Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 3, June 2024, pp. 2086~2094
ISSN: 2302-9285, DOI: 10.11591/eei.v13i3.6619

A convolution neural network integrating climate variables and spatial-temporal properties to predict influenza trends

Jaroonsak Watmaha, Suwatchai Kamonsantiroj, Luepol Pipanmaekaporn
Department of Computer and Information Science, Faculty of Applied Science, King Mongkut's University of Technology
North Bangkok, Bangkok, Thailand

Article Info

Article history:

Received May 1, 2023
Revised Nov 27, 2023
Accepted Dec 7, 2023

Keywords:

Climate
Convolution neural network
Forecasting
Influenza
Time series

ABSTRACT

The spread of influenza is contingent upon a multitude of outbreak-related factors, including viral mutation, climate conditions, acquisition of immunity, crowded environments, vaccine efficacy, social gatherings, and the health and age profiles of individuals in contact with infected individuals. An epidemic in the region impacted by spatial transmission risk from adjacent regions. A few influenzas epidemic models start highlighting the spatial correlations between influenza patients and geographically adjacent regions. The proposed model is based on the concept of climatic, immunization, and spatial correlations which are represented by a convolution neural network (CNN) for influenza epidemic forecasting. This study presents an integration of three determinants for predicting influenza outbreaks, multivariate climate data, spatial data on influenza vaccination, and spatial-temporal data of historical influenza patients. The performance of three comparison models, CNN, recurrent neural network (RNN), and long short-term memory (LSTM) was compared by the root mean squared error metric (RMSE). The findings revealed that the CNN model represents human interaction at intervals of 12, 16, 20, 24, and 28 weeks resulting in the best effectiveness of the lowest RMSE=0.00376 with learning rate=0.0001.

Corresponding Author:

Jaroonsak Watmaha
Department of Computer and Information Science, Faculty of Applied Science
King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok, Thailand
Email: s5904062910023@email.kmutnb.ac.th

1. INTRODUCTION

The common cold (flu) is an inhalable disease that is easily spread due to influenza viruses. The severity of flu symptoms can vary from mild to severe, it has the potential to impact both people and animals, such as porkers, birds, and horses. According to the centers for disease control and prevention (CDC), there were 402 million cases of flu in the United States from 1997 to 2016 [1], the highest number of cases occurred in 2014 at 38 million. However, not all cases may have been reported, and some patients may even die from the flu. The 2009 and 2010 flu outbreak affected 60.8 million Americans, while the 2012 and 2013 outbreaks resulted in 710,000 hospitalizations and 56,000 deaths. Anticipated to surpass the previous epidemic in terms of casualties, the present outbreak can be curtailed through timely intervention by healthcare authorities. To safeguard oneself from the common flu, it is wise to obtain annual vaccination and avoid proximity to those who are afflicted.

Influenza vaccines establish immunity and illness Influenza prevention. Influenza vaccines: successes and continuing challenges [2] presented maternal immunization, which has an estimated efficacy of

50–60%, is the most promising strategy to protect children who are not eligible for vaccination. This paper mentioned influenza vaccines have been shown to significantly reduce morbidity and mortality from influenza. Next-generation flu vaccines aim to evoke a broader, longer immune response to overcome the epidemic of seasonal influenza virus antigens and to deal with the emergence of an emerging strain of influenza virus on time. From Buynder *et al.* [3] have found the most effective prevention of Influenza in the elderly is vaccinated with an adjuvanted and unadjuvanted trivalent inactivated influenza vaccine (TIV). Essink *et al.* [4] conducted a study to evaluate whether an adjuvanted quadrivalent influenza vaccine (aQIV) and a licensed adjuvanted trivalent influenza vaccine (aTIV) could provide broader protection against influenza in older adults. The study concluded that aQIV could provide broader protection than aTIV against influenza in older adults.

Accurate prediction of influenza patients contributes to early public health preparedness measures that can lessen the impact of influenza seasons. Machine learning models are good at predicting infectious disease variables that have long time frames. Influenza forecasting has evolved through three eras. The first era used the compartment model, which forecasts the influenza spread. This model is effective in tracking the status of infected populations and includes types like susceptible-infected-recovered (SIR) [5], [6], susceptible-infected-recovered-susceptible (SIRS) [7], [8], and susceptible-exposed-infected-recovered (SEIR) [9], [10], which are deterministic and calibrated for capturing the dynamics of influenza outbreaks. In the second stage, statistical techniques such as Box-Jenkins, automated regression integration (ARIMA) [11], and generalized autoregressive moving average (GARMA) [12] were used to produce highly accurate forecasts of influenza.

In modern times, the application of deep learning models has become prevalent in disease diagnosis prediction. The predictive disease models include the stacked linear regression [13], support vector regression [14], binomial [15], and classification and regression trees [16]. The related dataset, such as flu count and vaccination, was provided in forecasting models to increase the accuracy of influenza prediction results. Whilst machine learning techniques excel in organizing data based on relevant variables, they are relatively costlier than statistical models as they require retraining with the introduction of fresh datasets.

A novel data-driven model for real-time Influenza forecasting [17] represented influenza prediction model via long short-term memory (LSTM) and automated regression integration (ARIMA). This paper showed the LSTM which is the deep learning model effected effectively influenza prediction in different geographical regions. This mentioned study provided two variables consisted of geographical proximity flu count and climatic variables. Humidity, temperature, precipitation, and sun exposure were climate factors. This model performed regional data on the Google flu trends (GFT) and center for disease control (CDC) influenza patient. While our present research included a vaccination dataset from the focus region and adjacent region integrated with climate variables and flu count history from the node and adjacent region for increased effective influenza forecasting.

Jiaming *et al.* [18] provided convolution neural networks (CNNs) for message hashing discusses a novel approach to word hashing using CNNs in which keyword features are encoded into a compact binary code. The network introduces a keyword attribute to maintain the restriction, as well as a position property which is a combination of the word and position properties. Additionally, the relative distances of each word in the text are encoded into vectors using a CNN experiment. The model does not require external tags/labels and was tested on two datasets using a multi-state hash method. Our CNN model was proposed to forecast influenza by processing flu count feature on node and adjacent regions, vaccination, and climate variables resulting in significantly high accuracy.

Our previous research [19] has shown that we can enhance the spatial transmission effects [20] of regional influenza using the CNN model from mentioned papers [17], [18]. In this paper, since spatial transmission is the main cause of the influenza epidemic. We improved human movement with the geographic topology including the local transmission effects [20], and represent it in our CNN influenza forecasting model. The CNN improved influenza forecasting accuracy by increasing the adjustment of the spatio-temporal factors to regional nodes and adjacent neighborhood regions to incur spatial correlations. Additionally, our study proposes a combination of climate variables, vaccinations, and spatiotemporal features from the United States territory region. In our influenza forecasting paper. We represented the encounter between Americans and people in an adjacent region in the United States where significant influenza spread with the CNN model.

2. METHOD

The CNN model has identified three critical variables in predicting influenza outbreaks in the United States. These variables are; i) climate variable patterns, including snowfall, precipitation, and temperature variations of the focus region, ii) the influenza vaccination of the focus region and adjacent, and iii) a spatio-temporal factor that accounts for flu counts in the focal region and neighboring areas. The CNN model was trained on a comprehensive dataset surrounding all regions of the United States.

2.1. Climate multivariate data

In this complicated dataset, the climate serves as an essential determinant. The data was sourced from the valued climate data online (CDO) [21] and includes maximum temperature (TMAX) measured in tenths of degrees C, lowest temperature (TMIN) measured in tenths of degrees C, precipitation (PRCP) measured in tenths of millimeters unit, and snowfall (SNOW) measured in tenths of millimeters unit. These datasets were gathered from the global historical climatology network (GHCN) for the United States. The weekly data was converted from two hundred million daily datasets of climate datasets. The model uses flu count data and then transforms data into scaled values for time-series analysis. The climatic data was collected quarterly from 2009-2016, making for a comprehensive analysis.

2.2. Influenza vaccination data

The second determinant is a vaccination immunization data. Immunization datasets collected from the CDC [22] which involed vaccine, region, month of each influenza season, and estimated vaccination coverage is resulted from percent of persons vaccinated in a particular place (percentage). We selected vaccination data derived from the sample surveyed in years 2009-2016 to estimate influenza imunization in each region. The 184,329 datasets have been converted into weekly data by utilizing the vaccination data of the region.

2.3. Spatio-temporal data

Within our research, we are taking into consideration the spatio-temporal data. We are utilizing flu count data from ten different regions across the United States spanning from 1997 to 2016. This data has been collected through the CDC [23] and the national center for immunization and respiratory diseases (NCIRD). The geographical regions of the United States can be seen in Figure 1. Additionally, we are referring to historical flu count data in specific locations, which can be found in Table 1. This includes the focus node and its adjacent regions. There are five groups of local transmission regions based on their proximity, region 1 has only one adjacent region, region 2 and region 10 have two adjacent regions, region 3 and region 9 have three adjacent regions. Region 4, region 5, region 6, and region 7 have four adjacent regions. The largest number of contiguous regions is region 8 with 5 adjacent regions. All region has a global influenza forecasting model in recurrent neural network (RNN), LSTM, and CNN.

Figure 1. The United States regional map

Table 1. Territory and their neighboring

Territory	Neighboring
Territory 1	Territory 2
Territory 2	Territory 1, Territory 3
Territory 3	Territory 2, Territory 4, Territory 5
Territory 4	Territory 3, Territory 5, Territory 6, Territory 7
Territory 5	Territory 3, Territory 4, Territory 7, Territory 8
Territory 6	Territory 4, Territory 7, Territory 8, Territory 9
Territory 7	Territory 4, Territory 5, Territory 6, Territory 8
Territory 8	Territory 5, Territory 6, Territory 7, Territory 9, Territory 10
Territory 9	Territory 6, Territory 8, Territory 10
Territory 10	Territory 8, Territory 9

2.4. Convolution neural network

CNN is an extension of the traditional multilayer perceptron (MLP) [24] that aims to achieve shift and distortion invariance by incorporating three key features: local receptive fields, shared weights, and spatial and temporal subsampling. CNNs consist of multiple trainable layers [25], each of which has a set of feature maps [26] as input and output. The output of each layer represents the extracted features from the input at different locations. Typically, each layer of CNN has three sublayers: a non-linearity sublayer, a filter bank sublayer, and a feature pooling sublayer. The final output of CNN is obtained by one or more fully connected layers after several convolution and pooling layers.

2.5. Proposed influenza forecasting model

In order to improve the accuracy of predicting influenza outbreaks, we utilized a 1D CNN with four channels to integrate climate determinants such as precipitation, snowfall, maximum and minimum temperatures of the focus region. Our fifth channel included two vaccine coverage features, measuring vaccine coverage data in the target region and neighboring regions. Additionally, we combined spatio-temporal features in our sixth channel, measuring flu count data in the focus and adjacent regions. By considering these factors, an influenza predictive model is based on our CNN. was able to accurately predict flu outbreaks in neighboring regions. This is particularly important as the interaction of people in the United States has been a significant factor in the spread of the virus. A diagram of our model is included in Figure 2.

For a CNN model approach, the associate dataset from every city in the United States was accumulated into a single dataset. The dataset was simultaneously trained in the different attributes size of vaccination and geography in each region dataset. Finally, all determinant datasets are flattened and concatenated to predict the influenza epidemic in different weeks. Our study suggested by a CNN model aimed to predict the occurrence of influenza by analyzing multiple sets of flu data using various time intervals and neural network structures. The advantage of this approach is that it effectively combines the impact of neighboring regions on the influenza spread. By using deep learning to analyze big data, we aim to identify effective patterns for predicting the spread of influenza.

The CNN model had six channels. In four channels, we establish input X into a 1D convolution in each channel. These four convolution inputs were derived from climate variables C, which is rainfall in the shortening *PRCP*, snowfall in the shortening *SNOW*, maximum temperature in the shortening *TMAX*, and minimum temperature in the shortening *TMIN*, incorporating data from prior weeks. As a conceptual, when dealing with input dataset C located in region n, the model employs (1):

$$C^n = \{C_{j1}^n, C_{j2}^n, C_{j3}^n, \dots, C_{jT}^n\} \tag{1}$$

Where C^n is the variable of climate; n is target region node | $n \in \{1,2,3,\dots,10\}$; j *is* the temporal sequences of four climatic variables $j \mid j \in$ {*PRCP*, *SNOW*, *TMAX*, *TMIN*}; and T is the time series from T=1 to T.

$$C^n = \left\{\begin{bmatrix} C_{prcp1} \\ C_{snow1} \\ C_{tmax1} \\ C_{tmin1} \end{bmatrix}, \begin{bmatrix} C_{prcp2} \\ C_{snow2} \\ C_{tmax2} \\ C_{tmin2} \end{bmatrix}, \begin{bmatrix} C_{prcp3} \\ C_{snow3} \\ C_{tmax3} \\ C_{tmin3} \end{bmatrix}, \begin{bmatrix} C_{prcp4} \\ C_{snow4} \\ C_{tmax4} \\ C_{tmin4} \end{bmatrix}\right\} \tag{2}$$

In (2) and Figure 2 the climatic of the region dataset is used in 4 time steps. As an theoretical, if we have a climate dataset for a certain region 8 with *PRCP* values of {8,5,3,4,…}, *SNOW* values of {5,4,2,7,…}, *TMAX* values of {6,2,8,5,…}, and *TMIN* values of {1,3,5,7,…}. We can use (3) to input this climatic data.

$$C^8 = \left\{\begin{bmatrix} 8 \\ 5 \\ 6 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 8 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ 5 \\ 7 \end{bmatrix}\right\} \tag{3}$$

The flu prediction model has added a fifth channel which is vaccination input data V. This channel uses two features: the first was the vaccination rate of the focus region node and the second was the vaccination rate v of the neighboring nodes k from the week i, as indicated in Table 1 of our model. If there is no data available for a particular node, it is assigned a null k value of 0 using the fillna function in Python. Both features are then fed into a 2D CNN as per (4):

$$V_k^n = \{v_{i,k}^n, v_{i+1,k}^n, \dots, v_{i+N,k}^n\} \tag{4}$$

Where V_k^n is the variable of influenza vaccination; $n\ is$ the target territory | $n \in \{1,2,3,\dots,10\}$; $i\ is$ the sequence of time; $v\ is$ the veritable rate of influenza vaccination in the contiguous vicinity; and $k\ is$ the adjacent node of the focus entity | $k \in \{1,2,\dots,K_{max}\}$ | $K_{max}\ is$ maximum number of adjacent regions.

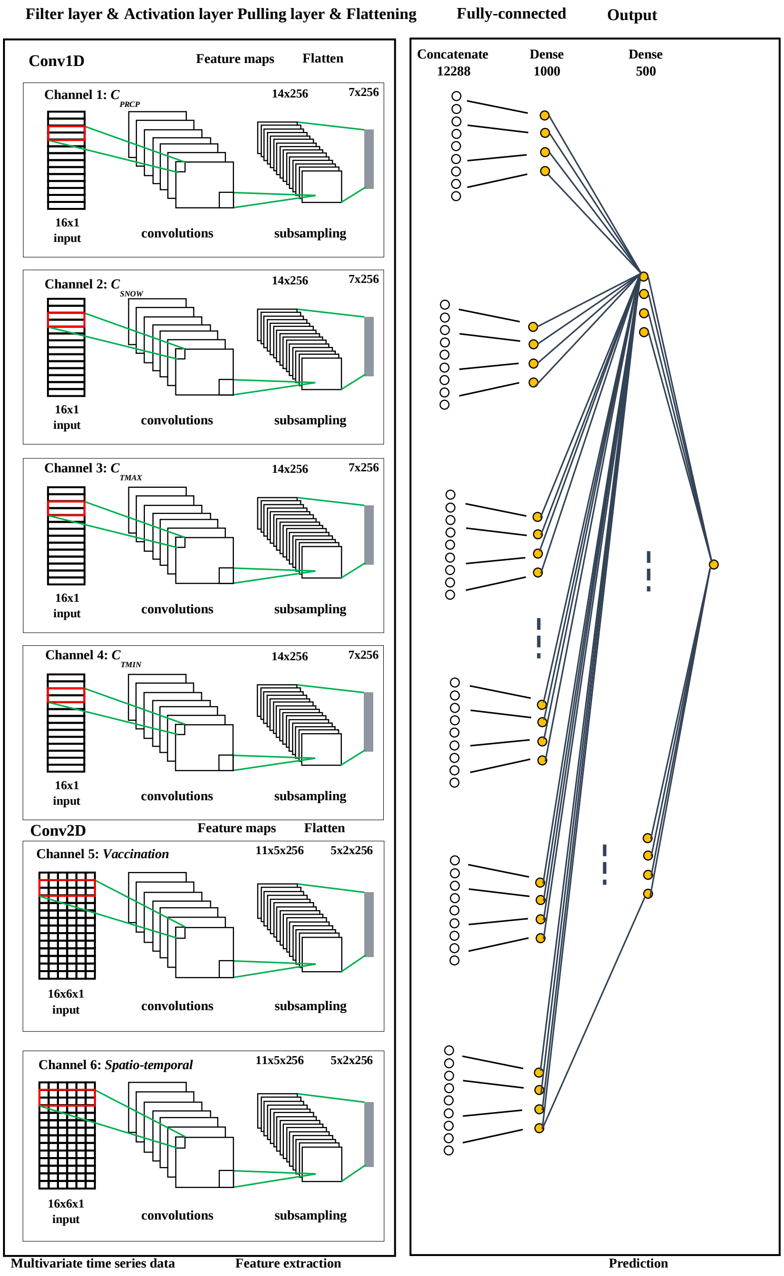


Figure 2. The dataset pattern utilized in the proposed global CNN for forecasting influenza

The datasets used for training consist of pairs $\{x_i, y_i\}$, starting from time i and continuing until it reaches end time t | where i is a member of the set $\{1,2,3,\ldots,N\}$ for N examples. The sequence of training is denoted as (x_i, y_i) in all its complexity, where $x_i^n = V_n^n \cup \{V_k^n\}$ and $y_i^n = V_n^n$, w=sliding window size. As an example, if the vaccination rate of Region 8 node is [4,2,7,4,7,5,3,...], and vaccination rate of five adjacent regions consisting of the 1st Region (Region 5)=[8,7,9,5,8,9,3,…], the 2nd region (region 6)=[6,2,1,8,4,9,7,…], the 3rd region (region 7)=[1,5,9,7,4,1,6,…], the 4th region (region 9)=[6,-,-,4,2,4,3,…], the 5th region (region 10)=[1,5,-,3,8,2,1,…]. The sliding window size is 4 and the input x, which is spatio-temporal, as (5):

$$x_i^8, y_i^8 = \left\{\begin{bmatrix} 8&6&1&4&6&1&4\\ 7&2&5&2&0&5&2\\ 9&1&9&7&0&0&7\\ 5&8&7&4&4&3&4 \end{bmatrix}\right\} \tag{5}$$

Through the sixth channel of the influenza prediction model, we have successfully formulated spatio-temporal A by attaching two distinct attributes. The first attribute draws upon the historical flu count of the focal node, while the second component integrates flu count information from node A and its neighboring nodes k from the week i, as indicated in Table 1 of our model. For nodes in the flu count A category, we assign a null k value of 0 using the fillna function in Python. Both features are then fed into a 2D CNN as outlined in (6):

$$A_k^n = \{a_{i,k}^n, a_{i+1,k}^n, \dots, a_{i+N,k}^n\} \tag{6}$$

Where A_k^n is the variable of spatio-temporal flu; n is the target territory | $n \in \{1,2,3, \dots ,10\}$; i is the sequence of time; a is the veritable count of influenza cases in the contiguous vicinity; and k is the adjacent node of the focus entity | $k \in \{1,2, \dots, K_{max}\}$ | K_{max} =Maximum number of adjacent regions.

The datasets used for training consist of pairs $\{x_i, y_i\}$, starting from time i and continuing until it reaches end time t | where i is a member of the set $\{1,2,3, \dots, N\}$ for N examples. The sequence of training is denoted as (x_i, y_i) in all its complexity, where x_i^n = $A_n^n \cup \{A_k^n\}$ and y_i^n= A_n^n, w=sliding window size. As an example, we will keep a record of the number of flu cases in region 8, which includes the values [2,4,7,1,1,5,3,…]. The flu count of five adjacent regions consisted of the 1st region (region 5)=[4,5,3,8,7,5,3,…], the 2nd region (region 6)=[7,3,9,6,4,9,7,…], the 3rd region (region 7)=[3,2,4,8,4,1,6,…], the 4th region (region 9)=[1,3,-,4,2,2,1,…], the 5th region (region 10)=[7,-,-,-,7,4,8,…]. A sliding window size of 4 will be used, and the spatio-temporal input x will as (7):

$$x_i^8, y_i^8 = \left\{ \begin{bmatrix} 4\ 7\ 3\ 2\ 1\ 7\ 2 \\ 5\ 3\ 2\ 4\ 3\ 0\ 4 \\ 3\ 9\ 4\ 7\ 0\ 0\ 7 \\ 8\ 6\ 8\ 1\ 4\ 0\ 1 \end{bmatrix} \right\} \tag{7}$$

3. RESULTS AND DISCUSSION

Using Python and Tensorflow libraries, we conducted a thorough examination of specific neural network models to determine the most effective one. We analyzed data related to climate, vaccination, and flu count from 1997 to 2016 using all ten regional models to attain the results. Our neural architecture was trained using a differential time step varying from twelve to twenty-eight weeks. The models we assessed included RNN, LSTM neural network, and CNN.

3.1. Evaluation metrics

In the evaluation of each model, two crucial metrics are employed- the mean absolute error (MAE) and the root mean square error (RMSE), as previously utilized in [27], [28]. To expand, the actual influenza count value is defined as A, while F represents the predicted influenza value, and N denotes the number of observations. The absolute values are represented by vertical bars, as agreed by the performance predictor. These metrics form the basis for the valuation of the model.

MAE is a statistical measure used to evaluate the performance of predictive prediction models. MAE measures the accuracy or precision of predictive models by computing the average of the absolute difference between predicted and observed data points. MAE is a powerful performance evaluation criterion, as it considers the size of errors without regard to their direction, making it particularly useful when assessing the overall model fit in regression and forecasting tasks. It is naturally expressed as a percentage and is defined by (8):

$$MAE = \frac{1}{N} \sum (A - F) \tag{8}$$

The RMSE is a statistical metric widely applied in various instruction, including statistics, machine learning, and engineering, to assess the accuracy of predictive models by calculating the square root of the mean of the squared differences between predicted and observed values, providing a comprehensive measure of the magnitude of errors and their dispersion in relation to the true values. To compute the RMSE. The square root of the mean of the difference between the predicted and actual values is squared. is taken, as demonstrated by (9):

$$RMSE = \sqrt{\frac{1}{N} \sum (A - F)^2} \tag{9}$$

3.2. Resulted of experiments

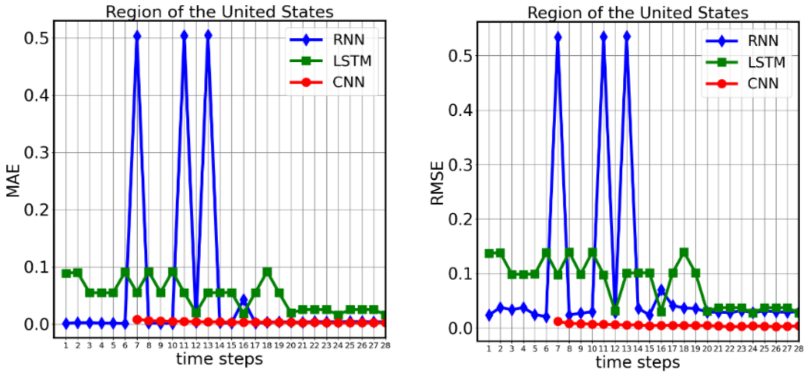
The study compared the predictive values of various models, including DENSE and different neural networks, at different time steps with a learning rate=0.0001. The CNN model was designed to work with varying filter sizes, depending on the adjacent neighboring regions. The study found that the CNN model was able to predict flu epidemics by analyzing climate variables at the focus node, vaccination, and flu counts on nodes and adjacent neighbors. Specifically, the CNN model performed better than the RNN and LSTM models in predicting influenza in regions with higher time steps. The study found that prediction was less effective in areas with a lower time step. Additionally, the study evaluated the performance of RNN, LSTM, and CNN models using MAE and RMSE, and the findings are presented in Table 2.

Table 2. The evaluation of predictive competences for influenza utilizing RNN, LSTM, and CNN methodologies

	Timesteps									
Model	12		16		20		24		28	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
RNN	0.00353	0.02802	0.04335	0.07010	0.00404	0.02938	0.00455	0.02840	0.00425	0.03025
LSTM	0.01998	0.03106	0.01882	0.02975	0.01995	0.03101	0.01662	0.02728	0.01668	0.02740
CNN	**0.00447**	**0.00596**	**0.00390**	**0.00514**	**0.00343**	**0.00446**	**0.00298**	**0.00386**	**0.00290**	**0.00376**

Within Table 2, one may observe the MAE and RMSE performance outcomes affecting influenza prediction through the employment of RNN, LSTM, and CNN models. The CNN model produced the most accurate predictions when using 12-week time steps, with an RMSE of 0.00596 compared to RNN value of 0.02802 and LSTM value of 0.03106. When reducing the time steps to 16 weeks, CNN model has lower MAE and RMSE value than RNN and LSTM models, with an RMSE of 0.00513, 0.07010, and 0.02975, respectively. These results are significant because the CNN model was able to extract and learn influenza spread patterns from limited time series data, resulting in a significant reduction in RMSE error. The CNN model achieved the lowest RMSE of 0.00376 with 28-week time steps, while RNN and LSTM had RMSEs of 0.03025 and 0.02740, respectively.

Figure 3 displays the outcomes of influenza forecasting in the United States using RNN, LSTM, and CNN models. The MAE error values showed in Figure 3(a) and RMSE error values are presented in Figure 3(b). The predictive models were compared with reliable flu CDC datasets. Our proposed CNN model proved superior in predicting influenza over a twelve-week period compared to RNN and LSTM models. The three critical factors influencing influenza prediction in the United States are climate in focus region, vaccine coverage, and the number of flu cases in each region and surrounding areas. We have provided a chart of performance errors for these indicators. The CNN effectively predicted influenza based on CNN model. In the future, by including spatio-temporal data, disease transmission studies can effectively utilize location data that involves a considerable population. The dataset of place coordinates, derived from Google Maps location data, can significantly contribute to this approach, and can provide more precise flu prediction.



(a) (b)

Figure 3. The MAE and RMSE errors of RNN, LSTM, and CNN influenza prediction models; (a) MAE and (b) RMSE

4. CONCLUSION

This study outlines a model that utilizes RNN, LSTM, and CNN to predict an influenza epidemic in the United States. The dataset used in the study consists of three factors: climate data, vaccination data, and flu count historical data. The spatial correlations between influenza patients and geographically adjacent regions have been represented by the CNN model impacting the accuracy of influenza epidemic predictions. The experiment was conducted using a default timestep of twelve, sixteen, twenty, twenty-four, and twenty-eight weeks. The CNN model at twenty-eight time steps was found to be the most effective. This research demonstrates that machine learning can be used to predict influenza with greater accuracy using time-series data and spatio temporal data. The study used data from ten regions in the United States, and we are hoping to expand their dataset to include adjacent nodes. We suggest that their model has the possibility to be applied to various clinical outbreaks and infectious diseases. They predict that in the future, disease transmission studies could be conducted by integrating location data of numerous individuals, in addition to spatio-temporal data. The dataset used in this study is derived from precise flu prediction data on Google Maps, indicating the exact coordinates of the locations.

REFERENCES

[1] C. for disease control and Prevention, “Overview of Influenza surveillance in the United States.” [Online]. Available: https://www.cdc.gov/flu/weekly/overview.htm. (accessed: Jan. 10, 2017).

[2] C. Mameli, I. Cocchi, M. Fumagalli, and G. Zuccotti, “Influenza Vaccination: Effectiveness, Indications, and Limits in the Pediatric Population,” *Frontiers in Pediatrics*, vol. 7, Jul. 2019, doi: 10.3389/fped.2019.00317.

[3] P. G. Van Buynder *et al.*, “The comparative effectiveness of adjuvanted and unadjuvanted trivalent inactivated influenza vaccine (TIV) in the elderly,” *Vaccine*, vol. 31, no. 51, pp. 6122–6128, Dec. 2013, doi: 10.1016/j.vaccine.2013.07.059.

[4] B. Essink *et al.*, “Immunogenicity and safety of MF59-adjuvanted quadrivalent influenza vaccine versus standard and alternate B strain MF59-adjuvanted trivalent influenza vaccines in older adults,” *Vaccine*, vol. 38, no. 2, pp. 242–250, Jan. 2020, doi: 10.1016/j.vaccine.2019.10.021.

[5] H. W. Hethcote, “The Mathematics of Infectious Diseases,” *SIAM Review*, vol. 42, no. 4, pp. 599–653, Jan. 2000, doi: 10.1137/S0036144500371907.

[6] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2011. doi: 10.2307/j.ctvcm4gk0.

[7] M. B. Hooten, J. Anderson, and L. A. Waller, “Assessing North American influenza dynamics with a statistical SIRS model,” *Spatial and Spatio-temporal Epidemiology*, vol. 1, no. 2–3, pp. 177–185, Jul. 2010, doi: 10.1016/j.sste.2010.03.003.

[8] J. Shaman, A. Karspeck, W. Yang, J. Tamerius, and M. Lipsitch, “Real-time influenza forecasts during the 2012–2013 season,” *Nature Communications*, vol. 4, no. 1, p. 2837, Dec. 2013, doi: 10.1038/ncomms3837.

[9] G. Chowell, M. A. Miller, and C. Viboud, “Seasonal influenza in the United States, France, and Australia: transmission and prospects for control,” *Epidemiology and Infection*, vol. 136, no. 6, pp. 852–864, Jun. 2008, doi: 10.1017/S0950268807009144.

[10] G. Chowell, H. Nishiura, and L. M. Bettencourt, “Comparative estimation of the reproduction number for pandemic influenza from daily case notification data,” *Journal of The Royal Society Interface*, vol. 4, no. 12, pp. 155–166, Feb. 2007, doi: 10.1098/rsif.2006.0161.

[11] K. Choi and S. B. Thacker, “An evaluation of influenza mortality surveillance, 1962-1979. II. Percentage of pneumonia and influenza deaths as an indicator of influenza activity,” *American Journal of Epidemiology*, vol. 113, no. 3, pp. 227–235, Mar. 1981, doi: 10.1093/oxfordjournals.aje.a113091.

[12] A. F. Dugas *et al.*, “Influenza Forecasting with Google Flu Trends,” *PLoS ONE*, vol. 8, no. 2, p. e56176, Feb. 2013, doi: 10.1371/journal.pone.0056176.

[13] J. C. Santos and S. Matos, “Analysing Twitter and web queries for flu trend prediction,” *Theoretical Biology and Medical Modelling*, vol. 11, no. S1, p. S6, May 2014, doi: 10.1186/1742-4682-11-S1-S6.

[14] A. Signorini, A. M. Segre, and P. M. Polgreen, “The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic,” *PLoS ONE*, vol. 6, no. 5, p. e19467, May 2011, doi: 10.1371/journal.pone.0019467.

[15] H. Nishiura, “Real-time forecasting of an epidemic using a discrete time stochastic model: a case study of pandemic influenza (H1N1-2009),” *BioMedical Engineering OnLine*, vol. 10, no. 1, p. 15, 2011, doi: 10.1186/1475-925X-10-15.

[16] S. C. Lemon, J. Roy, M. A. Clark, P. D. Friedmann, and W. Rakowski, “Classification and regression tree analysis in public health: Methodological review and comparison with logistic regression,” *Annals of Behavioral Medicine*, vol. 26, no. 3, pp. 172–181, Dec. 2003, doi: 10.1207/S15324796ABM2603_02.

[17] S. R. Venna, A. Tavanaei, R. N. Gottumukkala, V. V. Raghavan, A. S. Maida, and S. Nichols, “A Novel Data-Driven Model for Real-Time Influenza Forecasting,” *IEEE Access*, vol. 7, pp. 7691–7701, 2019, doi: 10.1109/ACCESS.2018.2888585.

[18] X. Jiaming *et al.*, “Convolutional neural networks for text hashing,” in *Proc. the 24th International Conference on Artificial Intelligence (IJCAI’15), AAAI Press*, 2015, pp. 1369–1375.

[19] J. Watmaha, S. Kamonsantiroj, and L. Pipanmaekaporn, “An Integrated Climate and Spatio-temporal Determinant for Influenza Forecasting based on Convolution Neural Network,” in *Proceedings of the 9th International Conference on Computer and Communications Management*, New York, NY, USA: ACM, Jul. 2021, pp. 104–110. doi: 10.1145/3479162.3479178.

[20] F. Xie, Z. Zhang, L. Li, B. Zhou, and Y. Tan, “EpiGNN: Exploring Spatial Transmission with Graph Neural Network for Regional Epidemic Forecasting,” in *Machine Learning and Knowledge Discovery in Databases: European Conference*, 2023, pp. 469–485. doi: 10.1007/978-3-031-26422-1_29.

[21] M. J. Menne *et al.*, “Global Historical Climatology Network - Daily (GHCN-Daily),” *NOAA National Climatic Data Center*, 2012.

[22] C. for D. C. and Prevention, “Influenza Vaccination Coverage for All Ages (6+Months).” [Online]. Available: https://data.cdc.gov/Flu-Vaccinations/Influenza-Vaccination-Coverage-for-All-Ages-6-Mont/vh55-3he6. (Accessed: Oct. 02, 2022).

[23] C. for disease control and Prevention, “National, Regional, and State Level Outpatient Illness and Viral Surveillance.” [Online].

Available: https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html. (Accessed: Jan. 10, 2020).

[24] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time-series," *The handbook of brain theory and neural networks MIT Press*, pp. 255–258, 1998.

[25] Y. L. Cun *et al.*, "Handwritten digit recognition with a back-propagation network," *Advances in neural information processing systems2. Morgan Kaufmann Publishers Inc*, pp. 396–404, 1990.

[26] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.

[27] R. Fildes, "The evaluation of extrapolative forecasting methods," *International Journal of Forecasting*, vol. 8, no. 1, pp. 81–98, Jun. 1992, doi: 10.1016/0169-2070(92)90009-X.

[28] N. G. Reich, J. Lessler, K. Sakrejda, S. A. Lauer, S. Iamsirithaworn, and D. A. T. Cummings, "Case Study in Evaluating Time Series Prediction Models Using the Relative Mean Absolute Error," *The American Statistician*, vol. 70, no. 3, pp. 285–292, Jul. 2016, doi: 10.1080/00031305.2016.1148631.

BIOGRAPHIES OF AUTHORS

Jaroonsak Watmaha received Bachelor of Business Administration (Business Information Technology) from Sukhothai Thammathirat Open University, Bangkok, Thailand in 2005. He received the Master of Science (Information and Communication Technotogy) from Sukhothai Thammathirat Open University, Bangkok, Thailand, in 2015. Currently he is the chief of Information Technology and Public Relations at Narathiwat Provincial Public Health Office. His research interests include informix, data science, and machine learning. He can be contacted at email: s5904062910023@email.kmutnb.ac.th.

Suwatchai Kamonsantiroj is currently a lecturer at Department of Computer and Information Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand. He holds a Bachelor's degree in Mechanical Engineering and a Master's degree in Information Technology Management. He also earned his Doctoral degree in Computer Engineering from Kasetsart University, Thailand, graduating in 2008. His current research interests include neural network, time series analysis, and artificial intelligence. He can be contacted at email: suwatchai.k@sci.kmutnb.ac.th.

Luepol Pipanmaekaporn is currently a lecturer at Department of Computer and Information Science, King Mongkut's University of Technology North Bangkok, Bangkok, Thailand. He holds both a Bachelor's and a Master's degree in Computer Science. He also earned his Doctoral degree in Computer Science from Queensland University of Technology, Australia, graduating in 2013. His current research interests include information retrieval, web mining, and data mining. He can be contacted at email: luepol.p@sci.kmutnb.ac.th.