



A PREDICTION MODEL FOR INFLUENZA EPIDEMICS USING ARTIFICIAL NEURAL NETWORKS

Sakorn Mekruksavanich

Department of Computer Engineering

School of Information and Communication Technology

University of Phayao

Thailand

e-mail: sakorn.me@up.ac.th

Abstract

Influenza is regarded as the most common disease in the world. Although some influenza patients recover from fever and other side effects within a week without requiring medical attention, the illness can cause serious sickness or death in others, especially children and the elderly. Information technology now plays an important role in the development of effective and comprehensive approaches to cope with epidemic diseases in humans. The development of artificial intelligence has greatly expanded the use of neural networks in the epidemiological field. Thus, this study examined the use of artificial neural networks (ANNs) to learn the historical patterns of disease incidence and forecast future occurrence. Experiments using the proposed model showed accurate prediction results. These results promoted the advantages of neural networks for supporting decision makers in developing long term strategies regarding the number of disease incidences. A web-based medical application, based on this architectural prediction model was developed. The model will assist in the planning and management of suitable strategies to reduce the number of influenza cases.

Received: September 22, 2015; Accepted: October 12, 2015

Keywords and phrases: artificial neural networks, influenza epidemics, prediction model.

1. Introduction

Infectious diseases are disorders caused by organisms, for instance, bacteria and viruses. These simple organisms live in and on human bodies. They are normally harmless and can even be helpful, however, under certain conditions, some organisms may cause disease. Influenza, one of the most common communicable illnesses, is a highly contagious airborne disease that occurs in seasonal epidemics throughout the world. This infectious disease is caused by the influenza virus which spreads easily from infected people to others by airborne infection (coughing, sneezing, etc.). Influenza is responsible for over 500,000 deaths every year [1].

According to the Bureau of Epidemiology, Ministry of Public Health, 65,000 people suffered from influenza in Thailand during 2014. Influenza epidemics can seriously affect all populations, but the highest risk of complications occurs among children under two years, adults over 65, pregnant women, and people of any age with certain medical conditions, such as chronic heart, lung, kidney, liver, blood or metabolic diseases, or weakened immune systems. Influenza is usually characterized by a sudden onset of high fever, cough, headache, muscle and joint pain, sore throat and a runny nose. The cough can be severe and persist for two or more weeks. Most people recover from the fever and other indications within a week without requiring medical attention. However, influenza can cause serious sickness or even death, especially in citizens at high risk.

An artificial neural network (ANN), a subfield of machine learning, provides a very general way of approaching many problems. When the output networks have continuous values, an ANN calculates predictions and with discrete values, a classification is performed instead. There has been much research worldwide on the development of neural networks for medical diagnosis [2, 3]. ANNs can be used to increase the accuracy and objectivity of forecasting disease. Furthermore, ANN research and its applications have been studied for five decades.

During the last ten years, ANNs have been broadly studied for influenza management, producing an increasing number of publications each year. Trtica-Majnaric et al. [4] designed a model to enable the successful prediction of the outcome of influenza vaccination, based on real historical medical data. Kou et al. [5] used ANNs to distinguish the stereotype of influenza A-type virus, based on the antigen reaction of hemagglutinin (HA) and neuraminidase (NA) with their antibodies. A novel technique for the stereotype classification of the influenza A virus isolated from humans was also proposed. Attaluri et al. [6] applied the technique of artificial neural networks for the prediction of important influenza virus antigenic types (H1, H3 and H5) and their hosts (human, avian and swine). This fulfilled a critical need for a computational system for influenza surveillance. However, an effective model of ANNs in influenza recurrence prediction was not appropriately employed.

Therefore, this study aimed to develop a prediction model for influenza disease outbreak. The model was implemented using a multi-layer neuron network with back-propagation. The results demonstrated the advantages of neural networks for predicting influenza epidemics with high accuracy.

The remainder of the paper is structured as follows. Section 2 introduces the concept of artificial neural networks (ANNs). Section 3 presents the methodology of the prediction model implementation. Results and discussions are presented in Section 4. In Section 5, related considerations of the use of ANNs in influenza medical diagnosis are elaborated. Finally, conclusions and suggestions for future research are presented in Section 6.

2. Artificial Neural Networks

An artificial neural network (ANN) [7, 8] is a connected network that simulates human brain functions. ANNs have recently become the focus of much attention, especially because of their wide range of applicability and the ease with which they can treat complicated problems. They are useful

tools for computational modeling, especially when the data relationship is not identified. ANNs can learn related patterns between input data sets and corresponding output values through a training process. After training, ANNs can then be used to forecast results from new input data. The working process of ANNs follows the learning patterns of the human brain. ANNs can process problems involving non-linear and complex data, even if the data are imprecise and noisy. An ANN consists of simple interconnected computational units called *neurons*, inspired by observed processes in natural networks of biological neurons in the human brain. The key feature of these interconnected networks is their adaptive characteristics, where learning by example is properly applied in problem solving. This prominent feature makes ANNs very appealing in application domains where the problem has little or incomplete understanding to be solved, but where training data sets of the problem are readily available. ANNs are presently recognized in the area of classification and prediction, where other related statistical techniques are traditionally practicable.

2.1. Multilayer perceptron neural network

A neural network processes generally the computational modeling that involving training the network with a representative data set. The network consists of a number of input and output layers. Between these layers, hidden layers consisting of some hidden nodes are presented. The number of hidden nodes and layers is empirically determined to optimize the performance of the network to obtain the best result. Multilayer perceptron neural network (MLP) of feed forward neural network using the back-propagation algorithm is commonly used in some applications. The general architecture of MLP is presented in Figure 1.

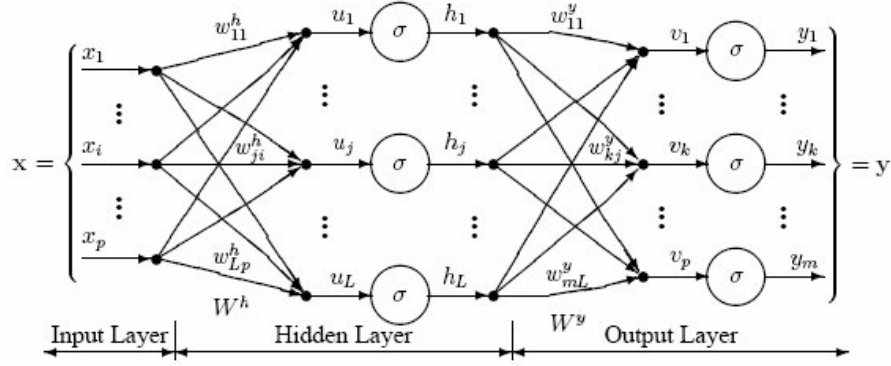


Figure 1. The multilayer perceptron neural network (MLP).

In Figure 1, u_n are nodes ordered in layer, x_n are inputs, o_n are outputs and $w_{n,n}$ are unidirectional connections with trainable weight. The MLP can be formulated as the following equation:

$$y_t = b_0 + w_{t1}x_1 + w_{t2}x_2 + \dots + w_{tm}x_m,$$

where y_t is the predicted value at t and w_{tj} is the weight related to the j th input at time t . The well-known algorithm of ANN training is back-propagation network (BPN). In back-propagation, the mean square error between calculated output and the desire value is back-propagated into the previous layer to minimize error. This is done by adjusting the node weight.

The MLP has an input layer, one or more hidden layers in the middle network and an output layer. Each layer contains at least three neurons.

- **Input layer** - The input layer accepts the input variable (x_1, \dots, x_p) of the problem and standardizes the values of each variable in the range of -1 to 1 . Then the distribution of these standardized values along with constant input called *bias* of value 1 is given to each hidden layer neuron by input layer. This bias value is then multiplied by a weight and added to the sum that is going into the neuron.
- **Hidden layer** - At each neuron in the hidden layer, a weight (w_{ji}) is multiplied to the value from each input neuron. Then a combined

value u_j is produced by adding the resulting weighted values from each hidden layer neuron. This weighted sum (u_j) is then given to the transfer function σ , producing the outputs of value h_j . The combined outputs obtained from the hidden layer neurons are then given to the neurons in the output layer.

- Output layer - At each output layer, neuron weight (w_{kj}) is multiplied to the value that is obtained from each hidden layer neuron. A combined value v_j is then produced by adding the resulting weighted values. The weighted sum (v_j) is then given to the transfer function σ , which outputs a value y_k . The y values are the outputs of the network.

3. Prediction Model Development

In this section, the methodology of the influenza prediction model is proposed. Five development processes of the model are shown in Figure 2.

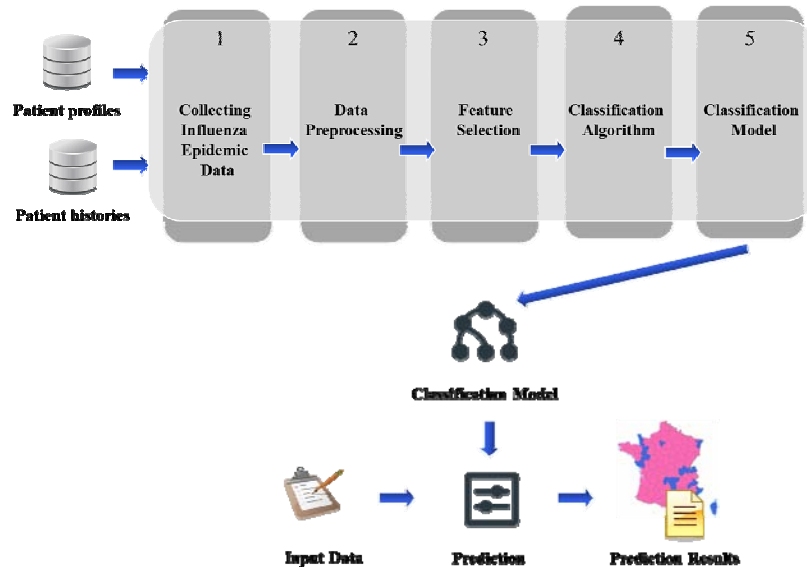


Figure 2. The development process of the proposed prediction model.

3.1. Collecting influenza epidemic data

Influenza patient data between 2007 and 2012 was obtained from the Northern Provincial Health Office, Thailand. Patients in the northern part of Thailand were chosen for this research study because the population catching flu in this area showed the highest ratio at 0.11%. The data consisted of 8,000 records, each having 13 attributes including address, age, date of sickness, blood pressure, sex, weight, and height coded as numeric values. The patient records included both males and females between one to 84 years old.

3.2. Data preprocessing

Only one kind of incomplete data was focused on for this study - missing data. Missing data can result from unexpected equipment malfunctioning or the data may not be completely recorded [9]. Data deletion without biased analysis was considered appropriate for these missing data; the probability of an observation being missing was independent of the missing data. Hence, after data preprocessing, the remaining data was composed of 7,806 records.

3.3. Feature selection

One of the most important parts in any proposed model is feature selection; this determines the appropriate attributes that contribute to the decision-making. This step requires detailed knowledge of the problem domain and the underlying data [10, 11]. Therefore, the selection of related attributes was based on the advice of doctors and also the review of the literature. Even though there were many attributes entered into the data repository of the Northern Provincial Health Office, only seven were identified as important variables that contributed to the prediction of influenza epidemics. Related attributes of the major risk factors are shown in Table 1.

Table 1. Related attributes

Field name	Data type	Values	Representation and normalization method
Latitude	Continuous numeric	90 to -90	Scale 0 to 1 (New value = positive value/maximum value)
Longitude	Continuous numeric	0 to 180	Scale 0 to 1 (New value = value/maximum value)
Month	Categorical	Jan-Dec	Jan = 1, Feb = 2, Mar = 3, ..., Dec = 12
Average age	Continuous numeric	0 to maximum value	Scale 0 to 1 (New value = value/maximum value)
Gender	Categorical	Male, Female	Male = 1 Female = 0
Moisture	Continuous numeric	0 to 100	Scale 0 to 1 (New value = value/maximum value)
Temperature	Continuous numeric	0 to maximum value	Scale 0 to 1 (New value = value/maximum value)

3.4. Classification algorithm

The proposed prediction model was implemented with ANN algorithms. The Weka toolkit was used to conduct experiments with the chosen algorithm. The Weka is an ensemble of tools for data classification, regression, clustering, association rules and visualization. The toolkit is a popular suite of machine learning software, written in Java. Weka is free, open source software issued under the GNU General Public License. The key process of this step was to train the neural networks with the selected data. Then the trained network was put through a testing phase to monitor performance. The main purpose of neural network training is to teach the network to make a good generalization about the new data set which it has never seen before. Therefore, the network should not memorize the network patterns, but instead learn the relationships among the variables involved in the network. The performance of the network is based on the testing and training results. The network with the highest testing results and the lowest training results was considered as the best network model in this exercise.

The achieved result of accuracy rates from different hidden nodes is presented in Figure 3.

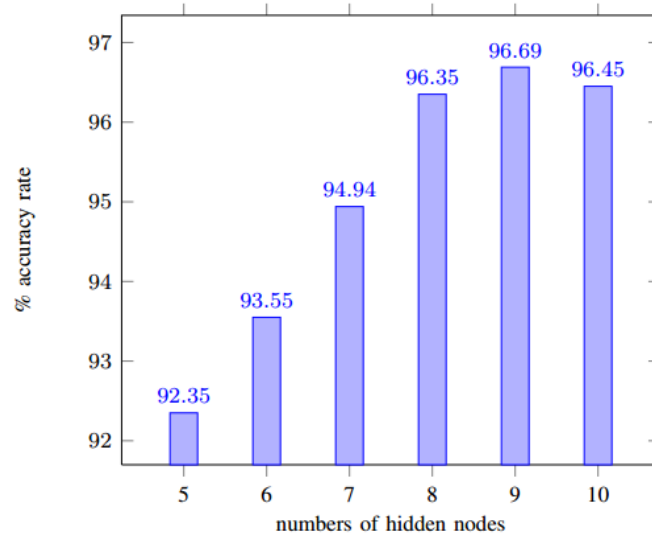


Figure 3. Accuracy rates from different hidden nodes.

3.5. Classification model

The results from the previous process determined a multi-layer perceptron with 3 layers of MLP, 7 input nodes, 9 hidden nodes and one output node as the best model. A learning rate of 0.3 and momentum rate of 0.2 were selected. The best stopping criterion was obtained at 1000 epochs. The network model performed well with a prediction accuracy of 96.69% when presented with the training data set. The Weka toolkit can calculate all these performance metrics after running a specified 10-fold cross-validation for detecting and preventing overfitting problems.

4. Results and Discussion

As stated in the previous section, the best model for predicting influenza epidemics was derived. To evaluate its accuracy and precision, the model was used to predict the influenza epidemic of 2013, using real data consisting of 1,000 records. The results of the prediction are shown in Table 2.

Table 2. The prediction results of the 2013 influenza epidemic

Class	Number of instances	% Accuracy rate
0 : not spread	357	95.79
1 : spread	643	96.42
Total	1,000	96.20

The prediction results determined the accuracy rate of predicting spreading areas (true positive) as 96.42% (342 of 357 instances), with the not spreading areas (true negative) at 95.79% (620 of 643 instances). The total accuracy result of all data sets was 96.20%. Therefore, the proposed model had good prediction accuracy and was useful for advance influenza prediction.

Three data mining techniques were also employed in this study to investigate the accuracy rate with the prediction model: random tree, random forest and tree J48 algorithms. These algorithms used the same influenza data sets for developing the models. The comparison of the results demonstrated that the ANN model was the most suitable for predicting influenza occurrences as shown in Figure 4.

The prediction model was also developed in the form of a web-based medical application. PHP and MariaDB were used as tools for implementing the application. The spatial warning system for influenza epidemics and the adjustable weight system of the model included in this application are shown in Figure 5 and Figure 6, respectively.

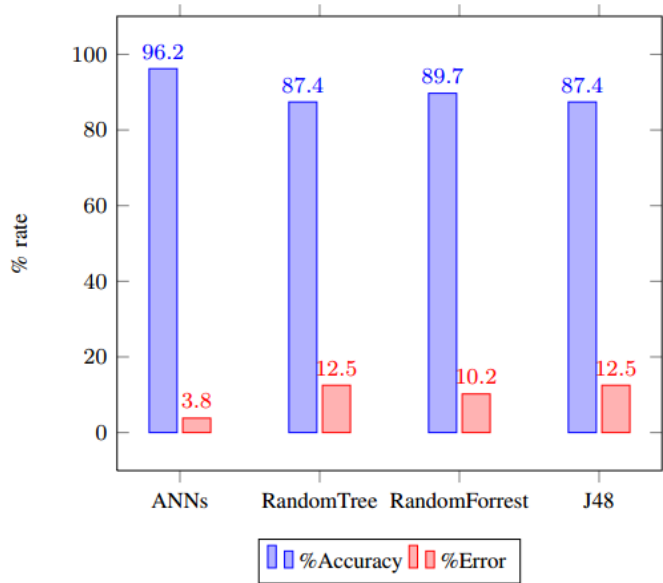


Figure 4. The comparison of results between the proposed model and others algorithms.

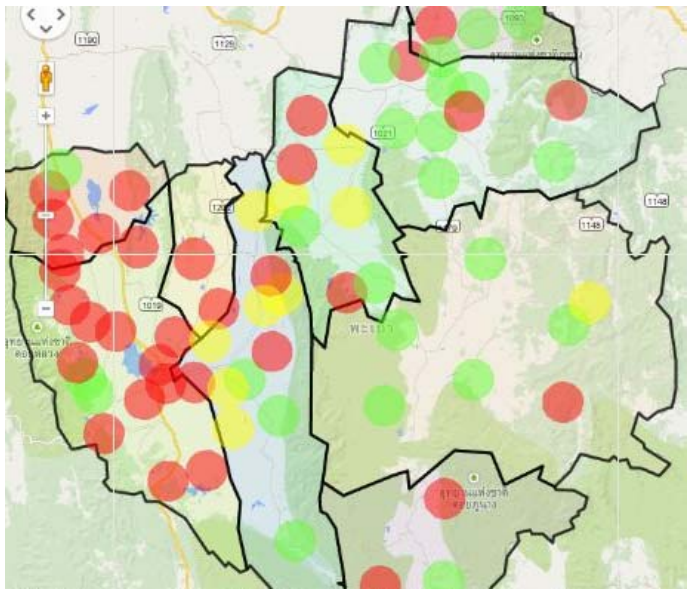


Figure 5. The spatial warning system.

NODE	w1	w2	w3	w4	w5	w6	w7	w8	w9	w10	threshold
1	-0.266	0.535	-97.865	-141.681	5.334	-39.653	112.532	-34.574	0.000	0.000	-70.482
2	-0.109	-0.199	0.992	0.868	-4.852	0.639	1.450	-0.894	0.000	0.000	-8.601
3	0.404	0.184	43.593	-11.079	41.397	54.288	15.613	91.298	0.000	0.000	-116.572
4	1.006	-0.152	-5.157	-1.682	-3.553	-1.048	0.177	0.610	0.000	0.000	-10.745
5	-1.564	0.841	0.207	0.807	-4.159	0.257	1.179	0.283	0.000	0.000	-8.257
6	-0.229	2.507	46.987	11.737	-5.107	0.308	-13.738	0.778	0.000	0.000	-43.364
7	-0.064	0.131	72.104	-46.434	-23.828	53.441	76.223	125.630	0.000	0.000	-115.602
8	0.694	0.240	-50.778	52.294	-33.241	29.993	-106.064	54.519	0.000	0.000	-110.422
9	0.045	-0.247	-7.085	0.341	5.175	-10.859	25.656	-3.844	0.000	0.000	-38.683
10	-0.540	0.470	2.234	-8.547	4.953	-8.681	-20.386	13.998	0.000	0.000	-38.385
11	-1.865	-1.439	-0.742	-0.721	-1.412	-0.719	-1.457	-1.868	-0.785	-2.122	0.755

Figure 6. The adjustable weight system.

5. Related Works

The literature survey contained several studies related to influenza medical diagnosis using data mining approaches with statistical and artificial neural networks.

Haung et al. [1] proposed a model to detect influenza transmissions. Their model was based on SinaWeibo, a Chinese micro-blogging website similar to Twitter and Facebook. They collected over 35.3 million tweets shared by all the metropolitan cities in China. Data was extracted by filtering techniques based on CDC ILI definition. They collected information related to infection centers, the city set, target city, related city and co-related city using a dynamic Bayesian network. They depicted their results of detection and transmission at city level. Culotta [12] analyzed messages posted on the micro-blogging site Twitter.com, and proposed several methods to identify influenza-related messages by comparing a number of regression models to correlate these messages with CDC statistics. Using over 500,000 messages spanning 10 weeks, the research discovered that the best model achieved a correlation of 0.78 with CDC statistics, by leveraging a document classifier to identify relevant messages.

Achrekar [13] proposed a SNEFT architecture model which contained crawler, predictor and detector components, to predict influenza activities

using information gathered from micro-blogging websites like Twitter and Facebook. In this framework, an autoregressive moving average (ARMA) model was used to predict ILI incidences. This model worked with certain accuracy. Tools mainly used in this model were ARMA model, ARX model and OSN crawler. Tweets were collected for 14 days, with 4.7 million recorded from 1.5 million unique Twitter users. The work also provided hourly and weekly basis results.

Tang and Yang [14] proposed a framework used to quantify users affected by influenza (swine flu) within a social networking community. They introduced a UserRank algorithm that incorporated link structure, content similarity, responding order and time of replies. They tested swine flu forums which were of small size, and had only 12 authorized members, each with 15.6 friends on average. They tested 90 threads and gave their results with a 100% precision rate.

Fujita et al. [15] considered the global dynamics of the epidemiological model for swine-human influenza. They incorporated the vaccination policy and anti-influenza medicine treatment into the model and analyzed their effects on the epidemic. The model included the two kinds of infected humans (infected with vaccine-sensitive strain or vaccine-resistant strain). Results paradoxically showed that the total number of the infected humans always increased monotonically if the prevalence rate of vaccination was increased. Blanco-Cocom and Vila-Vales [16] applied the Adomian decomposition method to obtain an analytical expression for the solution of a mathematical model for the transmission dynamics of influenza (type A). They used this analytical-numerical technique in order to produce simulations with different initial conditions, parameter values for different values of the basic reproduction number.

Recently, much research has been conducted on the investigation of influenza epidemics using social communities like Twitter and Facebook to collect real time data. Many data mining models have been constructed. However, there are still some limitations, for example lack of real time

information that can be used for predicting and detecting influenza epidemics.

6. Conclusions and Future Work

Influenza is one of the most common diseases that affects all ages of the population. In recent years, the number of cases and deaths has grown significantly as life expectancy has increased. Influenza is identified by physical symptoms including headache, muscle stiffness, cough, runny nose and fatigue. People normally recover within a week without requiring any medical attention. However, influenza can cause severe illness or death, especially in citizens at high risk. Thus, this study aimed to develop a performance model of ANNs for influenza epidemic prediction. The architecture model was a multi-layer perceptron with three layers constructed from 7,806 records of patient data. The proposed model gave a prediction accuracy of 96.20%. The statistical tests and simulations showed that the MLP model produced a superior performance using the test data compared with the other algorithms, random tree, random forest and tree J48. Web-based medical applications were also developed to assist related organizations regarding influenza prediction.

Future studies are planned to compare other artificial intelligence topologies to improve the accuracy rate of the model. Also, a greater variation of selection techniques will be used to further optimize the results. This will help reduce the advice times of doctors to perform early investigations of related attribute selection.

References

- [1] J. Huang, H. Zhao and J. Zhang, Detecting flu transmission by social sensor in China, IEEE International Conference on Green Computing and Communications (GreenCom), 2013 IEEE and Internet of Things (iThings/CPSCoM) and IEEE Cyber, Physical and Social Computing, Aug. 2013, pp. 1242-1247.
- [2] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis and D. I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Computational and Structural Biotechnology Journal 13 (2015), 8-17.

- [3] L. Dantas and M. Valenca, Using neural networks in the identification of signatures for prediction of alzheimer's disease, 26th IEEE International Conference on Tools with Artificial Intelligence, ICTAI 2014, Limassol, Cyprus, November 10-12, 2014, pp. 238-242.
- [4] L. Trtica-Majnaric, M. Zekic-Susac, N. Sarlija and B. Vitale, Prediction of influenza vaccination outcome by neural networks and logistic regression, *Journal of Biomedical Informatics* 43(5) (2010), 774-781.
- [5] Z. Kou, Y. Zhou and X. Qiang, Human influenza: a virus classification using a probabilistic neural network, *Kybernetes* 37(9/10) (2008), 1425-1430.
- [6] P. Attaluri, Z. Chen and G. Lu, Applying neural networks to classify influenza virus antigenic types and hosts, 2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), May 2010, pp. 1-6.
- [7] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed., Pearson Education, 2003.
- [8] J. Sonawane and D. Patil, Prediction of heart disease using multilayer perceptron neural network, 2014 International Conference on Information Communication and Embedded Systems (ICICES), Feb. 2014, pp. 1-6.
- [9] P. Ritthipravat, Artificial neural networks in cancer recurrence prediction, *International Conference on Computer Engineering and Technology*, Vol. 2, 2009, pp. 103-107.
- [10] A. J. Schroeder, Data mining with neural networks: solving business problems from application development to decision support, by Joseph P. Bigus, *JASIS* 48(9) (1997), 862-863.
- [11] A. R. Razavi, H. Gill, H. Hlfeldt and N. Shahsavar, A data pre-processing method to increase efficiency and accuracy in data mining, 10th Conference on Artificial Intelligence in Medicine, AIME2005 - Aberdeen, UK, ser. *Lecture Notes in Computer Science*, Vol. 3581, 2005, pp. 434-443.
- [12] A. Culotta, Towards detecting influenza epidemics by analyzing twitter messages, *Proceedings of the First Workshop on Social Media Analytics*, ser. SOMA'10, ACM, New York, NY, USA, 2010, pp. 115-122.
- [13] H. Achrekar, Online social network flu tracker: a novel sensory approach to predict flu trends, Ph.D. dissertation, 2013.

- [14] X. Tang and C. Yang, Identifying influential users in an online healthcare social network, 2010 IEEE International Conference on Intelligence and Security Informatics (ISI), May 2010, pp. 43-48.
- [15] N. Fujita, R. Miyazaki and Y. Takeuchi, Mathematical analysis of swine-human influenza epidemic models, Far East J. Appl. Math. 67(2) (2012), 75-114.
- [16] L. D. Blanco-Cocom and E. J. Vila-Vales, The use of the Adomian method for a sirc influenza model, Advances in Differential Equations and Control Processes 5(2) (2010), 115-127.