




Influenza-like illness prediction using a long short-term memory deep learning model with multiple open data sources

Chao-Tung Yang¹ · Yuan-An Chen¹ · Yu-Wei Chan² · Chia-Lin Lee^{3,4,5} · Yu-Tse Tsan^{6,7}  · Wei-Cheng Chan⁶ · Po-Yu Liu⁸

Published online: 21 February 2020
© The Author(s) 2020

Abstract

The influenza problem has always been an important global issue. It not only affects people's health problems but is also an essential topic of governments and health care facilities. Early prediction and response is the most effective control method for flu epidemics. It can effectively predict the influenza-like illness morbidity, and provide reliable information to the relevant facilities. For social facilities, it is possible to strengthen epidemic prevention and care for highly sick groups. It can also be used as a reminder for the public. This study collects information on the influenza-like illness emergency department visits to the Taiwan Centers for Disease Control, and the PM_{2.5} open-source data from the Taiwan Environmental Protection Administration's air quality monitoring network. By using deep learning techniques, the relevance of short-term estimates and the outbreak calculation method can be determined. The techniques are published by the WHO to determine whether the influenza-like illness situation is still in a stage of reasonable control. Finally, historical data and future forecasted data are integrated on the web page for visual presentation, to show the actual regional air quality situation and influenza-like illness data and to predict whether there is an outbreak of influenza in the region.

Keywords Influenza-like illness · LSTM · PM_{2.5} · Deep learning

1 Introduction

Human physiological disorders reflect an altered condition that interferes with or modifies the vital functions of various organs or body parts [1]. Kim et al. [2] examined whether the severity of posttraumatic stress disorder (PTSD) symptoms and perceived functional impairment in firefighters with current possible PTSD are

✉ Yu-Tse Tsan
janyuhjer@gmail.com

Extended author information available on the last page of the article

correlated with the use of mental health services. Gautam et al. [3] also studied the treatment of medical conditions in humans, and found the rate of predictive accuracy obtained by ant colony optimization (ACO) and neural network hybridization to be more promising than other individual or hybrid approaches. Based on their predicted human health implications, Van der Fels-Klerx et al. [4] thoroughly reviewed criteria for rating risks related to food safety and dietary hazards. The method to be used should be selected based on the criteria of the risk manager/assessor, the quality of data, and the method's characteristics.

Influenza is a direct and far-reaching problem. An uncontrolled flu epidemic can have a significant impact on all of society, for example, in the global H1N1 influenza outbreak around the world in 2009. According to data released by the World Health Organization, there were more than 1.3 million confirmed cases of H1N1 in the world, with a death toll of more than 14,000, presenting a significant challenge to the world's quarantine mechanism [5]. For Taiwan, many prevention and control measures have been proposed for incidents of this type, such as strict border control, mobilization, and isolation of medical facilities, public transportation bases, and campuses in order to maximize anti-epidemic measures in public places. It can be said that the influenza epidemic has a profound and long-lasting impact on people's lives, the safety of their property, human resources, material resources, and socioeconomic stability. If compelling trend predictions for flu-like epidemics can be achieved, relevant units can be provided. More reaction time allows the public to take prevention or treatment measures earlier, which can have a positive impact on the control of the flu-like epidemic. There are many medical studies confirming the relationship between air quality index (AQI) and influenza-like illness (ILI) throughout the world. Increased rates of culture-negative pneumonia and influenza were associated with increased concentrations of particulate matter $< 2.5 \mu\text{m}$ ($\text{PM}_{2.5}$) during the previous week, which persisted despite reductions in $\text{PM}_{2.5}$ from air quality policies and economic changes. Though unexplained, this temporal variation may reflect altered toxicity of different $\text{PM}_{2.5}$ mixtures or increased pathogen virulence [6–15].

Influenza-like illness (ILI) is defined as a medical condition that may be caused by influenza or other diseases such as parainfluenza viruses or adenoviruses with clinical manifestation in common with the influenza virus. According to the WHO surveillance case definition, ILI is defined as an acute respiratory infection with fever more than 38°C and with cough symptoms, as well as morbidity within the last 10 days. This study will collect information on the rate of influenza-like emergency visits from the Department of Health and Welfare and the $\text{PM}_{2.5}$ open-source data from the Environmental Protection Department's Air Quality Monitoring Network [16]. Deep learning techniques are used to predict the status of future influenza-like treatment and the effects of air pollution on the respiratory tract in inducing influenza-like symptoms, as well as the natural delay characteristics between a patient's inhalation of excessively high levels of contaminated gases and the onset of flu-like symptoms [17]. This system predicts the rate of ILI after a short period based on the current level of air pollution.

In this paper, the long short-term memory network (LSTM) of the recurrent neural network (RNN) will be implemented to solve the time-series prediction case. Also, it

can address the shortcomings of the disappearance of the RNN gradient [18, 19], so that LSTM can effectively reduce the weight of the current time. Therefore, the module can more effectively preserve the influence of special events before the distant future point. In combination with the Outbreaks algorithm published by the WHO, this algorithm can calculate the criteria of influenza-like outbreaks at various time points during the past 5 years, and can be used as supplementary data for training and subsequent data visualization. It may also function as a benchmark to aid decisions when presented.

Our experimental setup presents a visualization of the model based on historical data combined with the short-term forecast data predicted by the module. It allows the user to intuitively understand real-time air pollution data in a user-friendly system. The user can easily understand the current rate of influenza visits and the predicted rate in the short term. Hopefully, this influenza prediction platform can serve as a reference for administrative decision-making and as a basis for general public health measures, with the ultimate goal of effectively controlling and preventing influenza-like epidemics. The significance of our contribution lies in the ability to predict ILI based on the integration of the rate of influenza-like emergency visits and PM_{2.5} data using LSTM.

2 Background review and related works

In this section, we discuss several components that are included in the approaching in this paper: RNN, LSTM, and the mean absolute percent error (MAPE). The following sections discuss each component in more detail.

2.1 Recurrent neural network (RNN)

RNN is a type of feed-forward neural network. It is simply a network type that returns the value output from the neural network back to the input [20–22]. In this way, the output value of the last time point is transmitted back to the neural network, so that the weight calculation of each time point of the network model is related to the content of the previous time point, which means that the neural network is included. The concept of time, through such a mechanism, makes the neural network memory. Figure 1 describes recurrent neural network architecture. After each round of operation, the output of the web is retained up to the next transfer operation. Therefore, the time t will be $t + 1$ time points. The output is taken into consideration, so it has the characteristics of before and after memory.

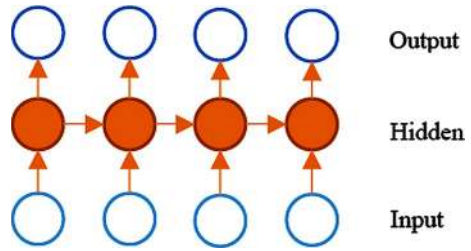
The following is the formula for RNN:

$$h_t = \sigma_h(W_h x_t + U_h y_{t-1} + b_h) \quad (1)$$

$$y_t = \sigma_y(W_y h_t + b_y) \quad (2)$$

where x_t is the vector of the input layer, h_t is the vector of the hidden layer, y_t is the vector of the output layer, W , U and b are the matrix or vector of the weight

Fig. 1 RNN schematic diagram



parameter, and σ_h and σ_y are the activation functions. From this formula, it can be clearly observed that the output previous time unit is added to the calculation of y_{t-1} .

2.2 Long short-term memory network

LSTM is a deformation module constructed on the basis of RNN, but the mechanism of adding gate allows the LSTM to effectively save events before long-distance time [23]. The weights of this method solve the defect in which the RNN is prone to gradient disappearance. Therefore, LSTM is more suitable for processing important events with longer intervals or delay time units in the time series than the RNN [24, 25] (Fig. 2).

The following is the formula for LSTM:

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{3}$$

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \tag{4}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{5}$$

$$c_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \tag{6}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot c_t \tag{7}$$

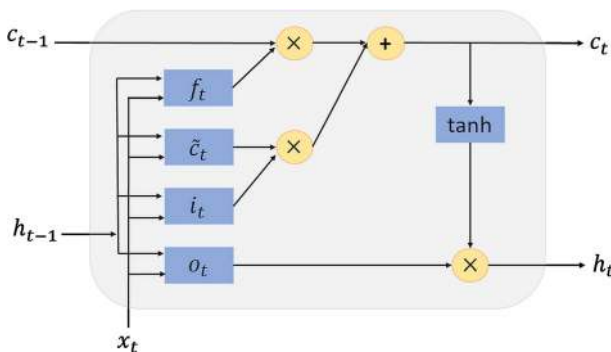


Fig. 2 LSTM schematic diagram

$$h_t = o_t \cdot \text{Tanh}(c_t) \quad (8)$$

$$y_t = h_t \quad (9)$$

where i_t is the input gate used to determine whether the current input value is added to the memory state, and f_t is the forget gate to determine whether the “forget” is previously stored. The memory state o_t is the output gate that will be used to determine the amount of memory output to the hidden layer, and also represents the degree of memory impact on the calculation.

2.3 MAPE

The mean absolute percent error (MAPE) as a percentage of the error reflects accuracy. Since the MAPE is a percentage, it can be easier to understand than the other figures on precision measurement. For example, if the MAPE is 5, the forecast is by 5% on average.

Nevertheless, a very large value of MAPE may sometimes be seen even though the model seems to suit the data well. Examine the plot to see if there are approximately 0 data values. Since MAPE divides the absolute error by the actual data, values close to 0 will inflate the MAPE significantly [26].

2.4 Related works

Yang et al. [27] proposed an analysis of air quality data and influenza-like illness (ILI) to determine the associations accurately and effectively. In their work, a novel integrated platform was implemented by building a cluster environment based on Hadoop and Spark. The experimental results showed the visualization of air quality and influenza-like illness data collected from 2016 to 2018 in Taichung, Taiwan. The association between air quality and influenza-like illness was also presented and discussed.

Liu et al. [28] focused on the association between invasive aspergillosis and ambient fine particulate air pollution. Two data sets were leveraged for this study: the National Health Insurance Research Database (NHIRD) was used to define invasive aspergillosis, while the Taiwan air quality index (AQI) monitoring network was used to profile the PM_{2.5} concentration in Taiwan. The findings of this study suggest a positive association between PM_{2.5} concentration and incidence of aspergillosis.

Lee et al. [29] demonstrated analysis and automated air pollution forecasting using RNN. A distributed computing environment was deployed based on RHadoop, which analyzed air pollution and presented a visualization using HBase from historical data. The short-term prediction of PM_{2.5} was presented, and the prediction accuracy based on the MAPE was measured.

Tang et al. [30] performed an integrated data analysis for the period from November 15, 2010, to November 14, 2016, to quantify the connection between AQI,

meteorological variables, and risk of respiratory infection in China's Shaanxi province. Their assessment showed a statistically significant positive correlation between the number of instances of ILI and AQI, and with a time lag of 0–3 days, the risk of pulmonary disease gradually improved with enhanced AQI.

Liu et al. [31] applied RNNs of LSTM to predict influenza patterns. They were the first to use various novel data sources to forecast influenza patterns, including virological surveillance, the geographic spread of influenza, trends in Google, the environment, and air pollution. They also discovered that there are several environmental and climatic variables that have an important correlation with the frequency of ILI.

Zhang et al. [32] used four distinct LSTM multi-step prediction algorithms to predict influenza outbreaks in multiple steps. The results showed that the greatest precision was attained by applying various single-output predictions in a six-layer LSTM framework. The MAPE for the US ILI rates from 2 to 13 steps forward were all less than 15%, with an average of 12.930%.

Huang and Kuo [33] integrated and applied a convolutional neural network (CNN) and LSTM to the $PM_{2.5}$ prediction method for controlling and predicting $PM_{2.5}$ concentration. Four measurement indexes, including mean absolute error (MAE), root mean square error (RMSE), Pearson correlation coefficient, and index of agreement (IA), were applied to the experiments to compare the overall performance of each algorithm. The experimental results confirmed that the predictive accuracy of the proposed CNN-LSTM model (APNet) was the highest in comparison with other machine learning models. The feasibility and practicability of the CNN-LSTM model in predicting the $PM_{2.5}$ concentration was also checked.

Tsai et al. [34] suggested an approach for predicting the concentration of $PM_{2.5}$ with LSTM using RNN. They leveraged Keras, a high-level neural network application programming interface (API) written in Python and able to run on top of TensorFlow, to construct a neural network and run RNN via TensorFlow with LSTM. Training data used in the network was collected from Taiwan's Environmental Protection Administration (EPA) from 2012 to 2016 and converted into 20-dimensional data at 66 stations across Taiwan. They conducted experiments to determine the importance of the predicted $PM_{2.5}$ concentration for the next 4 h. The results showed that the proposed method could effectively forecast the $PM_{2.5}$ value.

Qin et al. [35] introduced a new method for predicting pollutant concentration based on vast amounts of environmental data and deep learning techniques. The approach incorporated big data using two forms of deep networks. The method is based on a design that uses a convolutional neural network as the base layer, extracting input data features automatically. For the output layer, a long-term memory network was used to determine the time dependence of pollutants. With performance optimization, as a time series, the model was able to predict future concentrations of $PM_{2.5}$. Eventually, the outcomes of the forecasts were correlated with numerical model tests. The model's applicability and benefits were also analyzed. The experimental results indicated that it increased the accuracy of predictions compared with traditional models.

3 System design and architecture

3.1 TensorFlow

TensorFlow is an open-source code library that uses data flow graphs to express numerical operations. TensorFlow can help users easily construct models of machine learning and deep learning, reducing users' learning threshold [36]. TensorFlow also supports distributed computing and can run well on a variety of platforms. It can be supported from a single CPU to a multi-GPU system. In addition, it supports multiple programming languages. Both Python and C++ can be used to write TensorFlow programs.

3.2 Keras

Keras is a neural network library written entirely in Python and fully compatible with TensorFlow. Using Keras to write TensorFlow programs, building a neural network architecture will be simple and fast, allowing users to skip complex operations [18, 37]. The details of the complex module construction reduce the learning threshold in the field of machine learning. Keras supports most of the model solutions, supports multi-input and multi-output training, and can be used under the CPU and GPU system base. Its wide compatibility enables Keras to work on different operating systems such as Windows, Linux, and MacOS, and it can be used normally.

3.3 System design

Influenza is an acute respiratory disease that can easily lead to pneumonia and death. Early prediction and response to influenza epidemics is the most effective method of control. This paper intends to determine the relevance of $PM_{2.5}$ to influenza pneumonia death cases by using the government's open database to further build predictive modules through machine learning. The research method is to use the flu pandemic death database of the Department of Health and Welfare [38] and link to the EPA Air Quality Monitoring Network for $PM_{2.5}$ [39] open data to visualize the data in determining the correlation between $PM_{2.5}$ and influenza pneumonia death. The machine deep learning method is used for the prediction module of $PM_{2.5}$ causing death from influenza pneumonia, applied to the epidemic prevention policy. Figure 3 depicts the system design.

- Step 1: National Health Insurance Research Database and Air Quality Monitoring Network

The diagnostic data for *Aspergillus* infection was analyzed by Taichung Veteran General Hospital [28] with the National Health Insurance Research Database (NHIRD). AQI and $PM_{2.5}$ data obtained from the Environmental Protection Agency Air Quality Monitoring Network in Taiwan, combined with the above long-term big data, was used to prove the correlation between air pollution and respiratory diseases.

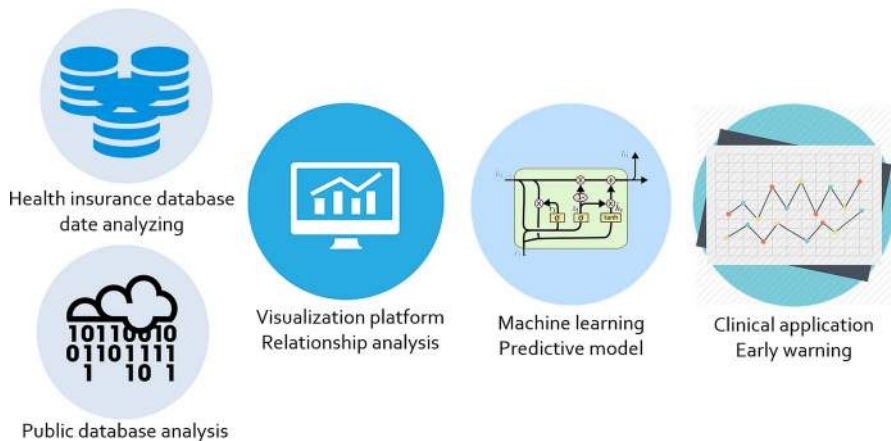


Fig. 3 System design

- Step 2: Use of visual data to build a visualization system and monitoring platform
Information on the rate of weekly emergency visits due to influenza-like illness, pneumonia, and enterovirus was gathered from the Department of Health and Welfare. This information was then combined to establish an associated visualization system and monitoring platform for air pollution and disease.
- Step 3: Establishing a predictive model
The 95% (confidence interval) upper limit was calculated as the prevalence threshold using the rate of weekly visits for the past 5 years. The LSTM model was used to predict the rate of visits in the next 4 weeks, and the threshold was used as an early warning.

3.4 System architecture

The architecture of the present study is shown in Fig. 4. The data set is collected from Taiwan Ministry of Health and Welfare. The data consist of influenza emergency department visits and the EPD data rate. $PM_{2.5}$ open air quality monitoring network data is stored in a MySQL database, and preliminary data cleaning is completed. The LSTM model is then built and trained through the TensorFlow library and using Keras' high-level tools, and the results of future short-term predictions are passed back to the MySQL repository. Finally, the data is visually presented on the web page by linking MySQL and PHP.

This study uses a linear method and a k -nearestneighbor (KNN) method, two methods for obtaining missing values. The linear compensation method is effective at filling missing values of fragmented time. However, when long-term continuous time values are missing, the linear complement method cannot accurately fill in missing values. Therefore, this study adopts two complementary methods of linear and KNN in parallel. If the missing value is fragmented, the linear complement method to is used fill in the missing value. Otherwise, if



Fig. 4 System architecture

the value is missing for more than 6 consecutive hours, the KNN complement method is used to fill in the value.

On the topic of data cutting, this study uses a ratio of 6:2:2 to cut the original data into three parts: the training set, the validation set, and the testing set. This arrangement is adopted in the hope that a part of the data can be completely excluded After training, one can use the testing set as brand-new data to objectively verify the model, and to determine whether the training effectiveness has reached the same level as the training data.

3.5 Provided services

This cloud system provides real-time air quality index information and population statistics of each city with various kinds of graphs on the dashboard. This data can be exported to Excel for further analysis. In addition, more detail of the AQI data formula is as follows:

- O_3 , 8 h: average of the last 8 h
- O_3 : real-time
- $PM_{2.5}$: $0.5 \times$ average of the last 12 h + $0.5 \times$ the last 4 h
- PM_{10} : $0.5 \times$ average of the last 12 h + $0.5 \times$ the last 4 h
- CO: average of the last 8 h
- SO_2 : real-time
- NO_2 : real-time

Table 1 describes the normal value index of ozone (O_3), fine aerosol ($PM_{2.5}$), aerosol (PM_{10}), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen dioxide (NO_2).

Table 1 The concentration of pollutants and air quality index value

Air poll.	O ₃ , 8 h	O ₃	PM _{2.5}	PM ₁₀	CO	SO ₂	NO ₂
Unit	ppm	ppm	μg/m ³	μg/m ³	ppm	ppb	ppb
0–50	0.000–0.054	–	0.0–15.4	0–54	0–4.4	0–35	0–53
51–100	0.055–0.070	–	15.5–35.4	55–125	4.5–9.4	36–75	54–100
101–150	0.071–0.085	0.125–0.164	35.5–54.4	126–254	9.5–12.4	76–185	101–360
151–200	0.086–0.105	0.165–0.204	54.5–150.4	255–354	12.5–15.4	186–304 (3)	361–649
201–300	0.106–0.200	0.205–0.404	150.5–250.4	355–424	15.5–30.4	305–604 (3)	650–1249
301–400(2)	–	0.405–0.504	250.5–350.4	425–504	30.5–40.4	605–804	1250–1649
401–500(2)	–	0.505–0.604	350.5–500.4	505–604	40.5–50.4	805–1004	1650–2049

Table 2 Software version

No.	Description	Version
1	Apache Hadoop	2.7.3
2	Cloudera	5.10
3	Apache Spark	2.0
4	ELK Stack	5.2
5	PHP	7.0
6	MySQL	7.0
7	Alluxio	1.4
8	Ceph	10.2.0 (Jewel)

3.6 System implementation

- The software deployment server that we use for experiments comprises five servers: one name node server, two data node servers, one ELK stack server, and one web service server. More detail for each server is presented in Table 2.
- Cluster deployment has become a well-known architecture design in the science of big data, because it can be more efficient in data processing. In this study, we created a Cloudera Manager cluster server.

Hadoop or Spark memory configuration recommendation

This part describes the way to configure YARN and MapReduce memory allocation settings depending on the hardware component specifications.

When we determine the appropriate YARN and MapReduce memory configurations for a cluster node, we have to start with the available hardware resources. The following are the values of each node:

- Random access memory (amount of memory)
- Cores (number of CPU cores)
- Disk storage (number of disks)

Table 3 Reserved memory recommendations

RAM per node (GB)	Rec. reserved memory (GB)	Rec. reserved Hbase RAM (GB)
4	1	1
8	2	1
16	2	2
24	4	4
48	6	8
64	8	8
72	8	8
96	12	16
128	24	24
256	32	32
512	64	64

General Formula Container

$$\#Container = \min(2 \times Cores, 1.8 \times Disks, (Total\ Available\ RAM) / Minimum\ Container\ Size) \quad (10)$$

If Hbase is included, the following formula is used:

$$\#Container = \min(2 \times Cores, 1.8 \times Disks, (Total\ available\ RAM - Reserved\ system\ memory - Reserved\ HBase\ memory) / Minimum\ container\ size) = \min(2 \times 8, 1.8 \times 1, (16 - 2 - 2) / 1) = (16, 1.8, 12) = 1.8 = 2$$

$$\#RAM\ per\ Container = \max(Minimum\ container\ size, (Total\ available\ RAM - Reserved\ system\ memory - Reserved\ HBase\ memory) / Containers) = \max(1, (16 - 2 - 2) / 2) = \max(1, 6) = 6$$

Reserved HBase memory is the RAM needed by system processes and HBase processes. The recommendation of reserve memory is described in Table 3.

Minimum container size recommendations is described in Table 4.

- **yarn.node.manager.resource.memory_mb:**
Containers \times RAM per container = $2 \times 6 = 12 \times 1024 = 12288$ MB
- **yarn.scheduler.minimum_allocation_mb:**
RAM per container = $6 \times 1024 = 6144$ MB
- **yarn.scheduler.maximum_allocation_mb:**
Containers \times RAM per container = $2 \times 1024 = 2048$ MB
- **mapreduce.map.memory.mb:**
RAM per container = $6 \times 1024 = 6144$ MB
- **mapreduce.reduce.memory.mb:**
 $2 \times$ RAM per container = $2 \times 6 \times 1024 = 12288$ MB
- **mapreduce.map.java.opts:**
0.8 times RAM per container = $0.8 \times 6 \times 1024 = 4915.2$ MB
- **mapreduce.reduce.java.opts:**
 $0.8 \times 2 \times$ RAM per container = $0.8 \times 2 \times 6 \times 1024 = 9830.4$ MB

Table 4 Minimum container size recommendations

RAM per node	Rec. minimum container size
Less than 4 GB	256 MB
Between 4 GB and 512 MB	2 GB
Between 8 GB and 1024 MB	2 GB
Above 24 GB	2048 MB

Table 5 YARN and MapReduce configuration setting value calculations

Configuration file	Configuration setting	Value calculation
yarn-site.xml	yarn.nodemanager.resource.memory-mb	= containers \times RAM per container
yarn-site.xml	yarn.scheduler.minimum-allocation-mb	= RAM per container
yarn-site.xml	yarn.scheduler.maximum-allocation-mb	= containers \times RAM per container
mapred-site.xml	mapreduce.map.memory.mb	= RAM per container
mapred-site.xml	mapreduce.reduce.memory.mb	= 2 \times RAM per container
mapred-site.xml	mapreduce.map.java.opts	= 0.8 \times RAM per container
mapred-site.xml	mapreduce.reduce.java.opts	= 0.8 \times 2 \times RAM per container
mapred-site.xml	yarn.app.mapreduce.am.resource.mb	= 2 \times RAM per container
mapred-site.xml	yarn.app.mapreduce.am.command-opts	= 0.8 \times 2 \times RAM per container

- **yarn.app.mapreduce.am.resource.mb:**
 $2 \times \text{RAM per Container} = 2 \times 6 \times 1024 = 12288 \text{ MB}$
- **yarn.app.mapreduce.am.command.opts:**
 $0.8 \times 2 \times \text{RAM per container} = 0.8 \times 2 \times 6 \times 1024 = 9830.4 \text{ MB}$

YARN and MapReduce configuration setting value calculations are showed in Table 5.

This Cloudera Manager is used to build related package services like Hadoop, HDFS, MapReduce, Spark, HBase, and Hive. The website can be used to monitor CPU usage, disk I/O, network I/O, and HDFS status.

3.7 Data preprocessing

This study collected information on the rate of influenza-like emergency visits from the Ministry of Health and Welfare and the PM_{2.5} Air Quality Monitoring Network. The total number in the six districts was more than 3600, collecting weekly averages from 2007 to 2019. The current data set is still in continuous collection in real time. This work uses 6:2:2 to divide the original data into three parts: training set, test set, and verification set, respectively. After the training is done, the test and verification sets are used as new data to verify the module in order to determine whether the training effect is correct. This method is used to

avoid over-fitting on the data set and underestimating the module performance evaluation data.

3.8 The outbreak calculation method

The outbreak calculation is based on the Outbreaks calculation method published by the World Health Organization [40]. Historical data of the past 5 years can be used to obtain the threshold of the influenza-like visit data for the current year. This means that the current influenza-like epidemic can be obtained through this calculation method. The criteria for judging and the data will be added to the data set of the training model, and at the same time, when the web page is visualized, it is the most direct criterion for the user to judge the current epidemic. The Outbreaks formula is as follows:

$$X \pm Z \frac{s}{\sqrt{n}} \quad (11)$$

where X is the average and Z is the confidence level coefficient. Taking this study as an example, when the confidence level is 95%, Z is 1.96, s is the standard deviation, and n is the sample size.

3.9 The LSTM model

In order to effectively predict influenza-like trends, this study will construct the LSTM model and use the air pollution index $PM_{2.5}$ and influenza-like visit rates. Outbreaks calculates values as supplementary data to train the model. There is no correct answer to the neural network model design, which must be determined according to the researcher's hardware environment, data characteristics, data set size, and user response time requirements. If the number of neural layers or the number of neurons are blindly increased, it may make the training effect worse, and also waste a lot of time and computing efficiency.

This research uses trial and error to find the best model architecture. The content includes the number of neural layers, the number of neurons, the design of the input and output layers, the activate function, and so on. The judgement criterion is of the highest accuracy. The MAPE evaluation index is used to select the architecture parameter with the lowest error. Therefore, after many experiments and adjustments, it was decided to set the membrane structure to four input layer neurons, four output layer neurons, 256 neurons in two layers of LSTM layers, and 128 hidden layers in two layers. Sixty-four neurons, that is, four future time units are predicted in four consecutive time units.

The input layer is set to four neurons according to trial and error results, which represents four data as an input for a single training. The output layer neurons are set to four according to the needs of the topic, which represents the prediction of the rate of ILI consultations in the next 4 weeks. The activate function of the output layer uses linear as an approximation prediction of the true value. The structure of

the input data set is the historical ILI consultation rate and $PM_{2.5}$ value for 4 weeks, and the data dimension is (2208, 4, 2). The output data is structured to predict the ILI consultation rate for the next 4 weeks. The data dimension is (2208, 4).

In order to avoid the problem of overfitting, the early stopping mechanism and the design of the dropout layer are added to the model. Early stopping monitors the error value of the training during the training, if it is under a certain time. If the error value does not continue to be effectively reduced, the training will be terminated early, and this method can be used to avoid over-learning. In addition to the addition of the dropout layer design, the dropout layer randomly stops the neurons of some hidden layers temporarily during training. Since the neurons that are paused during each training are randomly selected, it can be imagined. Each training is on a new model, and since the relevant neurons cannot appear at the same time each time, it is also possible to avoid the fact that there are joint features that are only retained when they appear with each other.

3.10 Model performance evaluation

After completion of the neural network module training, the model evaluation is carried out to understand the effectiveness of the model learning, which serves as the basis for the experiment and is also an important evaluation criterion for the model performance. In this study, the mean absolute percentage error (MAPE) was chosen as the criterion for the prediction performance of this module. The formula for the MAPE value is as follows:

$$\text{MAPE} = \left(\sum_{i=1}^n \frac{(|Y_i - \hat{Y}_i|)}{Y_i} \right) \cdot 100. \quad (12)$$

3.11 Web visualization

This study presents data analysis results through web pages. In the operation and display of web pages, the system is constructed based on the current mainstream HTML5 + JavaScript + CSS3. Users can connect web pages with multiple devices. In addition, the concept of responsive web design (RWD) and a bootstrap framework are added to the interface design. In this case, the implementation of a bootstrap grid system interface is to make the system interface work at different resolutions. With this environment running on the device, including on smart mobile devices and tablets, a good user experience can be achieved.

Period	Year	Week	PM25_Whole	PM25_Taipei	PM25_North	PM25_Central	PM25_South
202006	2020	6	15	10	11	16	20
202005	2020	5	15	14	15	14	17
202004	2020	4	26	21	25	29	34
202003	2020	3	17	13	14	17	21
202002	2020	2	27	17	24	29	32
202001	2020	1	16	9	12	18	18

Fig. 5 Schematic diagram of PM_{2.5} data

4 Experimental results

In this section, we describe our experimental results obtained from the designed methods.

Period	Year	Week	Whole	Taipei	North	Central	South
202006	2020	6	15.46	16.46	17.02	14.22	14.69
202005	2020	5	22.22	23.18	24.93	21.31	19.71
202004	2020	4	16.35	17.28	17.66	15.37	15.76
202003	2020	3	12.8	12.89	14.91	12.35	12.81
202002	2020	2	13.49	13.59	15.88	13.04	12.95
202001	2020	1	13.59	13.88	16.05	12.97	13.26

Fig. 6 Schematic diagram of ILI data in six districts

4.1 Data collection

The data set is obtained from the weekly average influenza-like emergency visit rate information of the six major subdistricts of Taiwan. The PM_{2.5} open data is collected from the Environmental Protection Department's air quality monitoring network with 77 air pollution stations. The data is organized into six district average weekly data sets that are consistent with the needs of the present study. The MySQL database stores these data set with a specific data format, as shown in Figs. 5 and 6.

This study also organizes historical air pollution data into air quality index (AQI), which is set by the Taiwan Environmental Protection Agency. The data set consists of ozone (O³), fine aerosol (PM_{2.5}), suspended particulates (PM₁₀), carbon monoxide (CO), sulfur dioxide (SO²), and nitrogen dioxide (NO²) in the air. The degree of health impact is converted into the sub-indicator values of different pollutants, and the maximum value is taken as the AQI value for the day. This study uses a full six-zone AQI value and ILI binding to influenza treatment. A line graph is displayed

on a web page system, as shown in Fig. 7, for a user to clearly observe the real-time AQI value and the relationship of AQI and ILI value. From the graph, one can observe the association of AQI and ILI weekly status in six areas: Taipei, North, Central, South, Kaoping, and East. The maximum value index indicator of AQI also can be shown in the balloon symbol.

4.2 Training result

In this section, the training module process is presented to visualize the loss value data. It will help researchers to understand whether the module training process has smooth convergence, and whether there is an excessive situation or not. Figure 8 shows the change of the loss value of the LSTM modules in the six districts. The six-zone model shows good performance during the training process. The training and verification loss values are consistent. There is no indication of over-fitting, and a stable error value is reached at the specified number of training times.

In addition, the training set, the test set, and the verification set are respectively implemented in the module to verify the prediction effect of the module. The central area data is taken as an example, and the actual values of the three data sets are respectively taken. The visualized graphs display the predicted values, as shown in Figs. 9, 10, and 11. The training effect of the observation module and the data set with the training can have the same prediction performance.

4.3 The evaluation results

This study selected MAPE as the standard for model evaluation, and MAPE values of the six major divisions of the whole station according to the planning of this

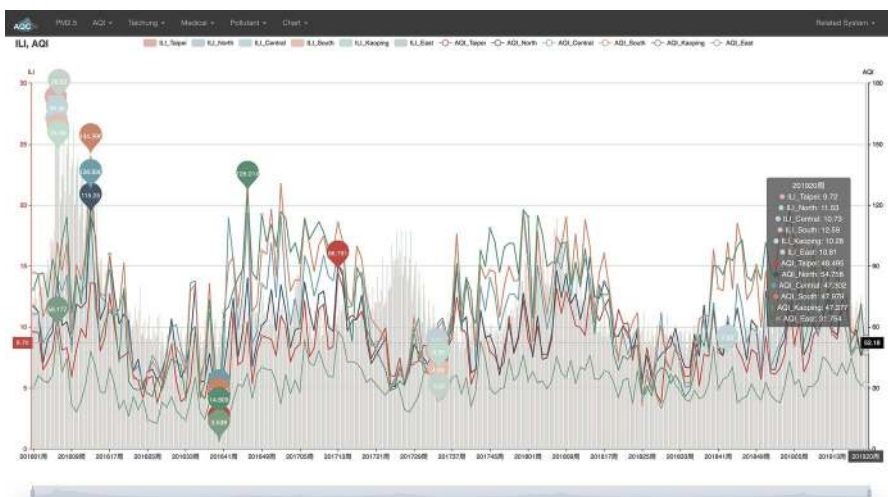


Fig. 7 AQI, ILI line chart

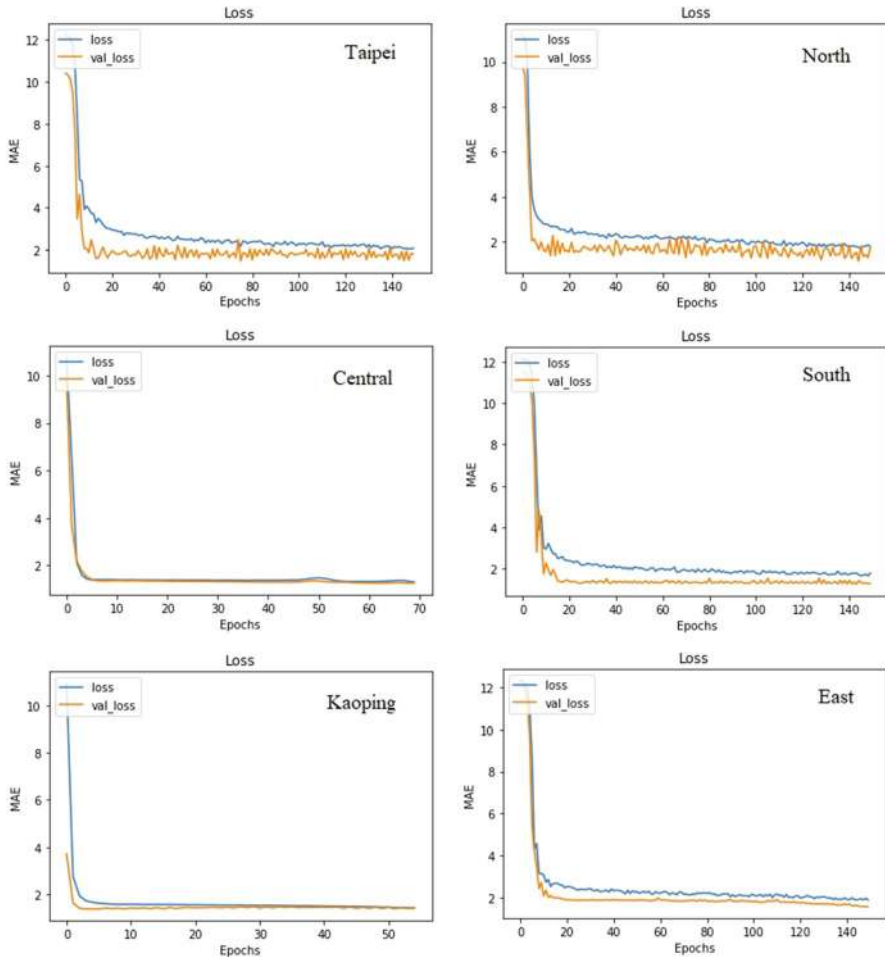


Fig. 8 The change of the loss value of the six-zone model

study were calculated. The model predicts the next 4 h to calculate the individual MAPE values in various hours. This calculation method can more clearly convey the change of model accuracy under the shorter and longer prediction time gap. Table 6 summarized the MAPE value comparison based on each area in Taiwan.

In this study, the LSTM model was better than the RNN model in predicting ILI consultation rates. The central area data for ILI consultation rate data were selected for training and to compare the effects of model predictions. Three model evaluation indicators, MAPE, RMSE, and MAE, are used to calculate the model errors, with results summarized in Table 7. According to the contents of the table, the evaluation results of the LSTM model have errors that are lower than those of the RNN model from the first week to the sixth week. This experiment confirms that LSTM has more advantages than RNN for the issue of ILI consultation rate prediction.

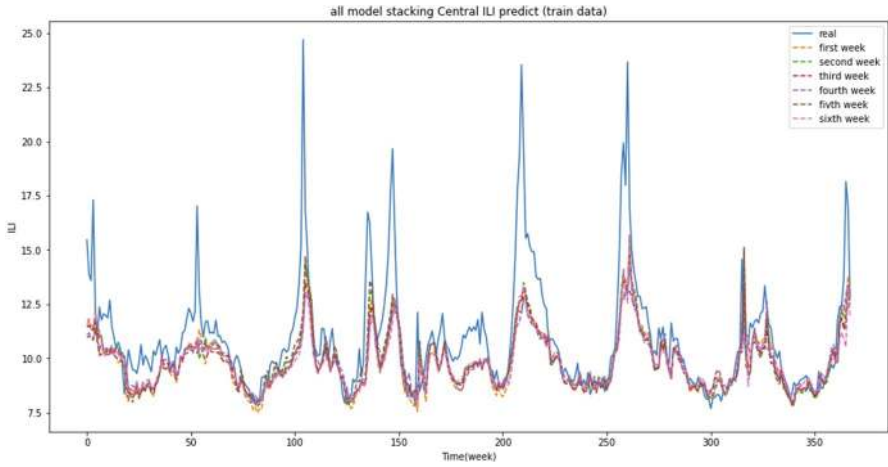


Fig. 9 The plotting of training

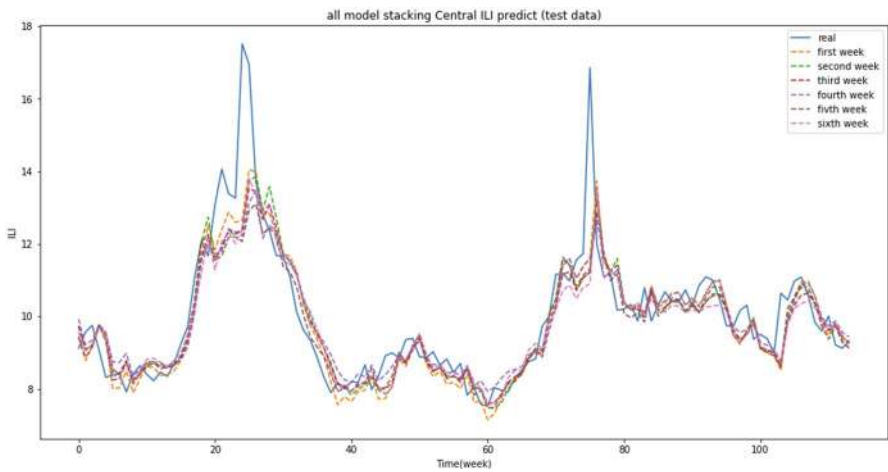


Fig. 10 The plotting of testing

4.4 Outbreak calculation method

The Outbreaks value and influenza-like treatment rate value are visualized on the web system to determine whether the influenza-like disease control status is still in a safe state. Figures 12 and 13 present the ILI visit rate and Outbreaks value for 2017 and 2018, respectively.

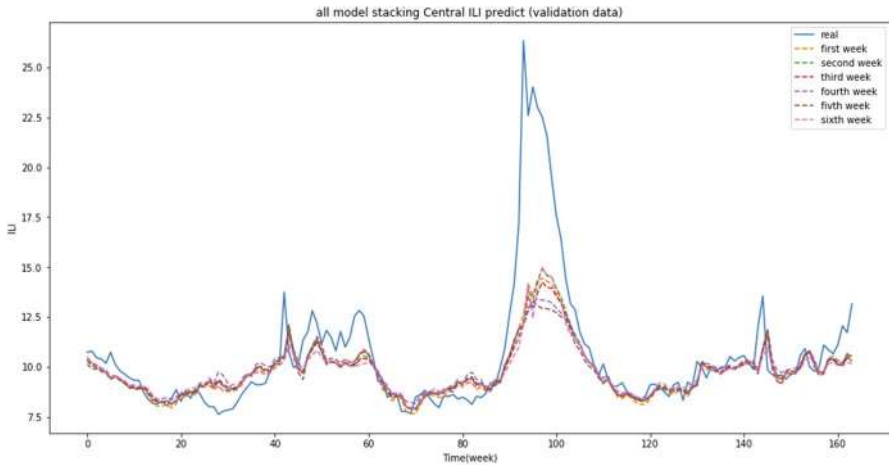


Fig. 11 The plotting of validation

Table 6 MAPE value comparison

Region	Data set	First hour	Second hour	Third hour	Fourth hour
Taipei	Training set	9.64	9.83	10.7	12.66
	Validation set	11.30	12.87	13.88	14.08
North	Training set	11.21	12.26	13.23	14.8
	Validation set	18.92	14.19	14.93	16.95
Central	Training set	7.97	9.45	10.12	11.08
	Verification set	10.48	11.76	12.62	13.56
South	Training set	6.02	7.55	8.47	9.39
	Verification set	9.61	10.77	10.8	11.08
Kaoping	Training set	8.92	9.76	10.74	11.87
	Verification set	9.48	10.79	11.25	12.31
East	Training set	9.15	10.53	11.82	12.97
	Verification set	12.98	14.82	15.02	15.6

Table 7 Comparison of central area model evaluation

Model	Evaluation	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
RNN	MAPE	0.479861111	0.513194444	0.566666667	13.48	14.43	14.04
LSTM		10.05	0.471527778	0.464583333	12.04	13.35	14.07
RNN	RMSE	02.33	02.49	0.128472222	0.134722222	0.146527778	0.150694444
LSTM		02.14	02.03	02.33	02.49	0.129861111	0.140277778
RNN	MAE	01.47	01.58	0.091666667	0.096527778	0.104861111	0.105555556
LSTM		01.36	01.45	01.45	01.58	0.092361111	0.099305556

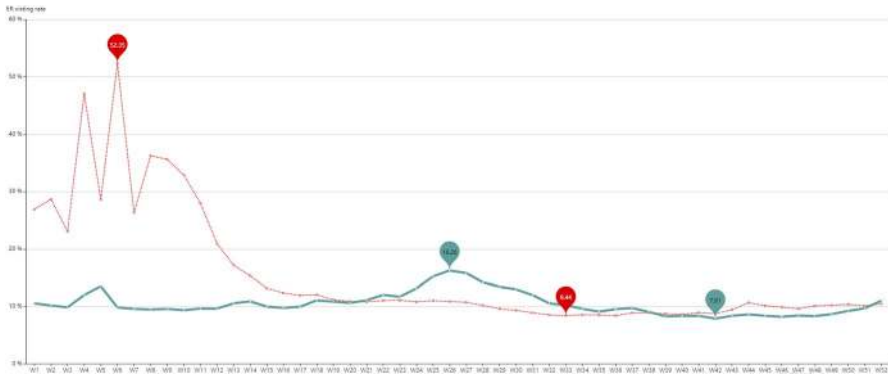


Fig. 12 2017 ILI visit rate and Outbreak threshold

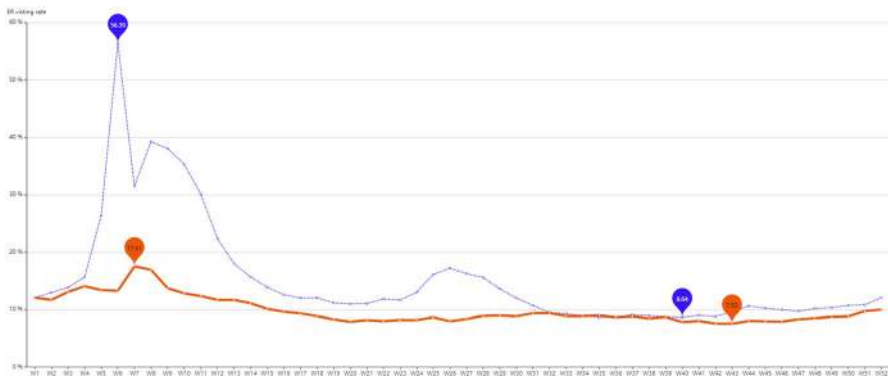


Fig. 13 2018 ILI visit rate and Outbreak threshold

4.5 Web visualization

Figure 14 presents the Outbreaks value in 2019, in which the middle region demonstrates the actual influenza-like visit rate; the next 4 weeks of influenza-like visits are presented in predicted values. Figure 15 shows the actual influenza-like visit rate in the six districts of the country and the predicted value of the next 4 weeks of influenza-like visits. This visualization allows users to know the future trend of influenza-like epidemics and whether it will occur or not. It can forecast the possibility of an ILI outbreak.

Figure 16 demonstrates the real-time monitoring system per area.

Figure 17 shows the real-time ILI time-series data.

The web visualization also has an additional feature to give the early warning. If the predicted visit rate is more than the threshold, then it is an early warning of influenza outbreak. Figure 18 describes the graph of an early warning system.

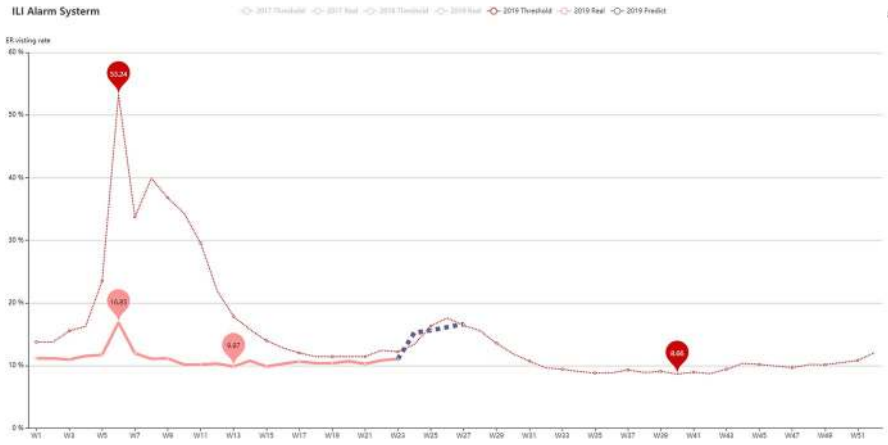


Fig. 14 Real-time ILI forecast data visualization

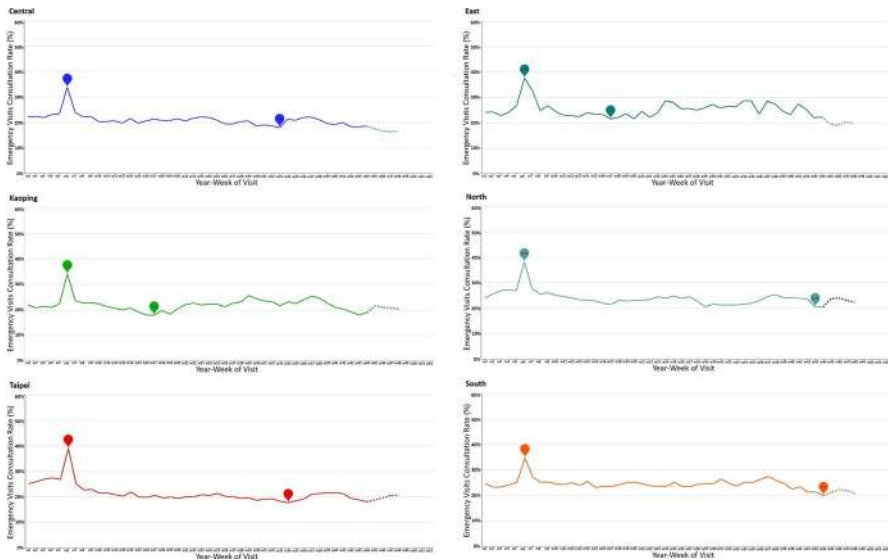


Fig. 15 Real-time ILI forecast data visualization for six districts

5 Conclusion and future work

This study collected data on air pollution and rates of doctor visitation due to ILI. After the air pollution and ILI treatment rate data were integrated with quality control and time units applied, the data from 2007 to the present were divided into weekly average values for six regions of Taiwan, and the Outbreaks algorithm was used to calculate the influenza-like epidemic. A threshold value was used as the auxiliary data for training and the decision basis for the user to

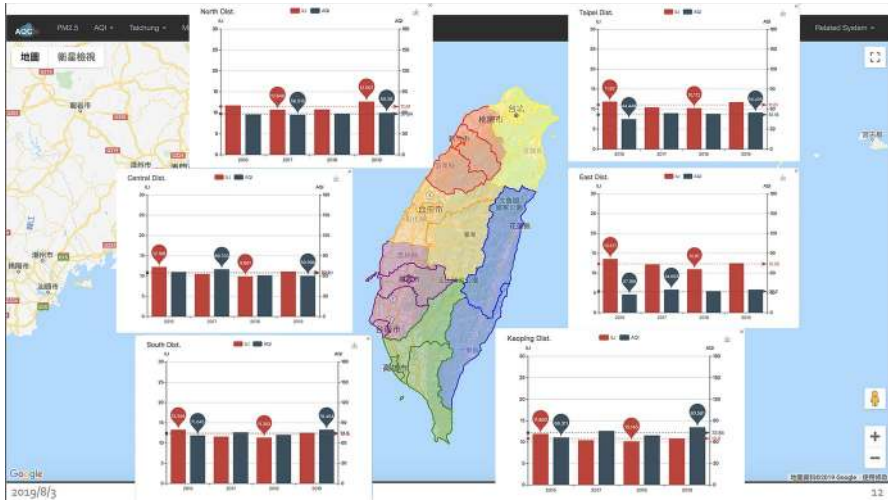


Fig. 16 Real-time monitoring system per area



Fig. 17 Real-time ILI time-series data

read the data. It also solved the problem with the large differences between the influenza-like out-of-control standard in the outbreak period and the non-perfect period, which can easily cause misunderstanding. The flu-like epidemic provides a more comprehensive and accurate understanding. We generated a simple linear regression model to estimate the correlation coefficient between AQI and ILI in previous papers. According to the model, we observed a significant correlation between $PM_{2.5}$ and risk of ILI (p value = 0.04) after adjusting for confounders.

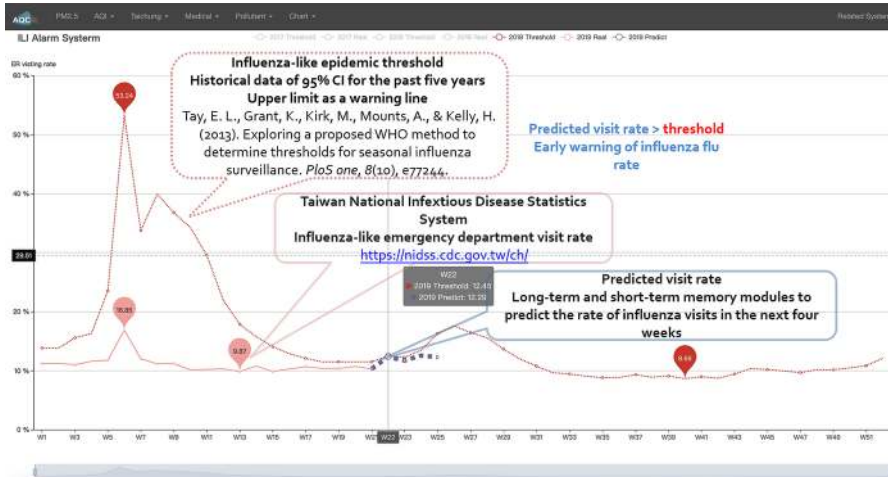


Fig. 18 Early warning of influenza outbreaks

This paper also implemented deep learning techniques to find correlations between data. The TensorFlow library and Keras tools were applied to construct neural network models for experiments. Long-term and short-term memory models (LSTM) were used to deal with time-series problems occurring in time-series data. The analysis of the influenza-like epidemic demonstrated very good performance based on the experiments. The analysis obtained a valid prediction of the rate of influenza-like visits in the next 4 weeks. The MAPE values for the module evaluation were all below 20, which is a valid result.

The web page system platform was built as a visual presentation interface. The integration of the AQI and the ILI treatment data demonstrated an interactive relationship between the two trends. The historical data of influenza-like visits and the Outbreaks threshold were graphically presented to observe the past epidemic control status of ILI, the yearly Outbreaks threshold, and real-time ILI visit rate data plus forecasts for the next 4 weeks. It showed the estimation of future flu-like epidemics and whether there would be an outbreak epidemic. This information will provide government agencies, medical-related units, and general users a clear and concise data reference for epidemic trends. The concept of responsive web design (RDD) was added to the web system design, and therefore the system interface can be connected with devices of different resolution, including smart mobile devices and tablets, and all users can achieve a good user experience.

In the future, it is possible to predict influenza-like epidemic trends with shorter historical data without sacrificing too much accuracy. In other words, it is possible to more rapidly predict outbreaks, and so the system can aid in achieving a faster response. This system can be combined with meteorological data, in which the daily maximum temperature difference and humidity data are highly correlated with influenza induction. In addition, past outbreaks of influenza have often been accompanied by variants of the influenza virus. The causes of recent pandemics have been

related to pigs, influenza, and avian flu mutations. Combining this data with data regarding poultry epidemics, such as rates of culling rate or diagnosis, as well as poultry transport flow data, might be helpful for early prediction of flu occurrence and the risk of spreading. Moreover, comparative experiments of similar algorithms can also be performed in the future.

Acknowledgements This work was funded by the Ministry of Science and Technology (MOST), Taiwan, under Grant Number MOST 108-2119-M-029-001-A. In addition, this work was also funded in part by the Taichung Veterans General Hospital (TCVGH), Taiwan, under Grant Nos. TCVGH-T1077803 and TCVGH-T1087804.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References


1. Kaur P, Sharma M (2019) Diagnosis of human psychological disorders using supervised learning and nature-inspired computing techniques: a meta-analysis. *J Med Syst* 43(7):204
2. Kim JE, Dager SR, Jeong HS, Ma J, Park S, Kim J, Cho HB (2018) Firefighters, posttraumatic stress disorder, and barriers to treatment: results from a nationwide total population survey. *PLoS ONE* 13(1):e0190630
3. Gautam R, Kaur P, Sharma M (2019) A comprehensive review on nature inspired computing algorithms for the diagnosis of chronic disorders in human beings. *Prog Artif Intell* 8(4):401–424. <https://doi.org/10.1007/s13748-019-00191-1>
4. Van der Fels-Klerx HJ, Van Asselt ED, Raley M, Poulsen M, Korsgaard H, Bredsdorff L, Frewer LJ (2018) Critical review of methods for risk ranking of food-related hazards, based on risks for human health. *Crit Rev Food Sci Nutr* 58(2):178–193
5. Lau JT, Griffiths S, Choi K-C, Lin C (2010) Prevalence of preventive behaviors and associated factors during early phase of the H1N1 influenza epidemic. *Am J Infect Control* 38:374–380. <https://doi.org/10.1016/j.ajic.2010.03.002>
6. Croft DP, Zhang W, Lin S, Thurston SW, Hopke PK, Masiol M, Squizzato S, van Wijngaarden E, Utell MJ, Rich DQ (2019) The association between respiratory infection and air pollution in the setting of air quality policy and economic change. *Ann Am Thorac Soc* 16(3):321–330
7. Croft DP, Zhang W, Lin S, Thurston SW, Hopke PK, van Wijngaarden E, Squizzato S, Masiol M, Utell MJ, Rich DQ (2019) Associations between source-specific particulate matter and respiratory infections in New York State adults. *Environ Sci Technol*. <https://doi.org/10.1021/acs.est.9b04295>
8. Hopke PK, Croft D, Zhang W, Lin S, Masiol M, Squizzato S, Thurston SW, van Wijngaarden E, Utell MJ, Rich DQ (2019) Changes in the acute response of respiratory diseases to PM_{2.5} in New York State from 2005 to 2016. *Sci Total Environ* 677:328–339
9. Strickland MJ, Hao H, Hu X, Chang HH, Darrow LA, Liu Y (2016) Pediatric emergency visits and short-term changes in PM_{2.5} concentrations in the U.S. State of Georgia. *Environ Health Perspect* 124(5):690–696
10. Weichenthal SA, Lavigne E, Evans GJ, Godri Pollitt KJ, Burnett RT (2016) Fine particulate matter and emergency room visits for respiratory illness. Effect modification by oxidative potential. *Am J Respir Crit Care Med* 194(5):577–586

11. Horne BD, Joy EA, Hofmann MG, Gesteland PH, Cannon JB, Lefler JS, Blagev DP, Korgenski EK, Torosyan N, Hansen GI, Kartchner D, Pope CA 3rd (2018) Short-term elevation of fine particulate matter air pollution and acute lower respiratory infection. *Am J Respir Crit Care Med* 198(6):759–766
12. Darrow LA, Klein M, Flanders WD, Mulholland JA, Tolbert PE, Strickland MJ (2014) Air pollution and acute respiratory infections among children 0–4 years of age: an 18-year time-series study. *Am J Epidemiol* 180(10):968–977
13. Pirozzi CS, Jones BE, VanDerslice JA, Zhang Y, Paine R 3rd, Dean NC (2018) Short-term air pollution and incident pneumonia. A case-crossover study. *Ann Am Thorac Soc* 15(4):449–459
14. Jones RR, Hogrefe C, Fitzgerald EF, Hwang SA, Özkaynak H, Garcia VC, Lin S (2015) Respiratory hospitalizations in association with fine PM and its components in New York State. *J Air Waste Manag Assoc* 65(5):559–569
15. Peters A, Breitenr S, Cyrus J, Stölzel M, Pitz M, Wölke G, Heinrich J, Kreyling W, Küchenhoff H, Wichmann HE (2009) The influence of improved air quality on mortality risks in Erfurt, Germany. *Res Rep Health Eff Inst* 137:5–77 (**discussion 79–90**)
16. Yang CT, Chen ST, Den W, Wang YT, Kristiani E (2019) Implementation of an intelligent indoor environmental monitoring and management system in cloud. *Future Gener Comput Syst* 96:731–749. <https://doi.org/10.1016/j.future.2018.02.041>
17. Feng C, Li J, Sun W, Zhang Y, Wang Q (2016) Impact of ambient fine particulate matter (PM 2.5) exposure on the risk of influenza-like-illness: a time-series analysis in Beijing, China. *Environ Health* 15(1):17. <https://doi.org/10.1186/s12940-016-0115-2>
18. Zhou Y, Chang F-J, Chang L-C, Kao I-F, Wang Y-S (2019) Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *J Clean Prod* 209:134–145. <https://doi.org/10.1016/j.jclepro.2018.10.243>
19. Hwang K, Sung W (2017) Online sequence training of recurrent neural networks with connectionist temporal classification. Department of Electrical and Computer Engineering Seoul National University. [arXiv:1511.06841](https://arxiv.org/abs/1511.06841)
20. Cinar YG, Mirisae H, Goswami P, Gaussier E, Ait-Bachir A (2018) Period-aware content attention RNNs for time series forecasting with missing values. *Neurocomputing* 312:177–186. <https://doi.org/10.1016/j.neucom.2018.05.090>
21. Graves A (2014) Generating sequences with recurrent neural networks. Department of Computer Science, University of Toronto. [arXiv:1308.0850](https://arxiv.org/abs/1308.0850)
22. Sutskever I, Vinyals O, Le QV (2014) Sequence to sequence learning with neural networks. Google. <http://papers.nips.cc/paper/5346-sequence-to-sequence-learning-with-neural-networks>. Accessed 20 July 2019
23. Kim T-Y, Cho S-B (2018) Web traffic anomaly detection using C-LSTM neural networks. *Expert Syst Appl* 106:66–76. <https://doi.org/10.1016/j.eswa.2018.04.004>
24. Li X, Peng L, Yao X, Cui S, Hu Y, You C, Chi T (2017) Long-short term memory neural network for air pollutant concentration predictions: method development and evaluation. *Environ Pollut* 231:997–1004. <https://doi.org/10.1016/j.envpol.2017.08.114>
25. Wen C, Liu S, Yao X, Peng L, Li X, Hu Y, Chi T (2019) A novel spatiotemporal convolutional long-short term neural network for air pollution prediction. *Sci Total Environ* 654:1091–1099. <https://doi.org/10.1016/j.scitotenv.2018.11.086>
26. Interpret all statistics and graphs for trend analysis. <https://support.minitab.com/en-us/minitab-express/1/help-and-how-to/modeling-statistics/time-series/how-to/trend-analysis/interpret-the-results/all-statistics-and-graphs/>. Accessed Date 25 March 2019
27. Yang CT, Chen CJ, Tsan YT, Liu PY, Chan YW, Chan WC (2018) An implementation of real-time air quality and influenza-like illness data storage and processing platform. *Comput Hum Behav*. <https://doi.org/10.1016/j.chb.2018.10.009>
28. Liu PY, Tsan YT, Chan YW, Chan WC, Shi ZY, Yang CT, Lou BS (2018) Associations of PM_{2.5} and aspergillosis: ambient fine particulate air pollution and population-based big data linkage analyses. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-018-0852-x>
29. Lee CF, Yang CT, Kristiani E, Tsan YT, Chan WC, Huang CY (2018) Recurrent neural networks for analysis and automated air pollution forecasting. In: International Conference on Frontier Computing. Springer, Singapore, pp 50–59. <https://doi.org/10.1007/978-981-13-3648-5-6/>
30. Tang S, Yan Q, Shi W, Wang X, Sun X, Yu P, Xiao Y (2018) Measuring the impact of air pollution on respiratory infection risk in China. *Environ Pollut* 232:477–486. <https://doi.org/10.1016/j.envpol.2017.09.071>

31. Liu L, Han M, Zhou Y, Wang Y (2018). LSTM recurrent neural networks for influenza trends prediction. In: International symposium on bioinformatics research and applications. Springer, Cham, pp 259–264. <https://doi.org/10.1007/978-3-319-94968-0-25>
32. Zhang J, Nawata K (2018) Multi-step prediction for influenza outbreak by an adjusted long-short term memory. *Epidemiol Infect* 146(7):809–816. <https://doi.org/10.1017/S0950268818000705>
33. Huang CJ, Kuo PH (2018) A deep CNN-LSTM model for particulate matter (PM_{2.5}) forecasting in smart cities. *Sensors* 18(7):2220
34. Tsai YT, Zeng YR, Chang YS (2018) Air pollution forecasting using RNN with LSTM. In: 2018 IEEE 16th International Conference on Dependable, Autonomic and Secure Computing, 16th International Conference on Pervasive Intelligence and Computing, 4th International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech). IEEE, pp 1074–1079
35. Qin D, Yu J, Zou G, Yong R, Zhao Q, Zhang B (2019) A novel combined prediction scheme based on CNN and LSTM for urban PM 2.5 concentration. *IEEE Access* 7:20050–20059
36. Xingjian SHI, Chen Z, Wang H, Yeung DY, Wong WK, Woo WC (2015) Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: Advances in neural information processing systems, pp 802–810. <http://papers.nips.cc/paper/5955-convolutional-lstm-network-a-machine-learning-approach-for-precipitation-nowcasting>. Accessed 20 July 2019
37. Pascanu R, Mikolov T, Bengio Y (2013). On the difficulty of training recurrent neural networks. In: International Conference on Machine Learning, pp 1310–1318. <http://proceedings.mlr.press/v28/pascanu13.pdf>. Accessed 20 July 2019
38. Taiwan National Infectious Disease Statistics System. <https://nidss.cdc.gov.tw/en/>. Accessed Date 20 Jan 2019
39. Taiwan Environment Protection Administration. <https://taqm.epa.gov.tw/taqm/tw/default.aspx>. Accessed Date 20 Jan 2019
40. Tay EL, Grant K, Kirk M, Mounts A, Kelly H (2013) Exploring a proposed WHO method to determine thresholds for seasonal influenza surveillance. *PLoS ONE* 8(10):e77244. <https://doi.org/10.1371/journal.pone.0077244>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Chao-Tung Yang¹ · Yuan-An Chen¹ · Yu-Wei Chan² · Chia-Lin Lee^{3,4,5} ·
Yu-Tse Tsan^{6,7}  · Wei-Cheng Chan⁶ · Po-Yu Liu⁸

Chao-Tung Yang
ctyang@thu.edu.tw

Yuan-An Chen
a961309@gmail.com

Yu-Wei Chan
ywchan@gm.pu.edu.tw

Chia-Lin Lee
u502107@yahoo.com.tw

Wei-Cheng Chan
fufu110017@gmail.com

Po-Yu Liu
pyliu@vghtc.gov.tw

¹ Department of Computer Science, Tunghai University, Taichung 40704, Taiwan, ROC

- ² College of Computing and Informatics, Providence University, Taichung 43301, Taiwan, ROC
- ³ Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan
- ⁴ Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan
- ⁵ Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan
- ⁶ Department of Emergency Medicine, Taichung Veterans General Hospital, Taichung 40704, Taiwan, ROC
- ⁷ School of Medicine, Chung Shan Medical University, Taichung 40704, Taiwan, ROC
- ⁸ Department of Internal Medicine, Taichung Veterans General Hospital, Taichung 40704, Taiwan, ROC