number of dengue cases is greater than other mosquito-borne diseases [2].

Introduction

In recent years, mosquito-borne diseases such as Zika, chikungunya, and dengue have become particularly problematic due to global climate change. Rising temperatures and changes in precipitation are considered to be associated with habitat suitability of mosquito vectors [3] and viruses [4]. To address such cross- border infectious diseases, countries have come up with various strategies to control and manage mosquito-borne diseases. In line with this, international efforts have been made to minimize the burden of global infectious diseases. In 2014, Global Health Security Agenda (GHSA) [5] has been launched in collaboration with the international organizations, member countries of GHSA, and non-governmental organizations in order to improve national and global capacities against global public health threat. In addition, various quarantine programs have been operated in and between countries borderlines and airports with cutting edge ICT technologies.

These efforts could be made more effective when the authorities have reliable predicted future trends or events [6], utilize their capacities more efficiently and provide timely alerts to the public. However, very few studies have been conducted to deal with imported disease, while much attention has been paid to the endemic diseases.

In this study, we aim to develop a prediction model for imported infectious disease by using the approach of ANN. We have chosen to model the imported cases of dengue in Korea, as the number of imported dengue cases is larger than other mosquito-borne diseases. Additionally, Japan, one of South Korea's neighboring countries, has recently experienced autochthonous dengue virus transmission, which has raised concerns about localization in Korea as well as in Japan [7].

Methods

As our prediction target was the monthly number of imported dengue cases, among the alternative types of ANN, our study used recurrent neural network (RNN) models, which has been developed to model the temporal sequenced data. Specifically, Elman algorithm [1] was used to develop an RNN and the model was implemented by an R package "RSNNS".

A conventional autoregressive integrated moving average (ARIMA) model was also developed to compare and verify the predictabilities between the RNN and conventional ARIMA modeling approach. The ARIMA model predicted the number of dengue cases that are likely to be imported from Indonesia in 2016, based on the reported number of imported dengue cases from the country between the year 2011-2015. The analysis was conducted by an R package "forecast."

To develop an RNN, the number of hidden layers and the number of nodes for each hidden layer need to be determined. The grid searching method was employed for the determination based on rooted mean squared error (RMSE), a measurement of the model performance under which a lower value indicates a better model fit. For the grid search, we chose a range of 1-3 for the number of hidden layers, and a range of 10-40 for the number of nodes for each hidden layer (29,791 combinations in total for the RNN model).

To this end, we have divided the dengue importation data into two sets, i.e., the training set versus validation set. The training data set included data which had a time period of 48 months from 2012 to 2015, and the validation set had data for over 12 months in 2016, which was the latest data set available during the time of our study. As the sequential external validation approach was

adopted, we used 12 RNNs for each data point in the validation set. In other words, we predicted the number of imported dengue cases from the given target country at the point in time in January 2016, using the RNN developed based on the data from January 2012 to December 2015. We then subsequently predicted the cases in February 2016 by using the data from 2012 to January 2016 and iterated the process similarly for predictions at other time points in 2016. Through the process, we have obtained the predicted number of imported dengue cases in 2016, and computed an RMSE by comparing the value to those in the observed data. Via the grid searching method used, a total of 29,791 RMSE values were calculated for models of each target country, and a model with the least RMSE was selected as the best-fitting model.

Results

The RMSE for the best-fitting RNN model was 14.152. In comparison to the ARIMA model, of which RMSE was 16.466, the RNN model showed improved predictability. The RNN-based prediction model can be utilized to improve the effectiveness of both national and individual level interventions for preventing the imported cases of dengue and its subsequent localization in Korea.

Conclusions

Since Dengue's enlistment in 2000 as nationally notifiable infectious disease, it has been reported to be one of the most common infectious disease imported into South Korea [8]. Concerning the rapid change in climate and disease patterns, the South Korean government should be prepared to be more responsive towards any potential imported cases of infectious disease (especially Dengue) by developing an active surveillance system and specific countermeasures.

In this context, our prediction model can be utilized to enhance the response system which is designed to reduce the number of imported cases in the future and prevent any possible localization of the disease. Our analysis also suggests areas for future research to further advance the prediction models for infectious disease importation in general. Since the current RNN-based prediction model still has its limitations, it would be crucial to put in more efforts in improving the model performance and applicability. The ANN algorithms need further development in order for it to effectively manage the availability of relevant big data in the near future.

Acknowledgement

References

1. Elman JL. 1990. Finding structure in time. *Cogn Sci*. 14(2), 179-211. https://doi.org/10.1207/s15516709cog1402_1

2. Choe YJ, Choe SA, Cho SI. 2017. Importation of travel-related infectious diseases is increasing in South Korea: An analysis of salmonellosis, shigellosis, malaria, and dengue surveillance data. *Travel Med Infect Dis*. 19, 22-27. PubMed https://doi.org/10.1016/j.tmaid.2017.09.003

3. Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, et al. 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. *eLife*. 4, e08347. PubMed https://doi.org/10.7554/eLife.08347

4. Alto BW, Bettinardi D. 2013. Temperature and dengue virus infection in mosquitoes: independent effects on the immature and adult stages. *Am J Trop Med Hyg*. 88(3), 497-505. PubMed https://doi.org/10.4269/ajtmh.12-0421

5. Arthur GF, Michael M, Leah FM, Maureen B, Mitsuaki H, et al. 2017. Contributions of the US Centers for Disease Control and Prevention in Implementing the Global Health Security Agenda in 17 Partner Countries. *Infect Dis J*. 23(13). doi:10.3201/eid2313.170898.

6. Eisen L, Eisen RJ. 2011. Using geographic information systems and decision support systems for the prediction, prevention, and control of vector-borne diseases. *Annu Rev Entomol*. 56, 41-61. PubMed https://doi.org/10.1146/annurev-ento-120709-144847

7.   Yoshimura Y, Sakamoto Y, Amano Y, Nakaharai K, Yaita K, et al. 2015. Four Cases of Autochthonous Dengue Infection in Japan and 46 Imported Cases: Characteristics of Japanese Dengue. *Intern Med*. 54(23), 3005-08. PubMed https://doi.org/10.2169/internalmedicine.54.4475

8.   Yeom J-S. 2017. Current status and outlook of mosquito-borne diseases in Korea. *J Korean Med Assoc*. 60(6), 468-74. https://doi.org/10.5124/jkma.2017.60.6.468